Lecture Notes in Production Engineering

Bernd-Arno Behrens
Alexander Brosius
Wolfgang Hintze
Steffen Ihlenfeldt
Jens Peter Wulfsberg *Editors*

Production at the leading edge of technology

Proceedings of the 10th Congress of the German Academic Association for Production Technology (WGP), Dresden, 23–24 September 2020

WGP Wissenschaftliche Gesellschaft für Produktionstechnik

Springer

Lecture Notes in Production Engineering

Lecture Notes in Production Engineering (LNPE) is a new book series that reports the latest research and developments in Production Engineering, comprising:

- Biomanufacturing
- Control and Management of Processes
- Cutting and Forming
- Design
- Life Cycle Engineering
- Machines and Systems
- Optimization
- Precision Engineering and Metrology
- Surfaces

LNPE publishes authored conference proceedings, contributed volumes and authored monographs that present cutting-edge research information as well as new perspectives on classical fields, while maintaining Springer's high standards of excellence. Also considered for publication are lecture notes and other related material of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of production engineering. The target audience of LNPE consists of advanced level students, researchers, as well as industry professionals working at the forefront of their fields. Much like Springer's other Lecture Notes series, LNPE will be distributed through Springer's print and electronic publishing channels. To submit a proposal or request further information please contact Anthony Doyle, Publishing Editor (anthony.doyle@springer.com).

More information about this series at http://www.springer.com/series/10642

Bernd-Arno Behrens · Alexander Brosius
Wolfgang Hintze · Steffen Ihlenfeldt
Jens Peter Wulfsberg
Editors

Production at the leading edge of technology

Proceedings of the 10th Congress of the German Academic Association for Production Technology (WGP), Dresden, 23–24 September 2020

Springer

Editors
Bernd-Arno Behrens
Institute of Forming Technology
and Machines
Garbsen, Germany

Alexander Brosius
Institute of Manufacturing Science
and Engineering
Dresden, Germany

Wolfgang Hintze
Institute of Production Management
and Technology
Hamburg, Germany

Steffen Ihlenfeldt
Institute of Mechatronic Engineering
Dresden, Germany

Jens Peter Wulfsberg
Laboratorium for Manufacturing
Technology
Hamburg, Germany

ISSN 2194-0525 ISSN 2194-0533 (electronic)
Lecture Notes in Production Engineering
ISBN 978-3-662-62137-0 ISBN 978-3-662-62138-7 (eBook)
https://doi.org/10.1007/978-3-662-62138-7

Editorial contact: Alexander Grün
This Springer imprint is published by the registered company Springer-Verlag GmbH, DE part of Springer Nature.
The registered company address is: Heidelberger Platz 3, 14197 Berlin, Germany

Preface

In 2020, the annual congress of the German Academic Association for Production Technology (WGP) will be held as a Webinar from September 23th to 24th under the slogan “Production at its limits – shaping change through innovation”. The WGP is hosting its annual congress for the 10th time in a row.

On behalf of the WGP, the organizing institutes are looking forward to exciting discussions with experts from research.

Production research permanently shifts the boundaries of what is feasible. Under the slogan “Production at its limits”, the contributions show production processes that advance into new areas in terms of methodology, use of resources or interdisciplinary.

But where does the search for new borders lead to? Which borders do we still have to cross, which ones do we prefer not to cross?

The focus of the congress is on production processes in border areas related to extreme velocity, size, accuracy, methodology, use of resources and interdisciplinarity. Challenges from the fields of forming machines and processes, cutting machines and processes, additive processes, automated assembly and robotics, machine learning and management sciences will be addressed.

The conference transcript summarizes the contributions from production science and industrial research. They provide the readership with an overview of current trends in production research and give an insight into ongoing research by the German Academic Association for Production Technology.

We wish all participants an interesting and inspiring WGP annual congress.

September 2020

Prof. B.-A. Behrens

Prof. A. Brosius

Prof. W. Hintze

Prof. S. Ihlenfeldt

Prof. J. P. Wulfsberg

Organization

Gottfried Wilhelm Leibniz Universität Hannover
Institut für Umformtechnik und Umformmaschinen
Prof. Dr.-Ing. Bernd-Arno Behrens
Norman Heimes, M. Sc.
Dipl.-Ing. Daniel Rosenbusch

Technische Universität Dresden
Institut für Fertigungstechnik
Prof. Dr.-Ing. Alexander Brosius
Dipl.-Ing. Sebastian Langula

Technische Universität Hamburg
Institut für Produktionsmanagement und -technik
Prof. Dr.-Ing. Wolfgang Hintze
Dr.-Ing. Carsten Möller

Technische Universität Dresden
Institut für Mechatronischen Maschinenbau
Prof. Dr.-Ing. Steffen Ihlenfeldt
Friederike Edner, M. A.

Helmut-Schmidt-Universität – Universität der Bundeswehr Hamburg
Laboratorium Fertigungstechnik
Prof. Dr.-Ing. Jens Peter Wulfsberg
Lennart Hildebrandt, M. Sc.

Contents

Cutting Machine Tools and Manufacturing Processes

Management Sciences

Forming Machine Tools and Manufacturing Processes

Experimental Characterisation of Tool Hardness Evolution Under Consideration of Process Relevant Cyclic Thermal and Mechanical Loading During Industrial Forging

F. Müller(✉), I. Malik, H. Wester, and B.-A. Behrens

Institute of Forming Technology and Machines, An der Universität 2, 30823 Garbsen, Germany
f.mueller@ifum.uni-hannover.de

Abstract. The near-surface layer of forging tools is repeatedly exposed to high thermal and mechanical loading during industrial use. For the assessment of wear resistance of tool steels, in previous work thermal cyclic loading tests were carried out to investigate changes in hardness. However, actual results of time-temperature-austenitisation (TTA) tests with mechanical stress superposition demonstrated a distinct reduction of the austenitisation start temperature indicating a change in the occurence of tempering and martensitc re-hardening effects during forging. Therefore, the superposition of a mechanical compression stress to the thermal cyclic loading experiments is of high interest. Tests are carried out in this study to analyse hardness evolution of the tool steel H11 (1.2343) under consideration of forging process conditions. The results show that the application of compression stresses on the specimen during the temperature cycles is able to restrict tempering effects while increasing the amount of martensitic re-hardening.

Keywords: Forging · Tool hardness · Phase transformation · Wear estimation · Martensitic re-hardening · Tempering

1 Introduction

High workpiece temperatures of up to 1250 °C during forging steel result in excessive heating of the surface layer of the forging tools [1]. Numerous investigations show that high surface temperatures in combination with strong cooling due to spray cooling lead to a structural change in the tool surface layer [2]. By this means, microstructural changes are caused leading to tool hardness changes depending mainly on the tool alloy, the maximum tool temperature and the cooling conditions [3] increasing the risk of tool failure or tool deformations [4]. However, recent studies have also proven, that mechanical stress strongly influences the austenitisation-behaviour of

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 3–12, 2021.
https://doi.org/10.1007/978-3-662-62138-7_1

hot work steels enabling the occurrence of martensitic re-hardening [5]. In general, hardness-changes have a decisive influence on the wear behaviour and thus on the tool life [6]. In each forging cycle the tools are exposed to a combination of thermal and mechanical stresses [7]. Statistical investigations on forging dies show that the main cause of failure of forging tools is due to approx. 70% abrasive wear and approx. 25% mechanical cracking [8]. In industrial practice, the type of damage is strongly dependent on the existing stress collectives. For example, increased wear appears due to thermally induced micro-cracks and abrasion in the tool surface [9]. The growth of tool wear also leads to geometric deviations and a reduction in component quality, which contradicts the demand for near-net-shape production and consistent product quality. In case of a significant wear progress or a tool breakage, high setup costs are incurred in addition to the costly production of new tools. Therefore, reliable information about the expected tool life is necessary for economical process control and the scheduling of set-up times. Moreover, for the design of wear-optimised tools a realistic prediction of the expected tool wear as a function of the forging cycles is required. In addition to the work of Klassen et al. [5], time-temperature-austenitisation (TTA) tests with mechanical stress superposition were carried out by Behrens et al. [10]. By varying the compression load between 30% and 80% of the elastic limit k_{f0} determined at 900 °C of hardened H11 tool steel, a distinct reduction of the Ac_1 temperature of approx. 40 °C was detected as shown in Fig. 1-A for every heating rate tested. The Ac_1 temperature of a tool steel is of high interest for its wear behaviour because at this temperature the phase transformation to austenite starts, which is retransformed to even harder martensite during the tool cooling [11]. In the context of this study, this effect is referred to as (martensitic) re-hardening. An exemplary application of the data from Behrens et al. [10] via an UPSTNO subroutine in the finite element software Simufact.forming 16 is presented in Fig. 1-B, indicating the area of a forging tool where the Ac_1 temperature is exceeded during the forging process. Including the consideration of mechanical stress on the Ac_1 temperature, the area, where martensitic re-hardening effects are expected ($T_{process} > Ac_1$), is significantly increased resulting in a different expected wear behaviour.

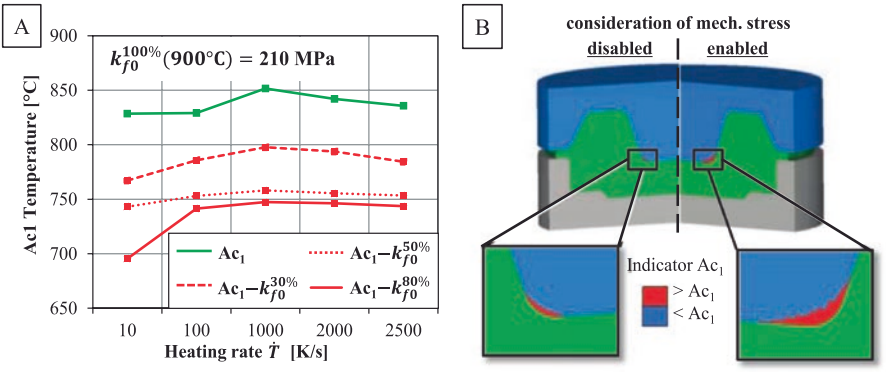


Fig. 1. Results of TTA tests with mechanical stress superposition on H11 tool steel [10] (A)/ Exemplary influence on the size of the of the martensitic re-hardening zone (red) with and without consideration of mechanical stress (B)

As a consequence, further tests are carried out in this study presenting the results of a cyclic thermo-mechanical loading to H11 tool steel. Peak temperatures are varied in regard to Ac_1 and mechanical loading in regard to the elastic limit of the material to test the influence of the parameters on tempering effects and martensitic re-hardening.

2 Methodology

For carrying out cyclic loading tests a forming dilatometer DIL805D by TA Instruments is used equipped with SiO_2 deformation punches (Fig. 2-A). Since investigations on wear-related topics are fundamentally about saving costs, the necessary testing procedure is also strongly connected to an evaluation of testing costs. Therefore, hollow samples are of high interest in order to not only to be able to achieve process relevant heating and cooling rates but also to minimise the amount of nitrogen cooling gas. This sample type is characterised by an increased specimen surface in comparison to the conventional specimen made of bulk material. Because of this decision, two main circumstances using the DIL805D had to be addressed:

1. The primary use case of the deformation unit for the DIL805 is the evaluation of mechanical properties using cylindrical bulk samples (Ø 5 mm x 10 mm). The default hydraulic force control parameters are therefore optimised for this sample type. Using hollow samples with these parameters results in high force deviations as shown in Fig. 2-B, especially during fast heating or quenching segments, where the sample length is rapidly changing. As a consequence the controller sensitivity has to be significantly increased by reducing the proportional value to xp = 0.007. This system value is the most influential parameter for the determination of the control strength in proportional relation to the systems inherent control power after a control deviation is measured. In this case, less power of the hydraulic pressure pump for the regulation of the punch force has to be engaged to accommodate for the reduced specimen cross section. This change reduces force deviations during fast temperature changes to less than 10% of the specified value while allowing heating rates of up to 600 K/s.

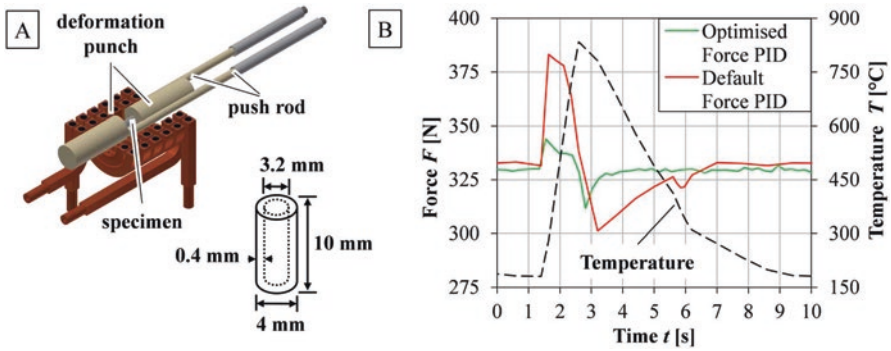


Fig. 2. Dilatometer DIL805D test apparatus and hollow specimen geometry (A)/Optimisation of force PID control parameters for hollow samples (B)

2. The DIL805D default programming capabilities are limited to a certain amount of test segments. Therefore, the DIL control software was extended by TA Instruments with a custom cycle generator module enabling the application of continuous thermal cycles with a constant mechanical stress superposition.

With this test-setup prepared, two types of tests were carried out in regard to the respective test matrices presented in Tables 1 and 2. At first, an extended cyclic re-hardening test is performed by applying sets of 25 thermo-mechanical load cycles with peak temperatures from 800 °C to 900 °C. Mechanical stress is superimposed with three levels in regard to the elastic limit k_{f0} of H11 tool steel determined at 900 °C. The aim of the test is to identify the lowest peak temperature where re-hardening effects can be observed by an increase in hardness. Also, this test is also used to investigate the relationship between the austenitisation behaviour characterised by TTA tests and the wear-relevant hardness. While the dilatometric TTA test used by Behrens et al. [10] is based on tactile measurements to identify phase transformation on a micrometre scale, the hardness evaluation features an optical measurement of indents for the determination of the hardness value. Therefore it must be assumed, that the detection resolution with this procedure is reduced, leading to the assumption that the measurable minimum temperature at which re-hardening occurs is higher compared to the TTA tests.

Table 1. Test matrix for the re-hardening study

Stress superposition [% k_{f0}]	Temperature range [°C]	Temperature Increment [°C]	Cycles	Repetitions
0	800–900	20	25	3
30				
50				
80				

Afterwards, thermo-mechanical loading tests with high cycle counts up to 2000 are carried out to estimate the effects during industrial use. Regarding peak temperatures 600 °C, 750 °C and 900 °C are used to ensure the formation of re-hardening as well as tempering effects. The thermal cycle profile using a heating rate of 500 K/s and the application of the mechanical stress superposition is identical in both parts of this study. Keeping a thermal cycle time of about eight seconds in mind, the repetition number had to be reduced to two because of the high testing time of over two hours per 1000 cycles.

Table 2. Test matrix for the high cycle loading tests

Peak temperatures [°C]	Stress superposition [%k_{f0}]	Investigated cycle numbers	Repetitions
600	0, 30, 50, 80	1, 10, 50, 100, 500, 1000, 2000	2
750			
900			

All tested samples are metallographically prepared for micro-hardness measurement at nine measuring points across the sample length as shown in Fig. 3-A. For this purpose, the samples are first cast in epoxy resin and subsequently wet grinded in several steps with SiC grinding paper ranging from a 220 to a 1200 grid. Afterwards the samples are polished three times using diamond suspension with an abrasive grain diameter of 6, 3 and 1 µm. Both operations are carried out on a Tegramin-30 sample preparation device by Struers. The embedded specimens are then etched with 5% nital acid for light microscopical images of the microstructure. For the micro-hardness measurement the standardised measuring method according to Vickers with a test load of 1.961 N (corresponds to HV0.2) is used.

3 Results and Discussion

3.1 Pretesting

In Fig. 3-B an exemplary overview of the microstructure after 10 loading cycles at 900 °C with 80% k_{f0} is presented showing a characteristic transition from the outer area of the specimen to the center. Because of heat transfer between the sample and the deformation punch leading to lower peak temperatures, the outer area is dominated by tempered ferrite with remaining martensite plates featuring a reduced hardness of 380 HV. The centre of the specimen, where the testing temperature is ensured, only consists of a fine-grained structure, which can be referred to as re-hardened martensite. Hardness in this area is significantly increased to 650 HV compared to the base hardness of 450 HV.

In this study hardness was only evaluated in the middle area of the specimen close to the central welding location of the thermocouple placed at evaluation point 5. This is achieved by statistically averaging the hardness values of the measuring points from position 4 to 6 while also calculating the standard deviation in this area to assess possible fluctuations in hardness.

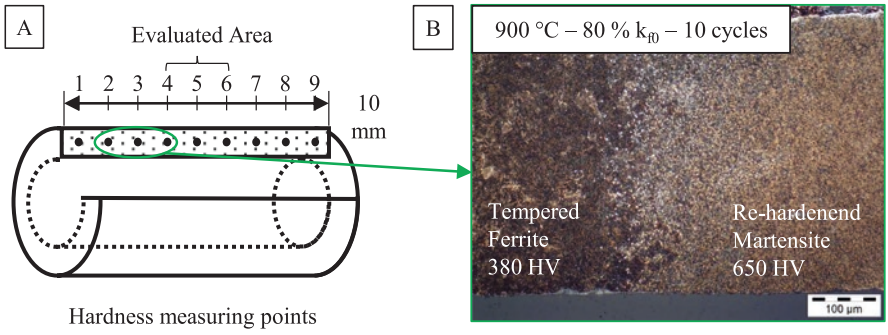


Fig. 3. Hardness measuring locations (A) and microstructural image of the transition area representing the hardness measuring locations from 2 to 4 showing tempered ferrite and re-hardenend martensite after thermo-mechanical loading (B)

3.2 Re-hardening Study

As assumed in Sect. 2, the results of the re-hardening study, plotted in Fig. 4, does not show a clear conformity to the measured Ac_1 temperatures via TTA testing. While without stress superposition (0% k_{f0}) re-hardening can initially be observed at temperatures over 880 °C, the superposition with 80% of the elastic limit k_{f0} will activate this effect already at temperatures of about 850 °C. In agreement with the results of Fig. 1-A the magnitude of the impact due to the mechanical stress level decreases with increased loading. While the measured data agrees with the finding that higher mechanical load leads to lower re-hardening start temperatures in theory, the difference between all results of the mechanical stress superposition tests are relatively minor in practice. The calculated standard deviation of ±15 HV are explained with the inherent measurement inaccuracy of the optical Vickers method. An exception is found in the hardness values of the 50% stress superposition series where the standard deviation values are significantly increased (approx. ±45 HV). The reason for this was found in the evaluation area defined in Fig. 3-A. While in the other test series at this zone either re-hardened or tempered microstructure was found exclusively, in the prominent test series a transitional microstructure comparable to Fig. 3-B was found. A possible explanation for this are slight offsets on the thermocouple welding location or a slight non-concentric placement of the sample in regard to the deformation punches of the dilatometer leading to deviations of the temperature field applied. Keeping this in mind, it must be concluded that the superposition with mechanical stress levels over 30% of k_{f0} leads to no distinct difference on the occurrence of re-hardening effects compared to each other.

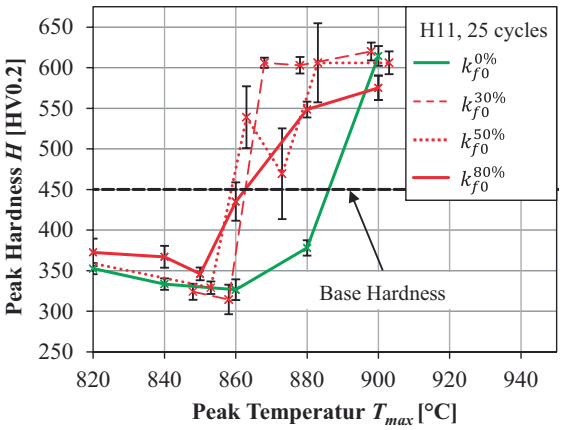


Fig. 4. Hardness over peak temperature for H11 after 25 thermo-mechanical loading cycles, heating rate: 500 K/s

3.3 High Cycle Loading

Because of the findings of the re-hardening study, only the results with mechanical loading of 80% k_{f0} and without additional loading are shown in Fig. 5 for the evaluation of the thermo-mechanical loading test at higher cycle counts and for better comprehension. In general, the results of measurements at 30% and 50% k_{f0}, which are not shown in this picture, are nearly identical to the results depicted below for 80% k_{f0}.

Regarding the three tested peak temperatures, three individual findings can be identified: After the loading at a peak temperature of 600 °C no measurable change in hardness could be observed under consideration of a minor standard deviation of less than 5 HV both with mechanical stress superposition and without. This finding indicates that the superposition applied has no influence on the material specific activation temperature for the occurrence of tempering effects.

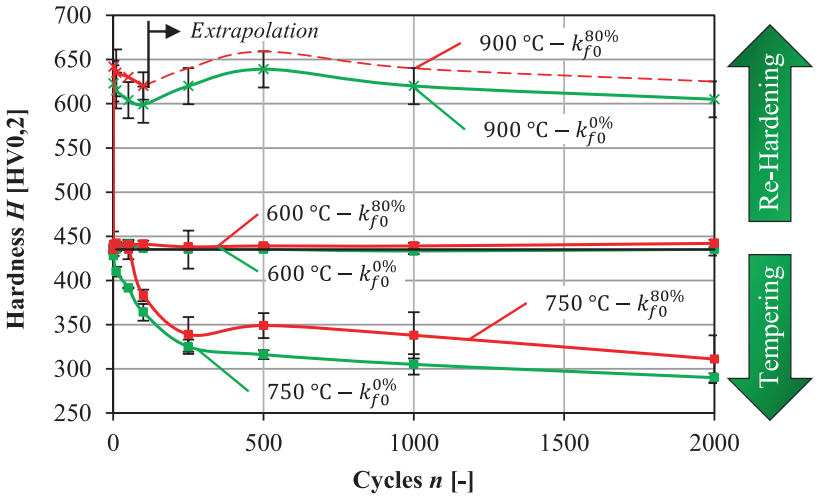


Fig. 5. Hardness results of the thermo-mechanical loading tests on H11 with cycle counts on a process relevant scale

At a peak temperature of 750 °C and up to 250 testing cycles no significant differences between all mechanical loading scenarios can be observed either. During all tests the hardness is reduced from 450 HV down to about 340 HV indicating tempering effects. However, after 500 cycles the reduction of hardness is slowed down by the application of mechanical stress superposition leading to a remaining hardness delta of approximately 30 HV over the full testing cycle spectrum. This finding is mainly explained by the diffusion properties of forcefully resolved carbon in the ordered martensitic matrix. At first, because of the high concentration gradient between matrix and microstructure, carbon can be transferred regardless of the overlaying mechanical stress leading to the identical drop in hardness. However, after the concentration compensation reached a critical point, the superposition with mechanical stress leads

to additional restraint on the martensitic structure slowing down the ongoing diffusion process. The standard deviation of the hardness values for this test series were also found to be in an acceptable range of less than ±10 HV indicating a homogenous microstructural distribution over the evaluation area of all related samples.

At a peak temperature of 900 °C, an immediate increase of the base hardness from 450 HV to over 600 HV is measured after all loading scenarios. During all tests at 900 °C with mechanical stress superposition a steady decrease of specimen length (about 4 µm per cycle) was also observed leading to severe deformation as shown in Fig. 6-A. To prevent a collapse of the specimen, which was found to be happen after a length decrease of about 500 µm, all tests with mechanical stress superposition had to be stopped after 100 loading cycles. Since the microstructure in the deformed area presented in Fig. 6-B is dominated by re-hardened martensite, as can be derived from Fig. 3-B, transformation-induced plasticity is determined as the reason for the sample deformation. This effect describes the occurrence of plastic deformations as a result of elastic mechanical stress during a phase transformation of the microstructure from austenite to martensite [12]. In the case of this study the transformation from martensite to the smaller austenitic structure during heating leads to a reduction of the sample length which is amplified by the mechanical load. During cooling the size increase in length direction due to the retransformation to martensite is also blocked by the deformation punches. Both effects combined cause an incremental reduction of the specimen length during each loading cycle. Still, up to a cycle count of 100, the measured hardness values were approx. 30 HV higher than the measurements with no external mechanical stress applied. However, because of the slightly increased standard deviation of both testing series of about ±15 HV, the influence of the superposed mech. loading is found to be minor in regard to the absolute achievable peak hardness. Nevertheless, the results of the test series with superposed mech. stress were extrapolated in accordance to the unloaded test results by adding a constant offset value of 30 HV to obtain a full data set for upcoming numerical material modelling.

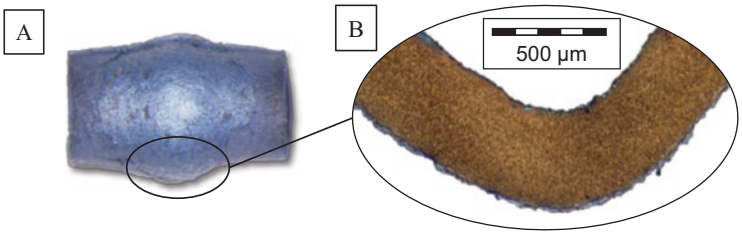


Fig. 6. Deformed specimen after 50 thermo-mechanical loading cycles at 900°C and 80% k_{f0} superposition (A) with martensitic microstructure formed by re-hardening (B)

4 Summary and Outlook

In the present study, the influence of mechanical stress superposition applied to thermal-cyclic experiments to reproduce the tool load in the surface layer while forging was investigated. In previous work, the decrease of the material characteristic Ac_1

temperature was already confirmed by dilatometric TTA tests. Now the occurrence of the associated martensitic re-hardening effects at reduced temperatures could be shown at the example of H11 tool steel. Regarding the two test series carried out in this study two main findings could be identified:

- While the TTA tests indicated a clearly measurable influence of the mech. stress level on the reduction of Ac_1, the results of the re-hardening study were only dependent on the amount of mech. stress to a lesser degree. In summary, it was found that as long as any mechanical load was applied, a significant reduction of about 20 °C to the minimum temperature necessary for re-hardening was observed.
- As a result of the high cycle loading tests, it was shown that tempering effects are influenced by an external mech. stress superposition, resulting in slower reduction in hardness. However, the maximum amount of hardness achievable due to re-hardening was found to be only marginally influenced by the application of mech. stress superposition.

These observations indicate for future work that for more precise wear estimations based on calculated process variables, the normal mechanical contact stress on the surface layer must also be taken into account. In the next step, a user subroutine for Simufact.forming 16 will be created to visualise the material data gathered in this study. Also, widely used nitride tool layers for additional wear resistance are in the focus of upcoming investigations. Since these layers represent a significant chemical modification of the surface layer, the austenitisation and the behaviour of the hardness evolution under thermo-mechanical load will be tested analogously to this study. Finally, laboratory forging tests are planned to validate the results of the material characterization and simulations.

Acknowledgements. The authors gratefully acknowledge the support of the German Research Foundation (Deutsche Forschungsgemeinschaft - DFG) and the German Federation of Industrial Research Associations (AiF) within the projects DFG 397768783 and AiF 19647 for this research work.

References

1. Jeong, D.J., Kim, D.J., Kim, J.H., Kim, B.M., Dean, T.A.: Effects of surface treatments and lubricants for warm forging die life. J. Mater. Process. Technol. **113**, 544–550 (2001)
2. Yu, Z., Kuznietsov, K., Mozgova, I., Böhm, V., Gretzki, T., Nürnberger, F., Schaper, M., Reimche, W.: Modeling the relationship between hardness and spray cooling parameters for pinion shafts using a neuro-fuzzy model strategy. J. Heat Treat. Mater. **67**(1), 39–47 (2012)
3. Caliskanoglu, D., Siller, I., Leitner, H., Jeglitsch, F., Waldhauser, W.: Thermal fatigue and softening behaviour of hot work tool steels. In: ICT Conference, Karlstad, 10.–13.09.2002, Issue 1, pp. 707–719 (2002)
4. Marumo, Y., Saiki, H., Minami, A., Sonoi, T.: Effect of the surface structure on the resistance to plastic deformation of a hot forging tool. J. Mater. Process. Technol. **113**(1–3), 22–27 (2001)
5. Klassen, A., Bouguecha, A., Behrens, B.-A.: Wear prediction for hot forging dies under consideration of structure modification in the surface layer. Adv. Mater. Res. **1018**, 341–348 (2014)

6. Kim, D.H., Lee, H.C., Kim, B.M., Kim, K.H.: Estimation of die service life against plastic deformation and wear during hot forging processes. J. Mater. Process. Technol. **166**, 372–380 (2005)
7. Bernhart, G., Brucelle, O.: Methodology for service life increase of hot forging tools. J. Mater. Process. Technol. **87**(1–3), 237–246 (1999)
8. Kannappan, A.: Wear in forging dies - A review of world experience. Metal Form. **36**(12), 335–343 (1969)
9. Persson, A., Hogmark, S., Bergström, J.: Temperature profiles and conditions for thermal fatigue cracking in brass die casting dies. J. Mater. Process. Technol. **152**, 228–236 (2004)
10. Malik, I.Y., Lorenz, U., Chugreev, A., Behrens, B.-A.: Microstructure and wear behaviour of high alloyed hot-work tool steels 1.2343 and 1.2367 under thermo-mechanical loading. In: Materials Science and Engineering, vol. 629 (2019)
11. Behrens, B.-A., Puppa, J., Acar, S., Gerstein, G., Nürnberger, F., Lorenz, U.: Development of an intelligent hot-working steel to increase the tool wear resistance. In: Tooling 2019 Conference & Exhibition, 13.05.–16.05.2019, p. 64 (2019)
12. Behrens, B.-A., Chugreev, A., Kock, C.: Macroscopic FE-Simulation of residual stresses in thermos-mechanically processed steels considering phase transformation effects. In: XIV International Conference on Computational Plasticity, Fundamentals and Applications COMPLAS (2019)

Modelling of Hybrid Parts Made of Ti-6Al-4V Sheets and Additive Manufactured Structures

J. Hafenecker[1,4(✉)], T. Papke[1,4], F. Huber[2,3,4], M. Schmidt[2,3,4], and M. Merklein[1,3,4]

[1] Institute of Manufacturing Technology (LFT), Friedrich-Alexander-Universität Erlangen-Nürnberg, Egerlandstraße 13, 91058 Erlangen, Germany
jan.hafenecker@fau.de
[2] Institute of Photonic Technologies (LPT), Friedrich-Alexander-Universität Erlangen-Nürnberg, Konrad-Zuse-Straße 3-5, 91052 Erlangen, Germany
[3] Erlangen Graduate School in Advanced Optical Technologies (SAOT), Paul-Gordan-Straße 6, 91052 Erlangen, Germany
[4] Collaborative Research Center 814 – Additive Manufacturing, Am Weichselgarten 9, 91058 Erlangen, Germany

Abstract. The current trend of mass customization pushes conventional production techniques to their limits. In the case of forming technology, limitations in terms of adaptability and flexibility emerge, while additive manufacturing lacks in the manufacturing of large, geometrically simple components. Combining both processes has potential to use the strengths of each process and thus realize time and cost efficient mass customization. As the interactions between the processes have not been fully investigated yet, in this work a distinct modelling approach in LS-DYNA is used to examine the influence of the additively manufactured elements on the formability. Namely, varying geometric properties and number of pins created with additive manufacturing are in the focus of this research. The used material is the alloy Ti-6Al-4V, which requires processing at elevated temperatures due to its low formability at room temperature. The results show a clear influence of the additively manufactured elements on the formability.

Keywords: Additive manufacturing · Forming · Titanium

1 Introduction

Mass customization describes the industrial trend to manufacture personalized products in high production volumes [1]. This trend is pushing conventional production methods to their limits [2]. One way to overcome these limits is the combination of conventional manufacturing processes with additive manufacturing (AM) [3] with its high degree of geometric freedom [4]. However, additive manufacturing shows a deficit with regard to the manufacturing time, especially for high production volumes [3]. In contrast, sheet metal forming is highly efficient in producing large quantities.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 13–22, 2021.
https://doi.org/10.1007/978-3-662-62138-7_2

Therefore, it is ideally suited for being combined with additive manufacturing for the production of hybrid components [5]. As a result, parts can be produced with less time, costs and energy compared to an additive only process [6]. A possible area of application for hybrid components are medical implants such as hip sockets. In this field, hybrid implants offer adaptability to patient-specific customizations, but also standard geometries, which can be produced with forming technology, to reduce production time and costs [7]. Instead of inefficiently producing the whole part with laser powder bed fusion (PBF-LB), only the customizable pins are built upon the sheet metal. As a result, standardized sheets with adaptable additively manufactured structured can be used in a deep drawing process. Thus, hybrid processes grant a leading edge in terms of manufacturing times of additively manufactured parts. Additionally, hybrid parts offer the adaptability conventional forming parts are not capable of. The used material is the titanium alloy Ti-6Al-4V because of its properties such as biocompatibility and specific strength [8]. However, high strength, low Young's modulus and limited plastic strain at room temperatures reduce the formability in cold forming processes [9]. At a temperature of 400 °C and higher, less energy is necessary for dynamic recrystallization and additional slip systems are activated [10]. Thus, experiments are performed at a temperature of 400 °C. Investigations on the combination of warm bending and PBF-LB to produce hybrid parts made from Ti-6Al-4V with one additively manufactured element (AE) show the manifold interactions between the two processes [11]. The combination of these two operations is influenced by the their sequence [11] and the interactions between the AM- [12] and the forming process [5]. Purpose of this research is to investigate these interactions, namely the influence of more than one AE on the formability of the hybrid parts made from Ti-6Al-4V in LS-DYNA to gain knowledge of the interactions. Since the AEs are the adaptable component of the hybrid part, different combinations have to be investigated since a strong influence on the formability is expected. With this intention, the following parameters are varied in this study:

- the influence numbers of AEs (*Num*) with their own layouts,
- different geometries (*Geo*),
- AE-diameters (*Dia*),
- distances to the middle (*Pos*),
- fillet radii (*Rad*).

2 Experimental Setup and Procedure

2.1 Material and Modelling Approach

In this research, the hybrid parts (Fig. 1 – right) consist of the two components: sheet material and additively manufactured elements, both made of Ti-6Al-4V. The sheet has a fine grained equiaxed $\alpha+\beta$-microstructure. The hexagonal close packed α-titanium has a limited formability at room temperature and causes anisotropic mechanical behaviour [9]. Without any heat treatment, the additively manufactured component consists of a martensitic α'-titanium. This granular structure is harder, stronger and less ductile compared to the equiaxed microstructure of the sheet

material. The results of tensile tests performed at 400 °C (Fig. 1 – left) show the differences and thus prove the need for different material models. In order to represent these different material properties in the numerical simulation, two different material models for sheet and AE are used. A second differentiation regarding the components of the hybrid part is made on basis of the element formulation. The sheet is modelled with shell elements and seven integration points across the thickness which is common for sheet metal forming simulation [13]. Solid elements are used for the additively produced pins due to their volumetric geometry. Both components are joined via common nodes (Fig. 2 – left) in a 3D-simulation. Due to symmetries, only a quarter of the hybrid part is modelled.

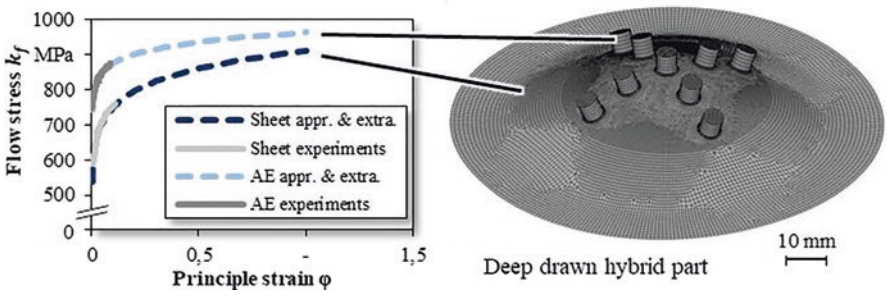


Fig. 1. Tensile flow curves of hybrid part's components at 400 °C (left); hybrid part (right)

The sheet is modelled with the LS-DYNA material model "233-CAZACU_BARLAT" that was found to represent the material behaviour precisely even at higher temperatures in [13]. This material model bases on the model of Cazacu and Barlat from 2006 [14]. However, this model is only applicable for shell elements. Thus, the material keyword "024-PIECEWISE _LINEAR_PLASTICITY", which bases on the modelling approach of von Mises, is used for the pins. The flow curves for both material models are extrapolated using the approach of Nemat-Nasser [15]. The schematic setup of the process is shown in Fig. 2 – right.

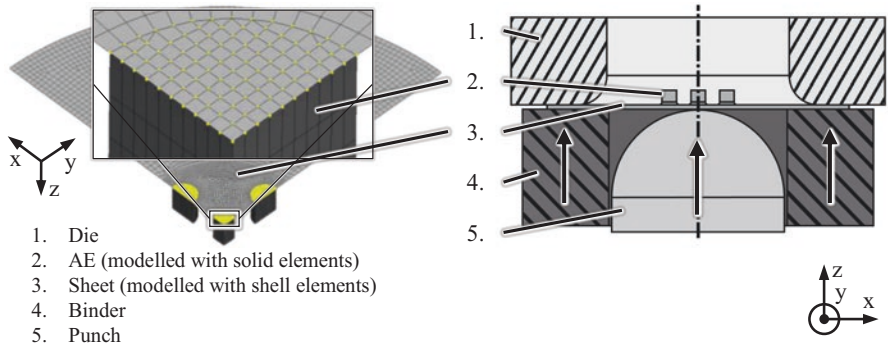


Fig. 2. Numerical representation of hybrid part with highlighted common nodes (left); schematic setup of the deep drawing process (right)

For the die, binder and punch rigid shell elements are used. The spherical punch has a diameter of 60 mm, the die clearance is 1.7 mm and the radius of the die is set to 10 mm. The sheet has a diameter of 105 mm, a thickness of 1.5 mm and the height of the pins is 5 mm. The drawing depth is set constant to 15 mm without usage of a failure criterion.

2.2 Sheet Material Modelling Validation

To assure that the process model can be used, the material model itself has to be validated first. Prior investigations used the same material model but a different sheet geometry (Fig. 3 – right), on which one AE was built after the forming process [16]. The same geometry is used for the recent material validation by thickness comparison. With this intention, the thickness distribution after the deep drawing process is calculated in the simulation and measured for real parts of a sheet without pins. The sheet thickness is measured along the x-axis in the x-z-plane. Since the numerical model only consists of a quarter of the real process, the thickness distribution is mirrored at the y-z-plane. The comparison of sheet thickness (Fig. 3) shows the high accuracy of the numerical model.

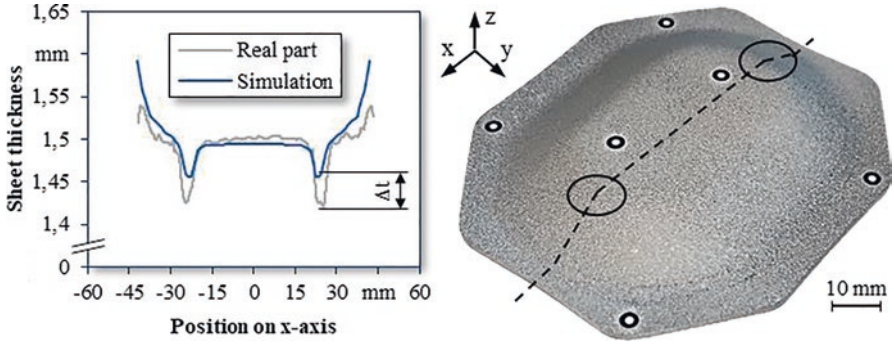


Fig. 3. Comparison of thickness distributions (left) between simulation and real part (right) along the x-axis for sheet thickness 1.5 mm; measurement path and critical spots marked (right)

Namely, maximum differences (Δt) in sheet thickness are lower than 3% of the thickness, which corresponds to 0.05 mm. Nevertheless, the critical sheet thickness reductions in the area of the punch radii are underestimated in the simulation. In particular, this is important for the spherical punch geometry where the shape is expected to lead to the highest thinning in the centre of the sheet. As this is the area where the pins are placed, the influence and resulting thickness reduction is even more critical.

2.3 Investigated Parameter Combinations

In this paper the results of the investigated influence of AE-diameter (*Dia*), AE-position (*Pos*), AE-geometry (*Geo*), number of additively build up pins (*Num*) and the size of the fillet radius at the transition (*Rad*) on the formability of hybrid parts

are presented. For a better understanding, the parameters and the different shapes and layouts are visualized in Fig. 4. The maximum sheet thickness reduction and the maximum stresses according to von Mises are calculated and evaluated in the simulations, as they are indicators for failure in terms of rupture during the forming process. Constant process parameters for all simulations are a forming temperature of 400 °C and a binder force of 20 kN. The drawing depth is set to 15 mm and the velocity of the punch to 200 mm/s. The coefficient of friction is assumed to be $\mu = 0.2$ in the contacts punch-blank, die-blank and binder-blank according to [16]. The diameter *Dia* varies between 3 mm and 7 mm, the position *Pos* between 10 mm and 20 mm and the fillet radius *Rad* between 0 mm and 1 mm. The layout of 8 and 9 pins is either circular (*Num* = 8-C, 9-C) or crosswise (*Num* = 8-X, 9-X). The reference setting is the combination of five pins (*Num* = 5), with a diameter of 5 mm, a radius of 0.5 mm, *Pos* = 10 mm and cylindrical pins.

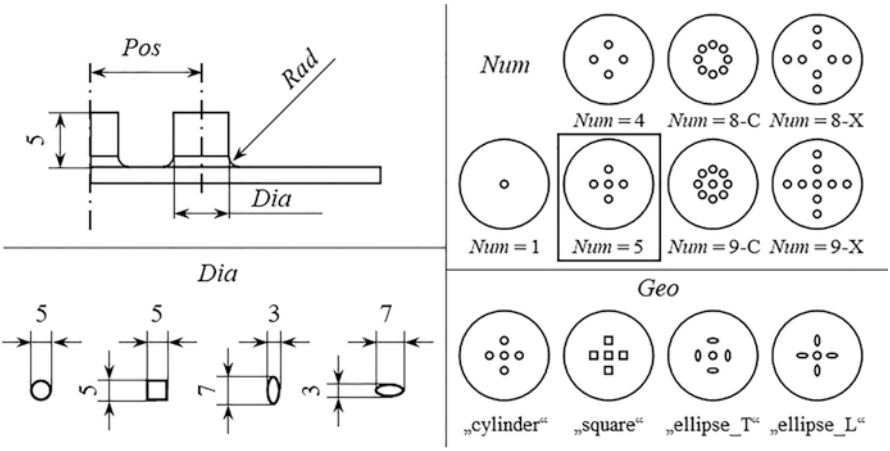


Fig. 4. Visual explanation of the investigated parameters; reference layout (*Num*) marked

3 Results and Discussion

Regarding formability the sheet thickness reduction and von Mises stresses are calculated for each simulations. Since high values in any of these outputs are indicators for failure in a real forming process, factor levels that result in low values are to be identified. The pins hinder the material flow during the forming process resulting in higher stresses and thickness reductions (Fig. 5). This effect arises at the transition area between the pins and the sheet. One reason is the geometric properties of the pins. Due to their height, the pins have a higher resistance towards forming. Thus, mainly the sheet material is formed under load.

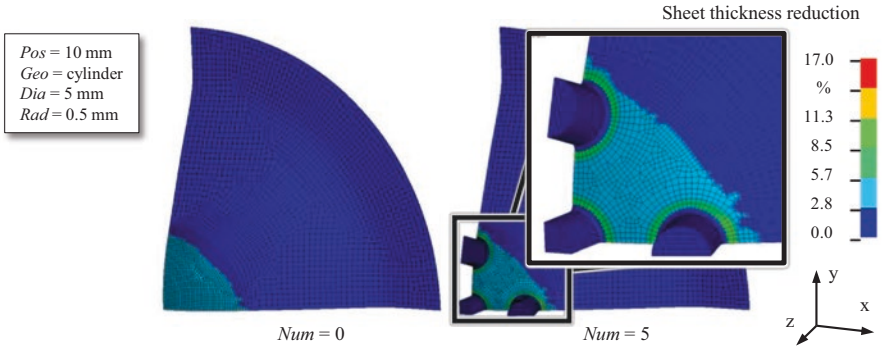


Fig. 5. Comparison of the calculated sheet thickness reduction between a sheet without pins (left) and with five pins (right)

Therefore, a lower ratio of "sheet with pin" to "sheet without pin" (Fig. 6) leads to a concentration of the forming on less material. As a result, the local stresses and resulting thickness reductions increase.

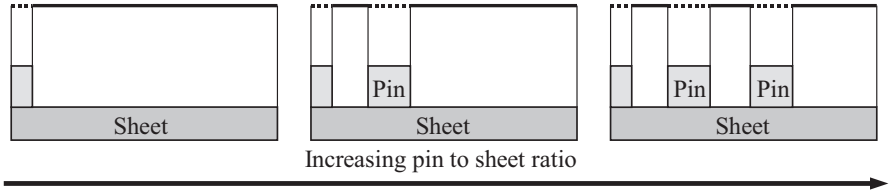


Fig. 6. Explanation of the pin to sheet ratio

Additionally, the sharp corners at the bottom of the pins have a stress concentrating effect. This can be reduced by using fillet radii (Fig. 7). With increasing radius (*Rad*), the stress concentrating effect and thus the thinning is reduced. However, the improvement for an increase from 0.5 mm to 1 mm is quite low.

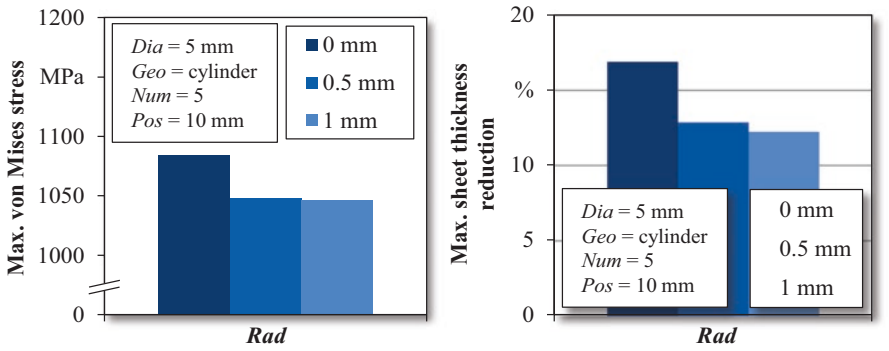


Fig. 7. Calculated stresses and thickness reductions for different fillet radii (*Rad*)

Unlike for the reference geometry “cylinder” that results in the lowest stresses, the geometry “ellipse_L” has by far the highest pin-sheet-ratio among the other geometries and results in the highest local stresses (Fig. 8). The lowest thickness reduction is found for the pin geometry “square”. Thus, the geometries “cylinder” and “square” are preferred due to their higher formability.

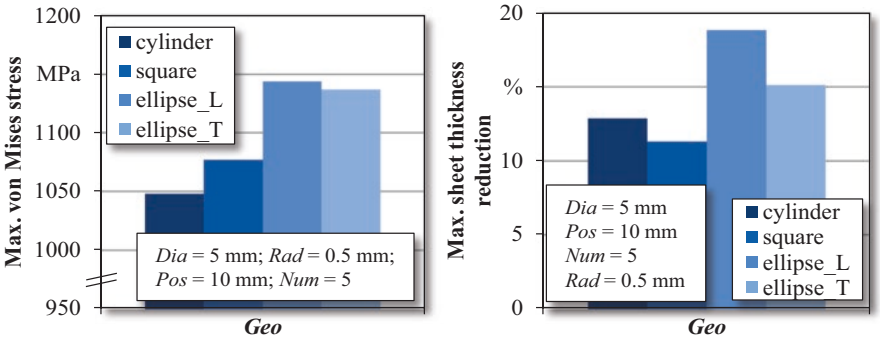


Fig. 8. Calculated stresses and thickness reductions for different pin geometries (*Geo*)

Additionally, this effect is visible for the crosswise layouts (8-X, 9-X) where again the pin-sheet-ratio is the highest compared to the other layouts (Fig. 9). The idea that this ratio matters only in a radial line is enhanced by a comparison between the results of the simulations with eight pins in a circular layout (8-C) and in a crosswise layout (8-X). Even though the amount of pins is the same, the maximum thickness reduction of the crosswise layout (8-X) is 50% higher. The circular layout has a higher formability compared to the crosswise layout.

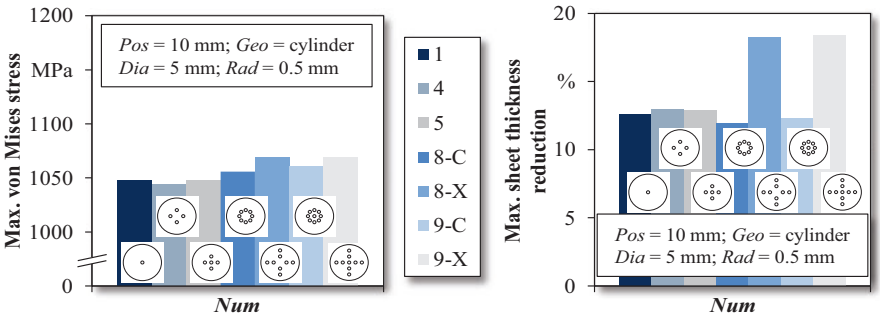


Fig. 9. Calculated stresses and thickness reductions for different AE-numbers and layouts (*Num*)

An increasing diameter (*Dia*) leads to higher stresses and thickness reductions as well (Fig. 10). Again, the increasing ratio of pin to sheet is an explanation for that. For later applications, the diameter is preferred to be as low as possible. The lower thickness reduction for *Dia* = 5 mm compared to 4 mm is so far inconclusive. However,

simulations with diameters of 4.9 mm and 5.1 mm lead to similar results as for 5 mm. Thus, this range needs to be investigated further.

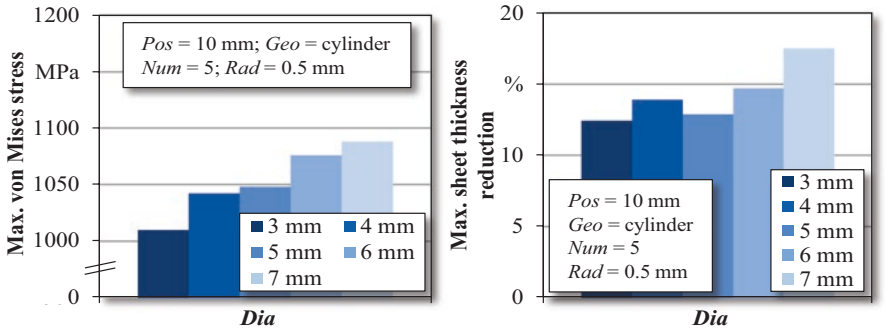


Fig. 10. Calculated stresses and thickness reductions for different pin diameters (*Dia*)

In contrast, the distance of the AEs to the centre (*Pos*) is almost not of concern for the material flow (Fig. 11). No clear influence is visible within the investigated parameters. Regardless of their position, the pins are placed in the forming zone, which is why the pin to sheet ratio has the highest effect on the formability.

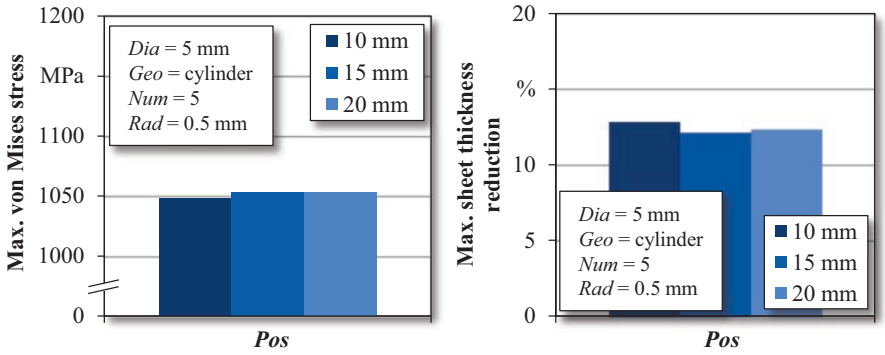


Fig. 11. Calculated stresses and thickness reductions for different pin positions (*Pos*)

4 Summary and Outlook

In conclusion, a new approach to manufacturing adaptable hybrid parts with less costs and time was investigated using numerical simulation. The used material modelling approach for the sheet material was validated and the numerical simulation was used to investigate the influence of additively manufactured elements (AE) on the formability of hybrid parts. The volumetric geometry of the pins leads to a higher resistance towards forming compared to the sheet. Thus latter mainly compensated the loadings.

It was shown that with increasing ratio of pin to sheet the formability of the hybrid parts was decreasing. With increasing AE-diameter (*Dia*) the formability decreases as well as for increasing number of pins (*Num*) in a radial line. The pin-sheet-ratio increases for both occasions and thus the stresses are distributed over a reduced proportion of sheet metal, which increases the loadings for the remaining sheet. Additionally, the transition zone between the pin and the sheet metal is the most critical part because of the stress concentrating effect causing higher thickness reduction in this area. The fillet radius lowers the stress concentrating effect of the transition yet does not remove it completely. However, stresses and thinning are reduced and thus formability increased. The AE-position (*Pos*) has the lowest effect on the formability since the pin-sheet-ratio stays constant. Future work will include the influence of a heat treatment before the forming process to smoothen the rather sharp gradient of mechanical properties between the two components. The interactions between the factors is another possibility for investigations. Finally, the gained knowledge will be applied in physical experiments to identify a process window and the maximal drawing depth for different parameter combinations. Thus, it is possible to overcome the limiting process times of additive manufacturing and to produce adaptable hybrid parts.

Acknowledgements. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project-ID 61375930 – SFB 814 - “Additive Manufacturing” TP B05.

References

1. Albers, S., Herrmann, A.: Handbuch Produktmanagement. Strategieentwicklung - Produktplanung - Organisation - Kontrolle, 3rd edn. Betriebswirtschaftlicher Verlag Dr. Th. Gabler/GWV Fachverlage GmbH, Wiesbaden (2007)
2. Zhu, Z., Dhokia, V.G., Nassehi, A., Newman, S.T.: A review of hybrid manufacturing processes – state of the art and future perspectives. Int. J. Comput. Integr. Manuf. **26**(7), 596–615 (2013)
3. Klocke, F.: Fertigungsverfahren 5. VDI-Buch. Springer, Heidelberg (2015)
4. Richard, H.A., Schramm, B., Zipsner, T.: Additive Fertigung von Bauteilen und Strukturen. Springer, Wiesbaden (2017)
5. Papke, T., Dubjella, P., Butzhammer, L., Huber, F., Petrunenko, O., Klose, D., Schmidt, M., Merklein, M.: Influence of a bending operation on the bonding strength for hybrid parts made of Ti-6Al-4V. Procedia CIRP **74**, 290–294 (2018)
6. Bambach, M.: Recent trends in metal forming: from process simulation and microstructure control in classical forming processes to hybrid combinations between forming and additive manufacturing. J. Mach. Eng. **16**(2), 5–17 (2016)
7. Papke, T., Junker, D., Huber, F., Hentschel, O., Schmidt, M., Merklein, M.: Verknüpfung additiver Fertigung und Umformtechnik in Form hybrider Bauteile und Werkzeuge. In: Witt, G. (ed.) Tagungsband Fachtagung Werkstoffe und Additive Fertigung, pp. 64–69 (2018)
8. Leyens, C., Peters, M.: Titanium and Titanium Alloys. Fundamentals and Applications. Wiley-VCH, Weinheim (2003)
9. Lütjering, G., Williams, J.C.: Titanium. Engineering Materials and Processes, 2nd edn. Springer, Berlin (2007)

10. Picu, R.C., Majorell, A.: Mechanical behavior of Ti-6Al-4V at high and moderate temperatures—Part II: constitutive modeling, vol. 326, pp. 306–316 (2002)
11. Butzhammer, L., Dubjella, P., Huber, F., Schaub, A., Aumüller, M., Baum, A., Petrunenko, O., Merklein, M., Schmidt, M.: Experimental investigation of a process chain combining sheet metal bending and laser beam melting of Ti-6Al-4V. In: Wissenschaftliche Gesellschaft Lasertechnik e.V. (ed.) Proceedings of the Lasers in Manufacturing LIM, München (2017)
12. Huber, F., Papke, T., Kerkien, M., Tost, F., Geyer, G., Merklein, M., Schmidt, M.: Customized exposure strategies for manufacturing hybrid parts by combining laser beam melting and sheet metal forming. J. Laser Appl. **31**(2), (2019)
13. Odenberger, E.-L.: Concepts for hot sheet metal forming of titanium alloys. Dissertation. Doctoral thesis/Luleå University of Technology, vol. 2009. Department of Applied Physics and Mechanical Engineering, Division of Solid Mechanics, Luleå University of Technology, Luleå (2009)
14. Cazacu, O., Plunkett, B., Barlat, F.: Orthotropic yield criterion for hexagonal closed packed metals. Int. J. Plast. **22**(7), 1171–1194 (2006)
15. Nemat-Nasser, S., Guo, W.-G., Nesterenko, V.F., Indrakanti, S.S., Gu, Y.-B.: Dynamic response of conventional and hot isostatically pressed Ti–6Al–4V alloys: experiments and modeling. In: Mechanics of Materials, vol. 33, pp. 425–439 (2001)
16. Schaub, A., Degner, J., Ahuja, B., Schmidt, M., Merklein, M.: Numerical investigation for superplastic forming tool development within the combined process chain of forming and additive manufacturing. In: Dimitrov, D., Oosthuizen, T. (eds.) International Conference on Competitive Manufacturing. COMA 2016: 27 January–29 January 2016, Stellenbosch, South Africa: Proceedings, pp. 173–179. Global Competitiveness Centre in Engineering Department of Industrial Engineering Stellenbosch University, Stellenbosch, South Africa (2016)

Investigation of a Superimposed Oscillation Compression Process for the Production of a Bulk Component

D. Rosenbusch(✉), P. Müller, S. Hübner, K. Brunotte, and B.-A. Behrens

Institute of Forming Technology and Machines, An der Universität 2, 30823 Garbsen, Germany
rosenbusch@ifum.uni-hannover.com

Abstract. In the scope of project A7 of the collaborative research center TCRC73 investigations of superimposed oscillation forming have been conducted. Positive effects regarding the reduction of plastic work, surface quality and mold filling in an ironing process could be demonstrated. Basic investigations also showed these effects for a compression process. By means of this research, the findings are transferred to the production of a functional component with an external gearing. For this purpose, a hydraulically working oscillating device is installed in an existing tool system in order to produce a demonstrator component. Forming experiments with and without superimposed oscillation are conducted. The components are examined concerning the plastic work required for forming as well as metallographic properties and surface quality. For this purpose, micrographs of the structure are taken and the flow lines and grain structure are analyzed. Hardness measurements are conducted and conclusions about the influence of superimposed oscillation on forming behavior are drawn.

Keywords: Superimposed oscillation · Gear · Sheet bulk metal forming

1 Introduction

In order to meet increasing demands on the economy and resource efficiency of production processes, a new production technique, Sheet Bulk Metal Forming (SBMF), has been developed in the scope of the Collaborative Research Center TCRC73. This manufacturing technology combines the positive characteristics of sheet metal and bulk metal forming. These include increased strain hardening due to forming, higher surface quality and a near-net-shape production [1]. In order to extend the limits of this technology, an SBMF process has been developed within the framework of subproject A7 at the Institute for Forming Technology and Machines (IFUM). In this process, a demonstrator component consisting of an internal and an external gearing is produced in three production steps [2]. The focus of subproject A7 is primarily on the investigation of dynamic

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 23–33, 2021.
https://doi.org/10.1007/978-3-662-62138-7_3

process forces in the main force flow of the forming machine. For this purpose, the internal gearing of the component is ironed under superimposed oscillation in the second stage of the SBMF process. As a result, mold filling is improved and plastic work reduced by means of superimposition oscillation [3]. In addition, ring-compression tests revealed that the surface quality also increases under superimposed oscillation [4].

Examinations of forming under superimposed oscillation date back to the 1950s. At that time, Blaha et al. observed a reduction in yield stress by means of superimposed oscillation forming of zinc single crystals in the ultrasonic frequency range [5]. Izumi et al. showed a yield stress reduction as a function of the oscillation amplitude in compression tests in a frequency range of 22 kHz [6]. Investigations for lower frequencies have been conducted by Osakada et al. for an oscillating tool die. Here, a forming force reduction of up to 57% could be demonstrated in a superimposed oscillation extrusion process with a frequency of 0.1 Hz [7]. In the industry, superimposed oscillation is exemplarily used for deep drawing processes by Jimma et al. [8]. A reduction in surface roughness could also be demonstrated by superimposed oscillation forming in comparison to oscillation-free forming [9]. The oscillation-induced reduction of plastic work can basically be explained by two effects, the surface-related effect and the volume-related effect [10]. The volume-related effect is defined as all vibration-induced effects inside the material. Kirchner et al. attributed the reduction in mean stress to the elastic relaxation of the material during stress release [11]. Surface-related effects describe the effects between tool and workpiece surfaces. Common explanations for a reduction in the average forming force due to surface-related effects are an oscillation induced reduction of the friction coefficient [12] and the erratic slide in stick-slip movements [13]. Furthermore, a solidification by plastic deformation under superimposed oscillation could be demonstrated. Kyryliv et al. investigated a method applying deformation under superimposed oscillation to the surface of cylindrical specimens of ferritic steel by means of a rotating centrifuge provided with balls. Here, solidification was reached by higher residual stresses due to superimposed oscillation in the range of $A = 5$ mm and $f = 24$ Hz [15].

In terms of the literature, investigations on superimposed oscillation forming in a frequency range of a few hundred Hertz are still largely unexplored. Ring compression tests at the IFUM showed significantly reduced friction as well as plastic work and increased surface quality in this frequency range [9]. In this present research, superimposed oscillation is applied to a compression process for forming a complex gear geometry. The plastic work required for forming is investigated as a function of the oscillation parameters. The microstructural properties of the components are analyzed in micrographs. Here, the focus is on an investigation of the flow lines. In addition, hardness measurements and an analysis of the quality of functional surfaces take place.

2 Experimental Setup

The experiments are conducted on the hydraulic press Wanzke GTA 250 at the IFUM. The tool system consists of a die, a deep drawing stamp and a compression stamp (see Fig. 1). A hydraulically operating oscillating device is installed beneath deep-drawing and compression stamp.

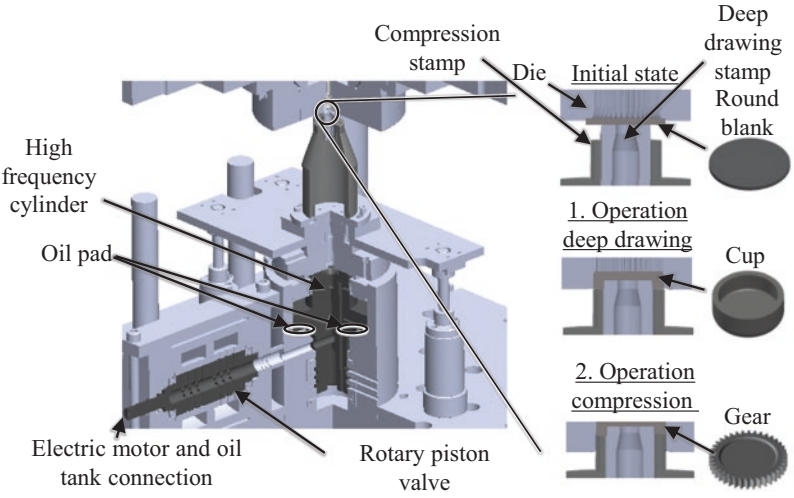


Fig. 1. Tool system for superimposed oscillation compression of a gear.

A high-frequency cylinder supported by an oil pad generates the oscillation. The oil pad is supplied with a volumetric oil flow by a rotary piston valve connected to pressurized oil tanks. The rotary piston valve is connected to an electric motor that can be used to adjust the oscillation frequency by varying the rotational speed [19]. The oscillating device can realize frequencies of up to $f = 600$ Hz and amplitudes of up to $A = 50$ µm [14]. For the experiments, round blanks with a diameter of 42 mm are cut out of a 2 mm thick DC04 steel sheet by waterjet cutting. The round blanks are positioned on the deep-drawing stamp in the tool system. During the experiments, a tool die moves down on the sheet metal blank and forms a cup. Subsequently, the oscillating device is activated and the compression stamp displaces the cup frame into the tooth cavity of the die. Tool die, compression stamp and deep-drawing stamp consist of the steel 1.3343 and are hardened to a value of 60 ± 2 HRC. The roughness of the gear geometry of the tool die is about $R_a = 0.8$ µm. The lubricant used is Beruforge 152 DL. All gears are compressed to the same tooth width of 4.35 mm in order to be able to compare the components. This compression height is chosen since the oscillating device reaches its operating limits according to the achievable forming force. To investigate the influence of both, frequency and amplitude, the amplitudes $A = 30$ µm and $A = 50$ µm as well as the frequencies $f = 200$ Hz and $f = 300$ Hz are varied. Therefore, the effects of changing amplitude at constant frequency as well as changing frequency at the same amplitude are considered. The experiments are repeated three times per parameter. The varied experimental parameters are shown in Table 1.

Table 1. Experimental parameters for the superimposed oscillated experiments.

Frequency, f	0 Hz	200 Hz	200 Hz	300 Hz
Amplitude, A	0 µm	30 µm	50 µm	30 µm

3 Results

In the following, the experimental results of the oscillation-free and superimposed oscillated experiments are evaluated. The influences of superimposed oscillation on the characteristics of the components during forming are investigated and analyzed. The focus is on the plastic work required for forming, the mechanical and metallographic properties and on an analysis of the surface quality of the components.

3.1 Influence on Plastic Work

The experiments are evaluated regarding the plastic work required for forming depending on the oscillation parameters. For this purpose, force-displacement curves are recorded during the forming process. The plastic work is calculated from the integral of the force over the deformation path using Eq. 1 [18].

$$W = \int_{s_o}^{s} F \, ds \tag{1}$$

With: W = plastic work
F = forming force
s_0 = initial state
s = forming path

Figure 2 shows the results of the determination of plastic work. The plastic work required for forming decreases with superimposed oscillation. At a frequency of $f = 200$ Hz and an amplitude of $A = 50$ µm, the plastic work reaches its lowest value of approx. $W = 2490 \pm 25$ J. At this process point, it is approx. 170 J below the value of oscillation-free forming. When the amplitude is reduced to $A = 30$ µm by keeping the frequency constant at $f = 200$ Hz, the plastic work increases and reaches a value of $W = 2520 \pm 5$ J.

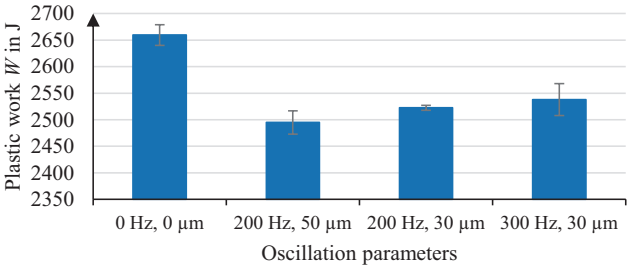


Fig. 2. Plastic work necessary for forming depending on oscillation parameters with different frequencies f (Hz) and amplitudes A (µm).

This can be explained by analyzing the force-path diagrams in Fig. 3. The area enclosed by the force-displacement curve is lower for an amplitude of $A = 50$ µm than it is for an amplitude of $A = 30$ µm. The amplitude is determined at the max. forming force. Thus, according to Eq. 1, plastic work also decreases. When the oscillation

frequency is increased from $f = 200$ Hz to $f = 300$ Hz by keeping the amplitude constant at $A = 30$ µm, no clear influence on the plastic work can be proven. Overall, there is a certain measurement deviation in the representation of the oscillation movement actually arriving at the sample due to natural frequencies in the tool system. This can explain the relatively high standard deviation for the 300 Hz oscillation in Fig. 2.

Nevertheless, it can be stated that the influence of the frequency on the deforma-

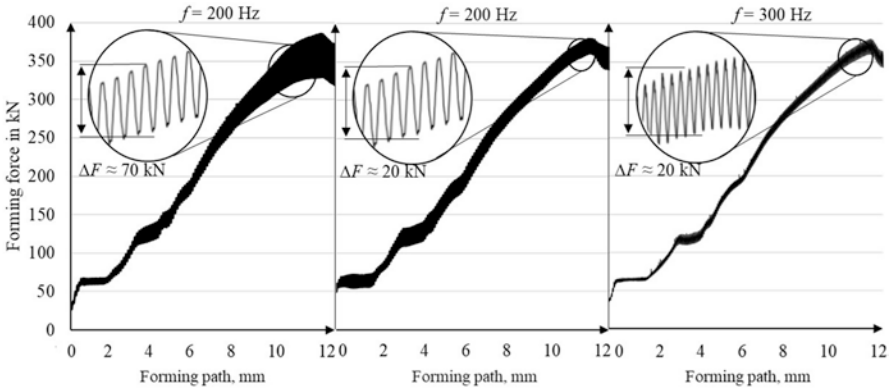


Fig. 3. Force-path diagrams for oscillation superimposed compression of the gear geometry.

tion energy is not as pronounced as the influence of the amplitude. Langenecker et al. explained a preferred influence of the oscillation amplitude on the reduction of yield stress [17], which can also be seen in the current investigation.

3.2 Metallographic and Mechanical Properties

The gears are examined regarding hardness depending on superimposed oscillated and oscillation-free forming. Since a superimposed oscillation with a frequency of $f = 200$ Hz and an amplitude of $A = 50$ µm seems to cause the highest reduction in plastic work, gears produced with these oscillation parameters are considered. The gears are cut open to the half of the tooth width. The hardness is measured at five positions above the tooth height with a Vickers hardness test with a test force of 1 N (see Fig. 4). Five gears are considered for each process parameter. The initial hardness of the DC04 sheets is 91–109 HV 1. For oscillation-free compressed samples, the highest average hardness values are measured in the area of the tooth flank with 184 to 190 HV 1 (positions 3 and 4 in Fig. 4). These values are significantly higher than the values measured in the area of the tooth head. In the tooth head, the average hardness for oscillation-free formed gears is in the range of 178 to 179 HV 1 (positions 1 and 2 in Fig. 4). This can be explained by higher degrees of deformation and thus higher strain hardening in the area of the tooth flanks. In the area of the middle lower tooth center (position 5 in Fig. 4), an average hardness of 184.4 HV 1 is measured for oscillation-free compressed gears. This can also be explained by higher degrees

of deformation compared to the tooth heads. For gears formed under superimposed oscillation, an average hardness of 188.4–184.2 HV 1 is measured at the tooth flanks (positions 3 and 4 in Fig. 4) and an average hardness of 181.6–182.6 HV 1 is measured at the tooth head (positions 1 and 2 in Fig. 4). Therefore, the hardness distribution between tooth flank and tooth head seems not to be as different as for gears compressed oscillation-free. An explanation for this is that the friction between component and tool surface is reduced by superimposed oscillation [16]. For the experiments under superimposed oscillation, the inner stresses in the tooth flank during compression may not be as different from the inner stresses in the tooth head due to the reduced friction between component surface and tool surface. Therefore, the hardness in the tooth wall of the components compressed under superimposed oscillation seems to be more homogeneously distributed than in the oscillation-free compressed components. It is noticeable that the hardness in the tooth wall area of the components manufactured under superimposed oscillation is very similar to the hardness of the components manufactured oscillation-free (see Fig. 4) and it is also more homogeneously distributed. Due to the relatively high standard deviation of the hardness values in the wall area, global conclusions for oscillation-induced softening or hardening effects cannot be drawn. The more homogeneous hardness distribution only indicates an oscillation-induced, more uniform material flow during forming.

Figure 4 illustrates that for gears compressed under superimposed oscillation the

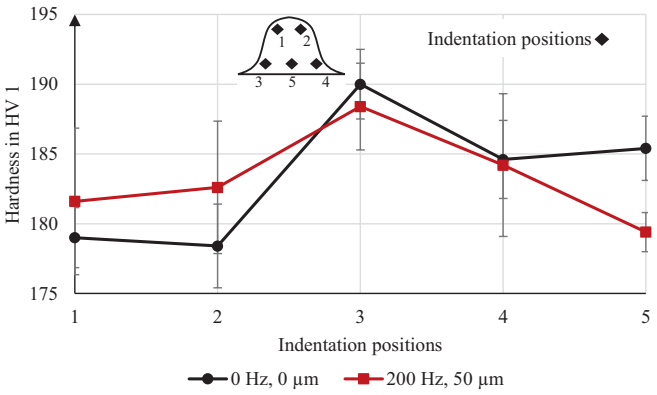


Fig. 4. Hardness gradients in oscillation-free and superimposed oscillation compressed gears.

average hardness value of 179.4 HV 1 lies below the average hardness value of 185.4 HV 1 for oscillation-free compressed gears in the lower middle tooth region (position 5 in Fig. 4). This can be explained by the fact that due to the cyclically alternating load under superimposed oscillation, the material flow is not as strongly steered in one direction as it is in the oscillation-free experiments. The microstructure is not as strongly deformed and therefore the material is not as strongly hardened in the middle tooth area as it is in the oscillation-free experiments. Since under superimposed oscillation additional contact effects, such as oscillating contact pressure between component and tool, occur in the area of the tooth wall, the hardness values for both process parameters are

more similar to each other in this area. In addition to the hardness distribution in the gears, their microstructure is examined as well. Micrographs of a tooth produced without superimposed oscillation are shown in Fig. 5 along with micrographs of a tooth produced under superimposed oscillation with $f = 200$ Hz and $A = 50$ µm. Due to volume constancy booth teeth have the same form filling. The height deviation in Fig. 5 results from a slightly inhomogeneous material distribution over the depth of the teeth. This is due to the fact that the tooth tip does not get in contact with the tool die and is therefore not under press quenching. The areas with the most conspicuous structural differences are exemplarily enlarged for the tooth head region, tooth middle region and tooth flank region. It can be seen that there is a strongly deformed structure in the area of the tooth root/flank (areas 1.) and 4.) in Fig. 5). The pronounced grain elongation indicates increased strain hardening, which explains the increased hardness values. In the area of the tooth head (areas 3.) and 6.) in Fig. 5), the grain elongation decreases and the grains become rounder. From this it can be concluded that strain hardening in the area of the tooth head is lower, which explains the tendency towards lower hardness values.

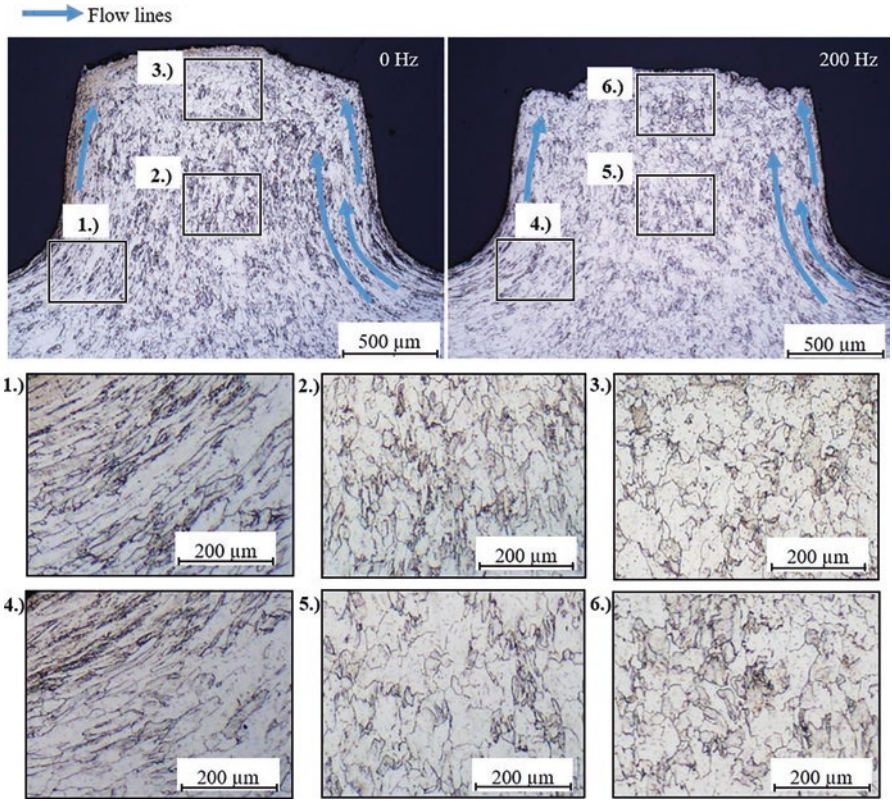


Fig. 5. Micrographs for an oscillation-free compressed gear tooth (left) and an oscillation superimposed compressed gear tooth (right)

In the area of the tooth center, the components produced without superimposed oscillation show more pronounced flow lines than the components produced with superimposed oscillation (see areas 2.) and 5.) Fig. 5). In addition, the components produced under superimposed oscillation tend to show pronounced flow lines in the area of the tooth head similar to the flow lines of the components produced without superimposed oscillation (compare areas 6.) and 3.) Fig. 5). This supports the assumption that the oscillation-induced effects causing a higher hardness primarily take place in the wall area of the gears due to the contact stresses between component and tool. In the area of the tooth center, superimposed oscillation leads to less directional grain stretching. This can be explained by cyclic loading, which does not force grain stretching in one direction. Therefore, lower hardness values are measured in the center area of the gears for components compressed under superimposed oscillation.

3.3 Influence on Surface Quality

The surface quality of the tooth flanks is one of the most important quality features for gears. The surface roughness of the tooth flanks is measured by the optical macroscope VR3200 of the company Keyence for gears compressed oscillation-free and under superimposed oscillation. The arithmetic average roughness S_a of the tooth flank surfaces is shown in Fig. 6. It is stated that an oscillation overlay affects the quality of the tooth flank surfaces. For a superimposed-oscillated compression process with a frequency of $f = 200$ Hz and an amplitude of $A = 50$ µm, the greatest surface smoothening is determined (see Fig. 6). For these oscillation parameters, the arithmetic mean roughness decreases from $S_a = 1.03 \pm 0.1$ µm with oscillation-free forming by approx. 22% to $S_a = 0.8 \pm 0.1$ µm with superimposed-oscillation forming. In addition, a tendency can be seen that the smoothening effect is reduced by reducing the amplitude from $A = 50$ µm to $A = 30$ µm at a constant frequency of $f = 200$ Hz.

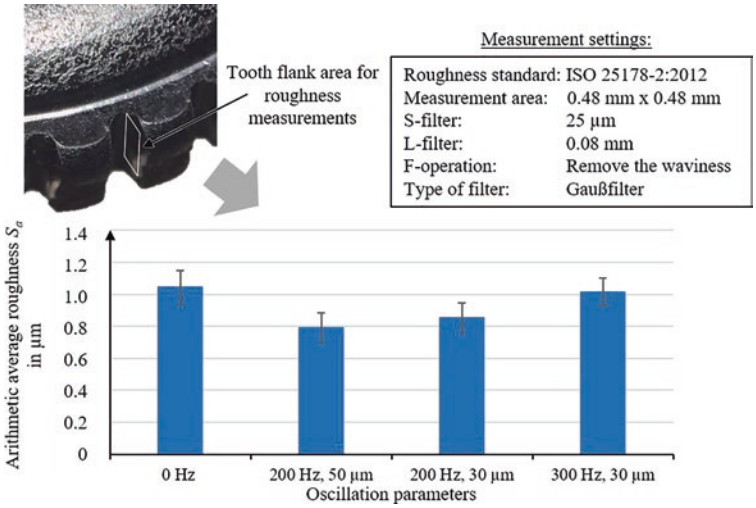


Fig. 6. Measurement of the surface quality of the tooth flanks

For an oscillation of $f = 200$ Hz and $A = 30$ µm the S_a value is 0.83 ± 0.12 µm and thus only reduced by about 19% compared to oscillation-free forming. In addition, the measurements show that the effect of surface smoothing is further inhibited by increasing the frequency up to $f = 300$ Hz with a constant amplitude of $A = 30$ µm. In this case, the surface roughness is only reduced in a range of about 2% compared to oscillation-free forming which is negligible. A high amplitude in combination with a frequency in the range of 200 Hz prove to be the most promising of all examined oscillation parameters with regard to an improvement of surface quality.

4 Conclusion and Outlook

In this research, the influence of superimposed oscillation on a compression process for producing a demonstrator component with external gearing was investigated. It was found that superimposed oscillation with a frequency of $f = 200$ Hz and an amplitude of $A = 50$ µm leads to the highest energy savings in forming. This effect was attributed to the increased influence of the oscillation amplitude. With regard to an increase in surface quality, these oscillation parameters ($f = 200$ Hz, $A = 50$ µm) also proved to be the most promising ones. This can be explained by oscillation-induced smoothing effects, which occur preferentially due to an increased amplitude. Furthermore, components manufactured with these oscillation parameters were compared to components manufactured oscillation-free with regard to their metallographic properties. An evaluation of the hardness curves and the microstructures showed that the hardness is highest at the tooth flanks due to an increased degree of deformation. In addition, the hardness in the tooth wall seems to be more distributed pronounced for gears produced under superimposed oscillation than it is for gears produced oscillation-free. This is explained by a cyclically alternating load during superimposed oscillated forming, leading to a more uniform structural deformation over the tooth height. Another explanation for the more uniform hardness distribution is an oscillation-induced reduction in the friction between tool surface and component surface, leading to lower contact shear stresses between component and tool and thus to lower internal stresses in the tooth flanks of the component. Due to the high standard deviations of the hardness measurements, it is hardly possible to make clear statements about oscillation-induced changes in work hardening. The given conclusions show tendencies which suggest an influence but are not clearly proven. More in-depth investigations are to be conducted in following researches. In addition, the frequency range below 200 Hz will be considered in more detail in future investigations.

Acknowledgments. Funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) –in the project A7 (grant number 116817829) within the scope of the Transregional Collaborative Research Centre 73 “Sheet Bulk Metal Forming” (TCRC 73).

References

1. Merklein, M., Allwood, J. M., Behrens, B. A., Brosius, A., Hagenah, H., Kuzman, Mori, K., K., Tekkaya, A. E., Weckenmann, A.: Bulk forming of sheet metal. CIRP Ann. Manuf. Technol. **59**(2), 725–745 (2012)
2. Behrens, B.A., Hübner, S., Müller, P., Besserer, H.B., Gerstein, G., Koch, S., Rosenbusch, D.: New multistage sheet-bulk metal forming process using oscillating tools. Metals **10**(5), 617 (2020)
3. Behrens, B. A., Bouguecha, A., Vucetic, M., Hübner, S., Rosenbusch, D., Koch, S.: Numerical and experimental investigations of multistage sheet-bulk metal forming process with compound press tools. In: 18th International ESAFORM Conference on Material Forming. Key Engineering Materials, vol. 651–653, pp. 1153–1158 (2015)
4. Koch, S., Vucetic, M., Hübner, S., Bouguecha, A., Behrens, B.A.: Superimposed oscillating and non-oscillating ring compression tests for sheet-bulk metal forming technology. Adv. Mater. Res. **794**, 89–96 (2015)
5. Blaha, F., Langenecker, B.: Dehnung von Zink-Kristallen unter Ultraschalleinwirkung. Die Naturwissenschaften **42**(20), 556 (1955)
6. Izumi, O., Oyama, K., Suzuki, Y.: Effects of superimposed ultrasonic vibration on compressive deformation of metals. Trans. Japan Inst. Met. **7**(3), 162–167 (1966)
7. Osakada, K., Wang, X., Hanami, S.: Precision forging of spline by flashless die forging with axially driven die. Ann. CIRP **46**(1), 209–212 (1997)
8. Jimma, T., Kasuga, Y., Iwaki, N., Miyazawa, O., Mori, E., Ito, K., Hatano, H.: An application of ultrasonic vibration to the deep drawing process. J. Mater. Process. Technol. **80–81**, 406–412 (1998)
9. Müller P., Rosenbusch D., Wehmeyer J., Hübner S., Behrens B.-A.: Investigations of forming force, friction values and surface qualities in ring compression tests using oscillating tools, Production at the leading edge of technology. In: Proceedings of the 9th Congress of the German Academic Association for Production Technology (WGP), Hamburg, pp. 73–81 (2019)
10. Bunget, C., Ngaile, G.: Influence of ultrasonic vibration on micro-extrusion. Ultrasonics **51**(5), 606–616 (2011)
11. Kirchner, H.O.K., Kromp, W.K., Prinz, F.-B., Trimmel, P.: Plastic deformation under simultaneous cyclic and unidirectional loading at low and ultrasonic frequencies. Mater. Sci. Eng. **68**(2), 197–206 (1985)
12. Rozner, A.G.: Effect of ultrasonic vibration on coefficient of friction during strip drawing. J. Acoust. Soc. Am. **49**(5), 1368–1371 (1971)
13. Starcevic, J.: Tribospektroskopie als neue Methode zur Untersuchung von Reibungsmechanismen: Theoretische Grundlagen und Experiment. TU Berlin-Dissertation (2008)
14. Behrens, B.A., Hübner, S., Vucetic, M.: Influence of superimposed oscillation on sheet-bulk metal forming. Key Eng. Mater. **554–557**, 1484–1489 (2013)
15. Kyryliv, V., Kyryliv, Y., Sas, N.: Formation of surface ultrafine grain structure and their physical and mechanical characteristics using vibration-centrifugal hardening. Adv. Mater. Sci. Eng. **2018**, 7 (2018)
16. Heß, B., Groche, P.: Untersuchungen zum oszillierenden Verzahnungsdrücken. Schmiede J. **03**, 38–40 (2014)

17. Blaha, F., Langenecker, B.: Plastizitätsuntersuchungen von Metallkristallen in Ultraschallfeld. Acta Metall. **7**(2), 93–100 (1959)
18. Resnick, R., Halliday, D.: Physics, Section 1–3 (Vol I and II, Combined edition). Wiley International Edition, Library of Congress Catalog Card No. 66-11527 (1966)
19. Behrens, B.-A., Hübner, S., Krimm, R., Wager, C., Vucetic, M., Cahyono, T.: Development of a hydraulic actuator to superimpose oscillation in metal-forming presses. In: 14th International Conference on Sheet Metal, pp. 217–222 (2011)

Towards an Adaptive Production Chain for Sustainable Sheet-Metal Blanked Components

P. Niemietz(✉), T. Kaufmann, M. Unterberg, D. Trauth, and T. Bergs

Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
p.niemietz@wzl.rwth-aachen.de

Abstract. Up to 250 fine-blanked components per luxury-class vehicle make a considerable contribution to the automotive production supply chain. High scrap rates in an industrial setting imply an inefficient use of the resources and process design and cause avoidable CO_2-emissions. This contribution focusses on a data-driven production cycle optimization of workpiece quality features and enabling efficient usage of resources in a production chain for the life cycle of fine blanked components. As one of the quality features in fine blanking, the die roll reduces the geometric accuracy and the functional surface area and thus needs to be removed in a secondary finishing process, typically grinding. An experimental setup is deployed using an industrial fine blanking and surface grinding machine on a shopfloor interconnected by an Edge-/Cloud-Architecture. This approach enables quality feature prediction and process adaption to enhance workpiece quality and efficient usage of resources, eventually leading to increasingly sustainable process chains.

Keywords: Grinding · Fine blanking · Networked process chains · Sheet metal forming

1 Introduction

The amount of energy and resource consumption is a key characteristic of sustainability in manufacturing technologies [1]. Regarding a production life cycle of a unique component, machine operating time, tool wear, workpiece material, and process utilities, such as energy and fluids, have to be interpreted on a process level. Additionally, used resources of complete process chains or plants need to be taken into account. Salonitis and Ball have presented different approaches to take energy efficiency in both levels into account [2]. Here, the polylemma in manufacturing technologies regarding time, flexibility, and cost, while achieving the best component quality, is expanded by sustainability as an additional dimension these days.

On the one hand, digitization enables machine tools to share information, data and knowledge, on the other hand, machines can learn from each other during production

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 34–44, 2021.
https://doi.org/10.1007/978-3-662-62138-7_4

processes. The reliable prediction of component result variables offers the potential to eliminate quality gates without the need of foregoing knowledge. Since knowledge of the individual workpiece geometry and surface area characteristics allows the parameters of the manufacturing process to be adapted accordingly and thus robustly compensates for internal and external disturbances or even prevents them and increases efficiency [3].

The process chain considered in this paper consists of a fine blanking line and a surface grinding process that is used to finish the surface of the fine blanked component to guarantee geometric accuracy by removing the die roll. Fine blanking is a precision forming process for manufacturing large quantities of (ideally) identical workpieces, e.g. for the aerospace and automotive services, thus reducing costs during production. To this end, fine blanking utilizes characteristic operating parameters that result in the production of components with high quality of the sheared surface and geometric accuracy [4]. Grinding is an important manufacturing and finishing process in metalworking industries and is used to improve surface quality as well as dimensional and shape accuracy. Thus, grinding is usually at the end of the value-added chain, where the individual component value is high. Accordingly, surface grinding is an essential and downstream production step for blanking processes [5]. Machining errors occur during grinding, particularly due to high temperatures in the contact zone between the abrasive grain and the workpiece and can thereby possibly compromise the quality characteristics of previous processes.

Utilizing modern sensor systems to monitor both manufacturing processes in detail already reveals valuable process insights, mitigating the uncertainty of the process outcome in fine blanking [6] as well as grinding [7]. To establish adaptive and sustainable process chains, the exchange of knowledge from fineblanking and grinding machines has to be enhanced as an example of an industrially relevant technology chain. By adjusting the grinding process at minimum machining time for each workpiece individually, or to individual batches of workpieces, the process can be run in its optimal setup for every batch, possibly reducing wear, use of lubricant and energy consumption and thus enhancing process resilience while enhancing sustainability.

This paper focuses on identifying potentials for the digitization of the aforementioned process chain with the higher purpose of increasing the sustainability of production processes across multiple silos. First, current approaches and state of the art are discussed in Sect. 2, followed by a detailed description of the model and suitable technologies that meet the discussed requirements in Sect. 3. In Sect. 4 the use case is described in detail and the current state of research is presented.

2 Implications of Digital Technologies for Manufacturing

Process chains in manufacturing are typically designed to achieve a global revenue optimum with a tradeoff between costs and quality of the outcome, while also considering sustainability issues. By introducing digital technology into the design of process chains, the availability of data within singular processes and the exchange of data between the processes is enabled, where initially no or only a little information has

been exchanged. This section sums up the current state of data analysis in the next subsection, followed by the potential of using digital technologies in manufacturing processes and, finally, presents the related works of approaches in grinding and sheet metal forming.

2.1 Prerequisites of Data Processing in Manufacturing

In manufacturing environments, sensory equipment usually comes in high frequencies, e. g. in mechanical vibration measurements. Thus, while designing digital solutions to store and analyze manufacturing data, requirements often surpass traditional IoT-Infrastructure approaches designed to handle massive amounts of low frequency sensor data and advanced options need to be considered.

Sensory information needs to be acquired from different sources in different manufacturing systems, often involving three critical components, namely the used material, the machine tool plus periphery, and the process itself. Thus, holistic sensor approaches combine analog sensors like force or acoustic emission sensors, machine control interfaces e.g. OPC-UA, and sensors specialized in the online measurement of surface condition, thickness, or hardness properties of the incoming material for each manufacturing process. Heterogeneous data coming from sensor systems have to be organized and enriched with the context of the manufacturing process since only data with proper provenance and contextual information can effectively be used to derive meaningful analytical models [8]. Thus, semantic data models and implementation of FAIR-Data Principles [9] allow for interoperable information systems that enable the exchange of information and dependencies between data. On top of the data model, storage systems and efficient processing models in the network and at the edge, as well as on compute clusters or in the cloud for large datasets have to be taken into account. Depending on the considered scenario, there are specific latency requirements that have to be met by the selected IT-system. The question emerging is not how to design IT-systems to achieve latency comparable to OT-systems, but rather what is needed for the specific use case and which concepts meet this requirement [10].

2.2 Current Situation

As of today, both grinding and fine blanking are mostly analog processes and machine tool manufacturers show little intention to realize concepts like cyber-physical systems due to the complexity of the matter and a lack of know-how in the required digital technologies.

For fine blanking, the die roll is one of the major quality features that affect subsequent manufacturing processes such as grinding, since it influences the geometric accuracy and especially reduces the size of the functional surface of the resulting workpiece. The die roll depends on many nonlinear relationships [11] and fluctuates highly in an industrial setting. Past research showed that there is a strong relationship between the die roll height, material, material thickness, punch-die clearance, and tool geometry [12]. Thus, traditional analytical approaches are inappropriate to model

this relationship to a satisfying extend, whereas machine learning (ML) approaches are promising [11]. Stanke et al. trained machine learning models with validated simulation data to accelerate die roll prediction as opposed to traditional FE simulation. The models were able to predict the die roll height given 8 variable parameters [11]. Since only scalar values as input data have been used for this model, these models do not learn any temporal information that varies significantly in an industrial setting [6].

In different work, the temporal punch-to-punch variation even within the same experimental setting (e.g. cutting-edge radius and sheet thickness) can be detected in the 2-dimensional representation of statistical features. However, it is not further considered due to the absence of an accurate model for interpreting those changes [13]. Furthermore, Havinga et al. [14] showed that force signature fluctuation can be used to predict part-to-part quality variations, which is the goal of every monitoring approach. Additionally, it has been shown that strain and force measurements contain multiple information about the current state of the tool, such as general and adhesive wear [13], misfeed, sheet metal thickness [15] or different lubrication regimen [16].

In grinding, however, due to the limited accessibility of the contact zone between abrasive grain and workpiece, a large number of influencing parameters cannot be measured, or can only be measured with difficulty and a high degree of uncertainty. It is therefore necessary to make valuable direct process signals indirectly measurable. One of the most frequent causes of workpiece damage in grinding is thermal surface damage due to excessive temperatures in the contact zone [17] resulting from forces, wear, and friction. Therefore, a prediction of the heat input due to several process parameters makes knowledge about metallurgical alterations, micro-cracks on and below the surface, and residual stresses of the material accessible. Experimental validation of different simulation methodologies for predicting the workpiece temperature, regarding the heat flux into the workpiece depth, is difficult and still hard to verify according to Wrobel [18]. The provision and treatment of the necessary cooling lubricant represent a significant proportion of the total CO_2-emissions of a machine tool, accounting for 20 % of total emissions [19]. A reduction of this share by up to 60 % [20], e. g. through demand-oriented supply in addition to considerable resource savings due to fewer rejects, allows a significant competitive advantage over low-wage countries.

2.3 The Gap

Combining several manufacturing processes into the balancing of resources as well as adaptive process control by knowledge of the value of preceding or cross-process relevant quality characteristics have, however, only been investigated to a limited extent. In this context, a methodology for integrative production planning [21], in which a set of requirements (i. a. highly adaptive processes themselves, independent information flow to prevent information flow) is defined and evaluated from Computer-Aided Design (CAD) to manufacturing, has been researched [22].

Summarizing the information of the preceding chapter, for both processes, namely fine blanking and grinding, various research has shown that available sensory signals contain rich information about the conditions of the process and can be used to

predict the quality of the individual components. In contrast, the potential of incorporation of the available data into models that represent the behavior of the two-step process chain, allowing per part or batch optimization of the subsequent process, has not been researched yet. This paper aims to present an approach that combines state of the art methods to overcome this challenge. Figure 1 illustrates the aforementioned two-step process chain, starting at raw material via first and second machining operation ending with a finished workpiece that meets the highest quality standards. Along the entire value-added process of the networked units, sensor, machine and workpiece data are merged to design downstream production steps adaptively and thus efficiently to improve sustainability in manufacturing processes.

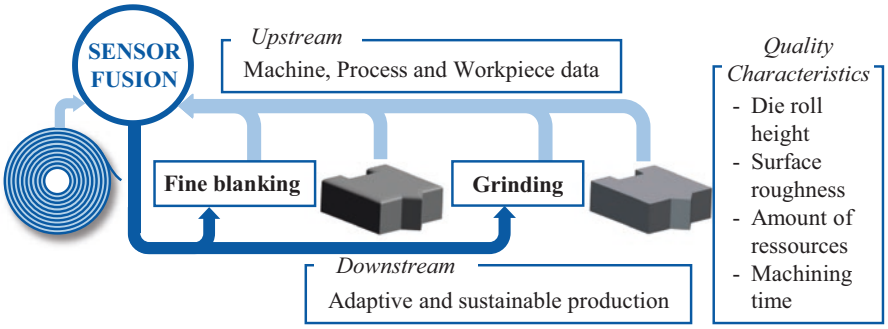


Fig. 1. Scalable two-step production chain containing up- and down-streaming data exchange for an integrative production planning.

3 Mitigating the Uncertainty in Fine Blanking - Grinding

In the approach presented by this paper, the objective is set on quantifying and understanding process variations and their impact on quality and wear characteristics throughout an economically important production chain consisting of fine blanking and grinding. Research shows that material properties [23], stamping process data [6], and grinding process data highly vary, even though all process parameters remain unchanged, caused by the dynamics of the machine tools, wear characteristics, material inhomogeneities, environmental changes or machine faults. Modelling these fluctuations on a sensor basis and enriching them with the varying workpiece quality enables the setup of subsequent production steps on a per workpiece basis. Such sensory systems can be used to widen process tolerances, to optimize production chain sustainability, and to increase revenue. The outcome of a successful implementation of this approach is a model that can be used to i) quantify the influence of varying material properties onto the process signals, ii) predict the height and width of the die roll of every produced workpiece, given the material properties and process signals, iii) predict the surface roughness, thermal damage of the workpiece external zone and the final geometry while machining fine-blanked workpieces and, lastly, iv) give

recommendations for actions on how to set up the fine blanking and grinding process accordingly to minimize the wear, improve quality features, and the number of operational resources required while maintaining highest workpiece quality.

3.1 Quantifying Process Variations

Variations in sensor signals and quality metrics occur in every manufacturing process and are often regarded as an unavoidable effect, since no process is ever performed under the exactly the same conditions, thus, naturally signals and metrics vary over time to a certain extend. In an attempt to mitigate this, studies have shown severe fluctuations on both, signal and quality level for hardness conditions of cold rolled coils [23], die roll of fine blanked components [24] and surface roughness of ground parts [7], see Fig. 2.

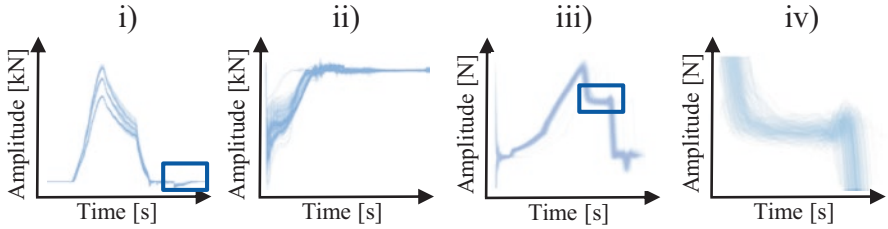


Fig. 2. From left to right: i) Stamping force of 300 punches of fine blanking ii) an expansion of the corresponding signals during punch ejection, iii) the tangential grinding force component during cylindrical grinding of 400 workpieces and iv) the corresponding level during finishing after roughing.

Furthermore, for grinding and fine blanking, fluctuations in force signals acquired during the process occur on a part-to-part basis that is even more significant considering specific segments within the signals [6]. Following the aforementioned order, the first step is to quantify and visualize these variations for each produced part for each sensor. This allows further investigations and data processing to determine which variations have a significant impact on the outcome of a process, and which actually can be regarded as noise. Figure 3 visualizes the possible influence of quality features of subsequent processes using the example of hardness after cold rolling of a 16MnCr5 (AISI: 5115) steel to a thickness of 4 mm, die roll height after blanking and surface roughness after grinding.

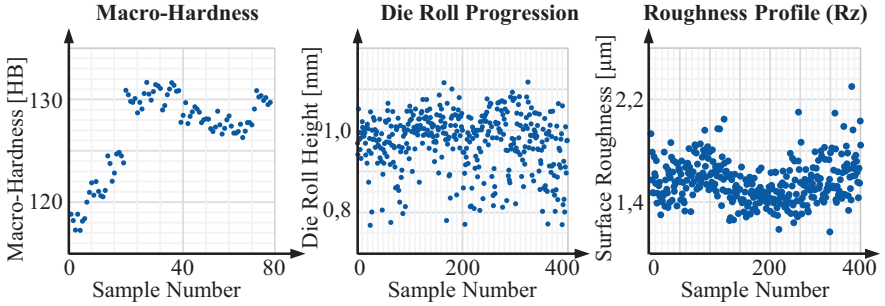


Fig. 3. From left to right: Measurements of subsequent samples in i) coil hardness along the length of a coil (max. 13 % variations), ii) the die roll of workpieces produced in 400 subsequent fine blanking operations (max. 29 % variations), iii) surface roughness of workpieces ground in 400 subsequent grinding operations (max. 77 % variations).

Methods that give meaningful insight into the variations in unsupervised time series data are labeled as time series data mining techniques and are traditionally performed in the working stages of data preparation and cleaning, extraction of meaningful features as a representation of the signal, reduction to a small number of significant features, and, eventually, clustering in homogeneous clusters within the feature space. Doing so allows for the identification of outliers/anomalies in sensor data and quantification of the variation of a signal by using different distance metrics within the derived feature space. A holistic implementation on all sensor signals yields an exact classification of which signals deviate from the previously identified "measured standard" for every working step and thus is a necessity for an adaptive process design.

3.2 Predicting Quality for Intermediary Steps

According to Rowe and Malkin, the main objective of using digital technologies in manufacturing technologies is to "produce higher quality components more efficiently with smaller batch sizes and more frequent changeovers" [25]. To mitigate the uncertainty throughout production chains and to enable the optimization potentially for every position on the coil during the fine blanking process and subsequently for the specific geometric accuracy of the fine blanked workpieces during grinding, these intermediate process outcomes need to be modelled by reliable prediction models.

By treating the development of the die roll height in fine blanking and surface quality in grinding as a machine learning problem, a model has to be trained where the die roll height can be viewed as the output $y \in \mathbb{R}$ of the model, which is a scalar value. The input for such models can potentially consist of various heterogeneous sensors, representing conditions of input material and conditions during the process. Sensor fusion models that combine significant information from e.g. process forces, magnetic barkhausen noise emission, material thickness etc. are viable approaches to this

problem. A corresponding model would have the goal to find a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$, so that

$$f(\vec{x}) \approx y, \tag{1}$$

where $\vec{x}$ refers to the n dimensional input of the model.

Promising approaches for such sensor fusion models are multi-channel convolutional neural networks, that are able to handle time-series data through spatial relationships in and between different time series [26]. With such a model material related, force related, and synergetic effects between material condition and force profiles on the development of the die roll height and the surface roughness can be considered in the prediction of the quality features.

3.3 From Monitoring to Adaption

Successful and accurate implementation and calculation of the aforementioned models for fine blanking and grinding allows the prediction of intermediate results and eventually the final result of the process chain can be estimated and optimized in terms of cost, quality, and sustainability of the process chain.

In the fourth step of the proposed methodology, the potential of information exchange between manufacturing steps in a production chain has to be investigated and modeled. While research presented in Sect. 2 indicates that all sensory signals acquired in each considered manufacturing process contain valuable information about individual process conditions and the resulting quality features, the cause-effect relations for data-driven approaches are still unclear for process chains. Models that allow for the identification and interpretation of such effects have to consider raw sensor signals, simulation data, workpiece related information, metadata, and output of analytical models of previous manufacturing steps. Using the example of grinding, the adjustment of process parameters while ensuring the surface integrity at the highest resource efficiency itself is challenging, but utilizing inputs such as die roll height, geometric shape and boundaries, magnetic barkhausen noise, stamping force profile, and data from simulations increases the complexity. In fact, an important subtask is to develop methods that detect the significance of parameter changes throughout process chains while indicating which process parameters may be neglected in future setups.

4 Defining the Roadmap

While the goal is clearly defined, the effort needed to make progress investigating the uncertainty and adaptability of currently relative static process chains is high and cumbersome. The current setup to investigate these problems includes the installation of various sensor systems, followed by sensor integration and data preparation. At the fine blanking line, a Feintool XFT 2500speed, various sensors for incoming material, such as surface roughness, thickness, magnetic barkhausen noise, as well as analog sensors integrated into the tool to measure process forces, namely punch, counterpunch, and blank holder force, acoustic emission signals and acceleration as

well as the acquisition of machine control information via OPC-UA interfaces have been installed. At the same time, a surface grinding machine of type Blohm Profimat MT608 has been equipped with force measurement systems, spindle power and axis feed rate diagnostics, coolant supply information (volume flow rate, pressure and temperature at every nozzle) and integrated into a holistic acquisition system.

Experiments to generate a good amount of data samples are to be executed on both machines, to start with modelling i) sensor signal behavior itself with various data mining techniques, ii) die roll width and height variations in fine blanking experiments, iii) corresponding surface roughness resulting from the surface grinding process and iv) optimization of the complete process chain for cost, quality, flexibility, and sustainability using data-driven approaches from all domains. To this extend, scientific data sets that encompass the material properties of a to be fine-blanked coil on every position, all sensor signals acquired during the fine blanking process on a workpiece level, and the resulting process signals and quality metrics after the grinding operations required are currently in preparation and a preliminary for further research. This approach will be carried out with the same material that has been used to uncover variations in hardness after rolling, die roll development after blanking, and roughness uncertainties after grinding (Sect. 3.1, 16MnCr5, AISI: 5115). In developing the methodology for an adaptive process chain, the authors are nevertheless endeavouring to keep the approach generic, so that it can be applied not only to other ferromagnetic blanking steels but also to improve in sustainability through adaptive process control available across manufacturing technologies.

Acknowledgements. This work is funded in parts by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612. This research is partially funded by the BMWi KI-Innovationswettbewerb – SPAICER - 01MK20015B.

References

1. Winter, M., Li, W., Kara, S., Herrmann, C.: Determining optimal process parameters to increase the eco-efficiency of grinding processes. J. Clean. Prod. **66**, 644–654 (2014)
2. Salonitis, K., Ball, P.: Energy efficient manufacturing from machine tools to manufacturing systems. Procedia CIRP **7**, 634–639 (2013)
3. Kusiak, A.: Fundamentals of smart manufacturing: a multi-thread perspective. Ann. Rev. Control **47**, 214–220 (2019)
4. Zheng, Q., Zhuang, X., Zhao, Z.: State-of-the-art and future challenge in fine-blanking technology. Prod. Eng. **13**(1), 61–70 (2018)
5. Jayakumar, T., Mukhopadhyay, C.K., Venugopal, S., Mannan, S.L., Raj, B.: A review of the application of acoustic emission techniques for monitoring forming and grinding processes. J. Mater. Process. Technol. **159**(1), 48–61 (2005)
6. Niemietz, P., Pennekamp, J., Kunze, I., Trauth, D., Wehrle, K., Bergs, T.: Stamping process modelling in an Internet of Production. Procedia Manuf. **49**, 61–68 (2020)
7. Kaufmann, T., Sahay, S., Niemietz, P., Trauth, D., Maaß, W., Bergs, T.: AI-based framework for deep learning applications in grinding. In: Proceedings of the IEEE 18th World

Symposium on Applied Machine Intelligence and Informatics (SAMI), pp. 195–200. IEEE, Herlany, Slovakia (2020)
8. Gleim, L., Pennekamp, J., Liebenberg, M., Buchsbaum, M., Niemietz, P., Knape, S., Epple, A., Storms, S., Trauth, D., Bergs, T., Brecher, C., Decker, S., Lakemeyer, G., Wehrle, K.: FactDAG: formalizing data interoperability in an internet of production. IEEE Internet Things J. **7**(4), 3243–3325 (2020)
9. Wilkinson, M.D., Dumontier, M., Aalbersberg, I.J.: The FAIR guiding principles for scientific data management and stewardship. Sci. Data **3**, 160018 (2016)
10. Glebke, R., Henze, M., Wehrle, K., Niemietz, P., Trauth, D., Mattfeld, P., Bergs, T.: A case for integrated data processing in large-scale cyber-physical systems. In: Proceedings of the 52nd Hawaii International Conference on System Sciences (HICSS), pp. 7252–7261. Wailea, USA (2019)
11. Stanke, J., Feuerhack, A., Trauth, D., Mattfeld, P., Klocke, F.: A predictive model for die roll height in fine blanking using machine learning methods. Procedia Manuf. **15**, 570–577 (2018)
12. Fuchiwaki, K., Mure, Y., Yoshida, K., Murakawa, M.: Prediction of die-roll in fine blanking by use of profile parameters. Procedia Eng. **207**, 1564–1569 (2017)
13. Hoppe, F., Hohmann, J., Knoll, M., Kubik, C., Groche, P.: Feature-based supervision of shear cutting processes on the basis of force measurements: evaluation of feature engineering and feature extraction. Procedia Manuf. **34**, 847–856 (2019)
14. Havinga, J., Van Den Boogaard, T.: Estimating product-to-product variations in metal forming using force measurements. In: Proceedings of the 20th International ESAFORM Conference on Material Forming ESAFORM 2017, vol. 1896, no. 1, pp. 070002. American Institute of Physics, Dublin, Ireland (2017)
15. Bassiuny, A.M., Li, X., Du, R.: Fault diagnosis of stamping process based on empirical mode decomposition and learning vector quantization. Int. J. Mach. Tools Manuf. **47**(15), 2298–2306 (2007)
16. Voss, M., Pereira, P., Rolfe, F., Doolan, C.: Using stamping punch force variation for the identification of changes in lubrication and wear mechanism. J. Phys. Conf. Ser. **896**(1), 012028 (2017)
17. Karpuschewski B., Inasaki I.: Monitoring systems for grinding processes. In: Wang L., Gao R.X. (eds.) Condition Monitoring and Control for Intelligent Manufacturing. Springer Series in Advanced Manufacturing, pp. 83–107. Springer, London (2006)
18. Wrobel, C., Mattfeld, P., Trauth, D., Klocke, F.: Modeling of the temperature field in the workpiece external zone as a function of the grinding wheel topography. Procedia CIRP **77**, 291–294 (2018)
19. Gutowski, T., Dahmus, J., Thiriez, A.: Electrical energy requirements for manufacturing processes. In: 13th CIRP International Conference on Life Cycle Engineering, vol. 31, no. 1, pp. 623–627. Elsevier, Leuven (2006)
20. MAV Homepage. https://mav.industrie.de/peripherie/anlagen-verfahren/smarte-pumpen-auslegung. Accessed 15 Mar 2020
21. Nyhuis, P., Wulf, S., Klemke, T., Hirsch, B.: Integrative factory, technology, and product planning-systemizing the information transfer on the operational level. Prod. Eng. **4**(2–3), 231–237 (2010)
22. Denkena, B., Dittrich, M.-A., Jacob, S.: Methodology for integrative production planning in highly dynamic environments. Prod. Eng. **13**(3–4), 317–324 (2019)

23. Unterberg, M., Niemietz, P., Trauth, D., Wehrle, K., Bergs, T.: In-situ material classification in sheet-metal blanking using deep convolutional neural networks. Prod. Eng. **13**(6), 743–749 (2019)
24. Bergs, T., Niemietz, P., Kaufman, T., Trauth, D.: Punch-to-punch variations in stamping processes. In: Proceedings of the IEEE 18th World Symposium on Applied Machine Intelligence and Informatics (SAMI), pp. 213–218. IEEE, Herlany (2020)
25. Rowe, W.B., Yan, L., Inasaki, I., Malkin, S.: Applications of artificial intelligence in grinding. CIRP Ann. **43**(2), 521–531 (1994)
26. Münzner, S., Schmidt, P., Reiss, A., Hanselmann, M., Stiefelhagen, R., Dürichen, R.: CNN-based sensor fusion techniques for multimodal human activity recognition. In: Proceedings of the 2017 ACM International Symposium on Wearable Computers, pp. 158–165. Association of Computing Machinery, New York, USA (2017)

Investigation on Noise Reduction During Cutting of High-Strength Materials Based on Machine Acoustic Simulation

D. Friesen(✉), R. Krimm, S. Fries, K. Brunotte, and B.-A. Behrens

Institute of Forming Technology and Machines, Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany
friesen@ifum.uni-hannover.de

Abstract. The increasing use of high-strength materials in sheet metal processing results in challenges for cutting presses. The required cutting force increases with the strength of the material. After material separation, a sudden release of the energy stored in the machine leads to intense system vibrations and noise emission. Noise protection and reduction solutions such as machine cabins or active vibration damping systems are expensive, require much space or frequent maintenance. Due to this, forming machines must be acoustically optimized. In this article, selected structure measures for the machine-side noise reduction are derived and evaluated in the context of simulation studies using an acoustic machine model.

Keywords: Noise reduction · Forming machines · Acoustic machine simulation

1 Introduction

In many mechanical engineering areas, innovative high-strength materials are increasingly used. The use of such materials in the vehicle body contributes significantly to the reduction of weight while increasing crash safety.

An essential procedure in sheet metal processing is shear cutting. A large part of the noise emission arises from the pulse-like release of energy after material separation (cutting impact). The sudden drop in cutting force causes the components of the forming machine and the surrounding air to vibrate. In particular when cutting sheet materials with high strength, the sound emission occurs with high intensity, since a high cutting force has to be applied and consequently a high energy is released. This results in sound emission with a broad frequency spectrum and high amplitudes. The generated noise can lead to long-term health damage, which is the reason for a mandatory noise limit during machine operation. In addition, it leads to a reduction in the ability of the operating personnel to concentrate and, overall, to a decrease in productivity, as production times are often restricted for reasons of noise protection.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 45–55, 2021.
https://doi.org/10.1007/978-3-662-62138-7_5

Noise reduction measures on cutting presses are usually directly related to the reduction of excitation or take the form of encapsulation. In [4] and [8], for example, presented hydraulic systems for cutting shock damping achieved reductions in sound pressure levels of up to 10 dB. Further investigations on the reduction of ram vibrations and thus the excitation can be found in [1] and [9]. Secondary noise protection measures include enclosures or sound insulation cabins. With full-automated operating high-speed presses, a complete enclosure of the machine is usually possible [3, 6]. The development or purchase of sound insulation cabins is associated with high costs, while increasing the space requirement and limiting the machine accessibility. Active solutions such as cutting-shock absorbers are expensive, require high maintenance and are not suitable for applications with high stroke rates. In outlook of the newly emerging challenges in the processing of innovative sheet metal materials, the machines used for this must be acoustically optimized.

In this paper a procedure for detecting and analysing noise emissions during the cutting of high-strength materials is presented. Additionally, a systematic identification and a simulation-based evaluation of primary design measures to reduce noise development on the example of an automatic punching press is introduced.

2 Measurement of the Noise Emissions of a Cutting Press

First, the acoustic emissions of an exemplary forming machine are measured during the shearing process. As measuring device an acoustic camera from CAE Software & Systems, consisting of an optical camera and 112 MEMS microphones, is used. The function of the camera is based on the beamforming principle, whereby the localization of the sound sources is carried out by evaluating the runtime differences of the sound signals in relation to each microphone. The measurement considered below were conducted during a cutting process on a 630 kN eccentric press. A closed circular cut was carried out on a stainless steel sheet (X5CrNi18-10) with a tensile strength of approx. 700 N/mm^2 and a sheet thickness of 1.5 mm. Figure 1 shows the experimental setup. This results in a maximum force of about 270 kN and a frequency spectrum of the force excitation up to about 5 kHz.

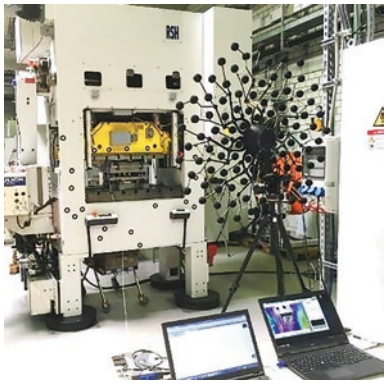


Fig. 1. Measurement setup

In post-processing, the recorded frequency spectrum of the sound pressure level as well as the local and temporal sequences of the sound emissions can be visualized. Figure 2 shows a spectrogram of the cutting process. It provides acoustic information about the events occurring during the cutting process, such as clutch engagement (1) and disengagement (2) of the eccentric shaft, the impact (3) of the blank holder on the sheet metal, the cutting impact (4), scrap dropping noises (5) as well as peripheral noises (e.g. 6). The process spans a broad frequency range beyond the hearing limit. It should be noted that the sound pressure levels relevant to noise protection (above 85 dB) tend to concentrate in the lower frequency range up to approx. 1.5–2 kHz. This could also be observed in another, significantly larger knuckle-joint press with a nominal force of 2000 kN.

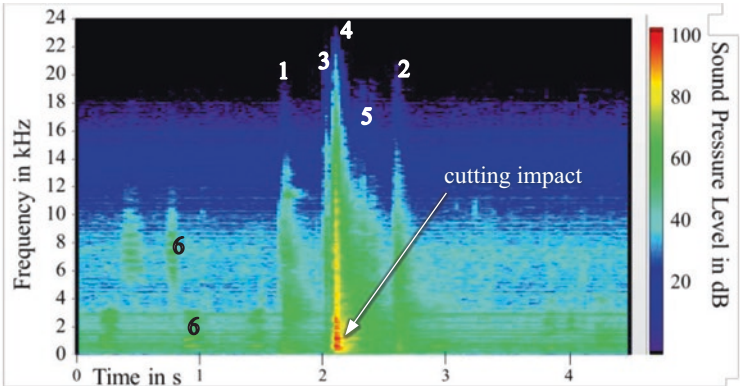


Fig. 2. The recorded spectrogram

The localization of sound emissions is realized by color visualization. Figure 3 shows some selected examples. The impulse that occurs immediately after the material separation is transferred directly to the surrounding air in the workspace. In post-processing, the "loudest" source can be located more precisely with approx. 105 dB, which is located in the contact zone between the cutting punch and the sheet metal edge after the material separation (left).

In addition to the directly radiated emissions, the machine structure is excited to mechanical vibrations, which are transmitted in form of structure-borne noise via the drive train to the surface of the structure (middle, right). These sound emissions occur in a frequency range up to approx. 2 kHz. Also, in the case of an excitation in the higher frequency range (which can be up to 10 kHz, depending on the sheet strength and thickness [2]) the relevant highest levels on the structure were measured in a similar range up to approx. 2 kHz. The "filtering out" of the high-frequency parts of the excitation is related to the transmission behavior of individual elements in the force flow between the process excitation and the radiation surfaces on which factors such as inertia, elasticity and play have an influence. The separation between directly and indirectly radiated emissions in the measurement is difficult. More precise statements can be made later in the context of a simulation of the indirect radiation.

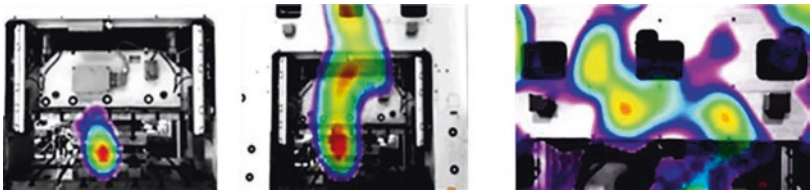


Fig. 3. Visualization of the sound emissions during the cutting process (l: punch/sheet metal contact, m: transmission into the structure and radiation, r: indirect structural emissions)

The measurements were carried out under the same basic conditions (tool, press) for different material strengths and stroke rates. Table 1 provides an excerpt from this for an overview of the maximum pressure levels for the materials used (DC04, DP600 and X5CrNi18-10). It also lists the corresponding maximum cutting force.

Table 1. Measured sound pressure levels with variation of the stroke rate and the sheet material (cutting impact)

	60 spm	200 spm
DC04 (130 kN)	90 dB	92 dB
DP600 (220 kN)	98 dB	100 dB
X5CrNi18-10 (270 kN)	105 dB	107 dB

3 Modeling

3.1 Elastic Multibody Simulation of Structural Dynamics

In order to map the elastodynamics of the considered forming machine the multi-body simulation method is used, whereby the components in the force flow and the sound transmitting or radiating components are modeled elastically [cf. 2].

The model is validated experimentally. The measured curve of the cutting force is used as a model input and the calculated path-time process of the ram is compared with the measured one. Figure 4 shows the simulation result in case of the cutting impact (left) as well as the result of the validation during the cutting process (right). The bending deformation of the structure is largely responsible for the sound radiation [7].

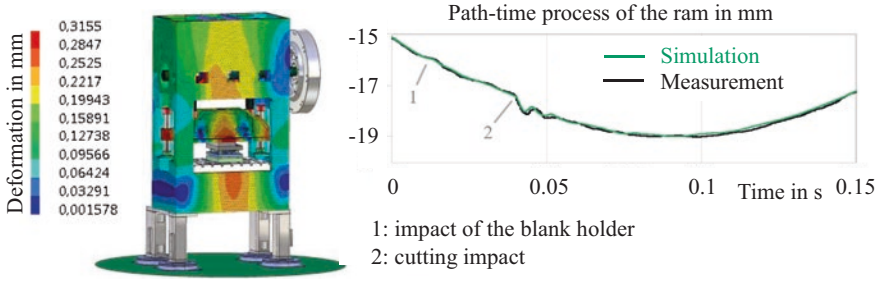


Fig. 4. Machine simulation at cutting impact (left) and result of model validation (right) [cf. 2]

3.2 Mapping of the Machine Acoustics

In case of indirect sound radiation, the acoustic parameters are calculated as a function of the structural loads. These excite mechanical vibrations of the structure components, which are transmitted to the structures surface in form of structure-borne noise and excite the surrounding air. The result is a sound power spectrum, which leads to a directional and distance-dependent sound pressure distribution.

The FE-MBS model is expanded by an acoustic model to convert the structural loads into acoustic quantities. For this purpose, an acoustic 3D fluid grid is created, which encloses the structural components and maps the surrounding air. The excitations from the structural dynamics are recorded via the internal structure-side boundary of the fluid grid, a 2D intermediate network being derived from the structure network for its transmission. The outer boundary of the fluid grid must be complied with the non-reflection boundary condition, so that no reflections distorting the acoustic solution take place at the outer edge of the fluid grid. For this purpose, an artificial reflection-free layer in form of a "Perfect Matched Layer" (PML) is generated [10]. Microphone grids are used to calculate and display the acoustic field variables such as sound pressure. These are 2D or 3D acoustic grids, which can have any shape in the free field around the structure and in which the spread of the field sizes in space is calculated.

The load-results of the structural dynamics simulation are transformed into frequency-dependent loads and are used in a vibroacoustic solution as structural excitation at the force-action positions of the structure. The decoupling between the structural dynamics simulation and the acoustic simulation leads to a significant reduction of computing time. However, the feedback of the air waves to the structure is hereby neglected. Figure 5 illustrates the procedure for acoustic modeling.

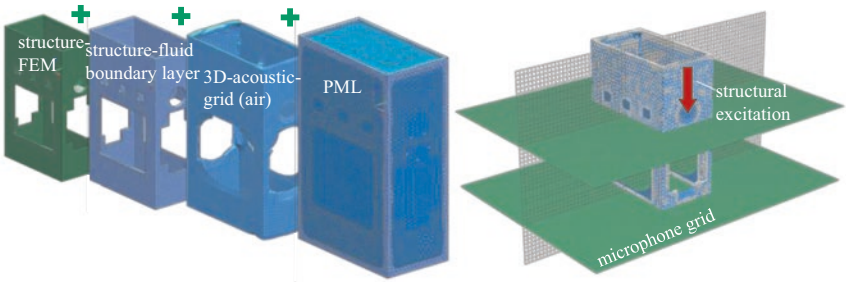


Fig. 5. Acoustic modeling

4 Measures of Noise Reduction

4.1 Basics

The "basic equation of machine acoustics" for indirect sound radiation represents the basis for the derivation of design optimization measures on the press structure [5, 7]:

$$P(f) = \rho_L c_L S \sigma(f) h_T^2(f) F^2(f).$$

This equation is used to illustrate the relationship between the excitation *F(f)* and the resulting sound power *P(f)*, taking relevant influencing variables into account. In addition to the specific impedance of the air $\rho_L c_L$, there are the radiating structure surface *S*, the body transmission admittance $h_T(f)$ and the degree of radiation $\sigma(f)$.

According to the Equation above basic "rules" for reducing the sound power can be derived. These are reduction in excitation (1), reduction in body transmission admittance (2) and reduction in the degree of radiation and the radiation area (3):

1. In theory, all measures that reduce the cutting force or the contact forces between the cutting punch and the sheet metal edge after material separation can be used to reduce the excitation.
2. The body transmission admittance is the opposite of the impedance (resistance) of the structure in terms of sound transmission within the structure. A reduction in admittance can generally be achieved by influencing the eigen behaviour of the structure. This includes shifting the eigenfrequencies into the low-excitation areas, e.g. by the redistribution of the wall thickness.
3. Design measures, through which the sound radiation can be influenced, are the compact design and the increase of the coincidence frequency.

The conversion of the sound power into sound pressure takes place depending on the distance as well as the direction (L_p: sound pressure level, L_W: sound power level, *Q*: directional factor, *r*: distance to the sound source):

$$L_P = L_W - \left| 10 \lg \left(\frac{Q}{4\pi r^2} \right) \right| [\text{dB}].$$

Accordingly, (3) also refers to design rules that favorably influence the direction of radiation. The coincidence frequency is the frequency at which the wavelength of the air coincides with the bending wavelength of the structure, which is unfavorable for the radiation. In order to avoid this, the coincidence frequency should be set as high as possible in relation to the excitation spectrum. The constructive possibilities for this can be derived from the following Equation [7]:

$$f_C = \frac{c_L^2}{2\pi} \sqrt{\frac{m'}{EI'}} \text{ [Hz]} .$$

Here c_L is the speed of sound, m' [kg/m^2] is the area-related mass distribution and EI' [kg m^2/s^2] is the width-related bending stiffness of the structure. Accordingly, a high

mass distribution with constant or reduced bending stiffness leads to an increase of the coincidence frequency.

4.2 Simulation Results

Based on the approaches described above, simulation calculations varying the machine structure were carried out. The sound power spectrum and the resulting location and frequency-dependent sound pressure distribution served as the basis for the analysis and evaluation of the construction measures. According to the measurement results, within the simulations a frequency range between 0.8 and 2 kHz was considered. Following, a selection of the measures is presented using the example of the frame, which represents the largest uncovered radiation area.

Figure 6 left shows a frequency response in case of the initial construction of the machine frame. It illustrates the according amplitude response of the maximal sound power level around the machine frame. In this case, there are distinct amplitude increases resulting from the excitation of the frame by corresponding eigenfrequencies. The high levels unfavorably occur almost over the entire considered frequency range. The representation of the pressure field (Fig. 6 right) allows the localization of the resulting pressure or the pressure level directly on the structure and in the surrounding area. Following, selected pressure level fields will be presented for the frequencies with the highest power level.

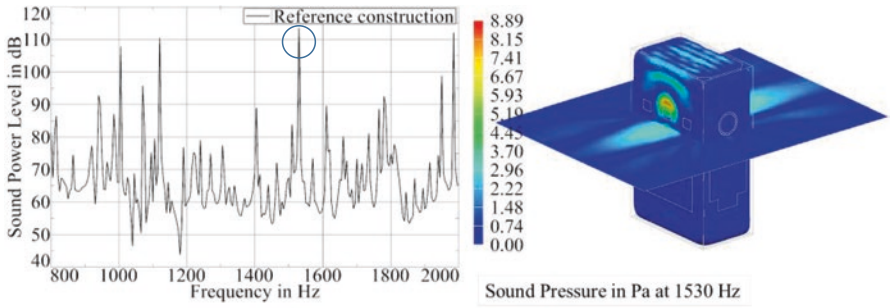


Fig. 6. Frequency response (left) and sound pressure field at 1530 Hz (right)

Figure 7 shows this as an example for the frequencies 1530 Hz and 1985 Hz. The result at 1530 Hz is comparable with the measurement and with the deformation simulation with regard to the relevant radiation area on the upper front wall. A negative aspect is that in this case the highest levels are radiated directly into the operator space. A sound pressure level of approx. 100 dB was calculated at a distance of approx. 1 m, which corresponds approximately to the measurement. At the higher frequencies (Fig. 7, right) the emission sources are no longer locally concentrated, but rather "scattered" over the entire structure. This is natural in the case of the excitation of higher eigenmodes, since here the intensity of the oscillation is widely distributed.

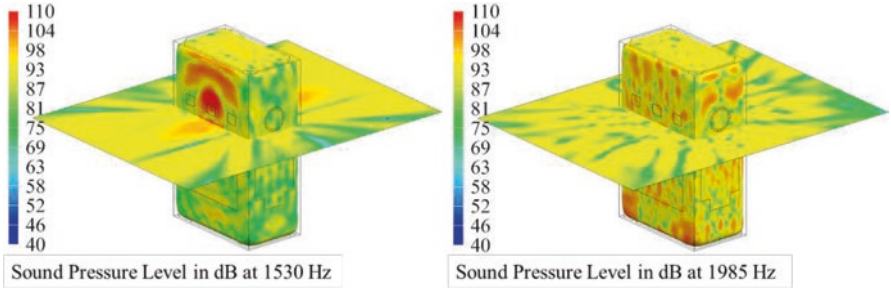


Fig. 7. Sound pressure level field at 1530 Hz (left) and at 1985 Hz (right)

The aim of the investigation is to define measures that influence the amplitude response in a target-oriented manner. This includes that in the relevant frequency range as few as possible increased power peaks occur or that they concentrate in a low-excitation range. Furthermore, the aim is to achieve a pressure distribution that is as evenly as possible in the area surrounding the structure.

From an acoustical point of view, a compact construction without sharp edges should be preferable [5]. Figure 8 shows the result for a construction with rounded edges ($r = 100$ mm). There is a positive influence on the radiation characteristics observable (right). In this case, the direction of radiation is more favorable, particularly with respect to the front of the structure.

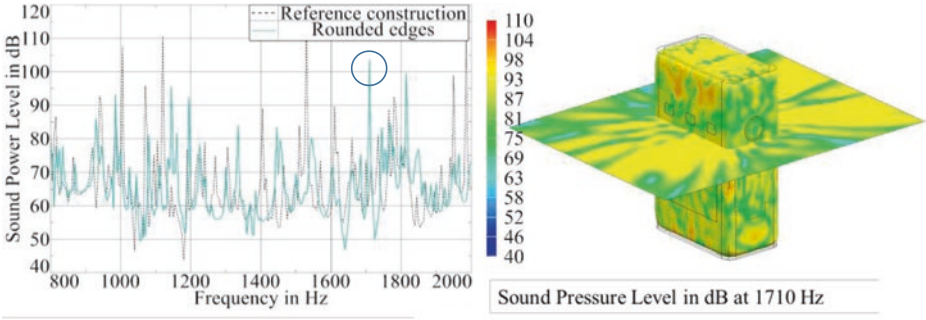


Fig. 8. Frequency response (left) and sound pressure level field at 1710 Hz (right)

The variation in the wall thickness influences the admittance and the degree of radiation in contrast [7]. Figure 9 shows the result of a combination of the rounded edges with an increase in the thickness of the front and rear walls (from 40 to 45 mm). As a result there are comparatively fewer and lower power peaks. The radiation is evenly distributed around the structure.

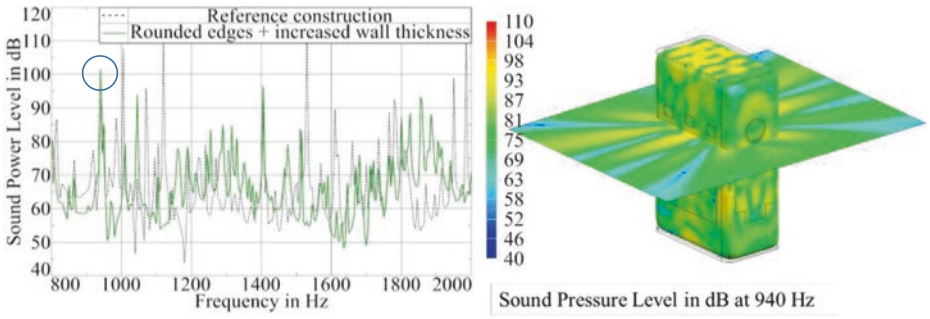


Fig. 9. Frequency response (left) and sound pressure level field at 940 Hz (right)

Local stiffeners are another way of specifically influencing structure-borne noise. Figure 10 shows an exemplary stiffening (left) at the inside edges of the structure and the resulted frequency response (right). Due to an increase in stiffness, the eigenmodes in the upper frequency range dominate as expected.

In general, the stiffening elements increase the eigenfrequencies, which means that they have a positive effect on noise reduction, particularly in the low-frequency range. In a targeted coordination with the relevant range of excitation, significant level reductions can be achieved.

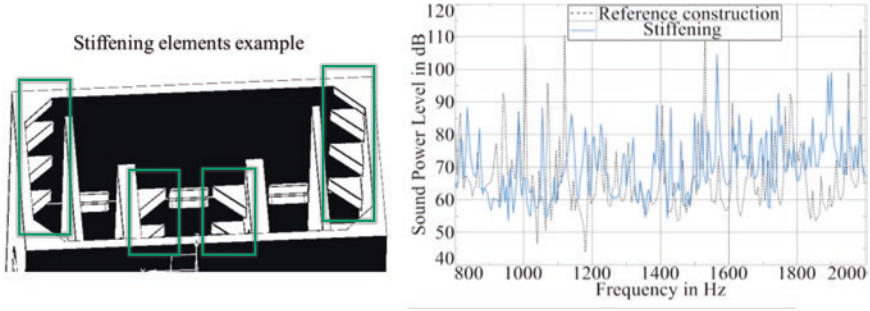


Fig. 10. Exemplary stiffening (left) and resulted frequency response (right)

In order to counteract the negative effect of stiffening in the upper frequency range and at the same time to limit the increase in mass, a further measure to influence the degree of radiation is examined in the following example. Figure 11 shows a result for a design version with an increased mass distribution while reducing the wall thickness and maintaining the shown stiffening from Fig. 10. The combination has the desired effect. The power peaks were largely reduced in the considered frequency range.

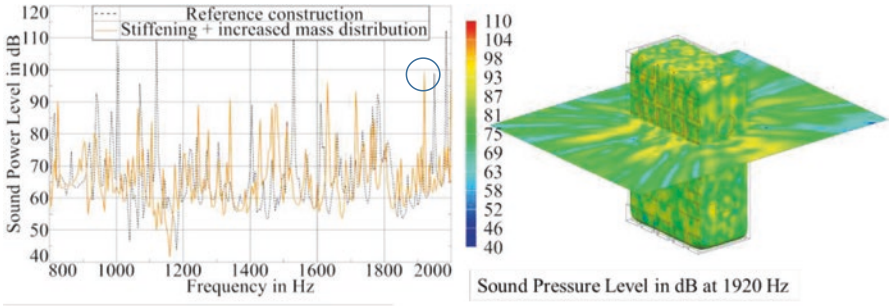


Fig. 11. Frequency response (left) and sound pressure level field at 1920 Hz (right)

5 Summary

In this article a procedure for detecting and analyzing sound emissions during cutting high-strength sheet metal materials, as well as evaluating measures to reduce these emissions, is presented. The investigations were focused on primary design measures that contribute to noise reduction at the machine structure. For the purpose of testing them, an acoustic machine model was created. In the context of simulations, various design measures for noise reduction were implemented and evaluated, of which a selection was presented here.

Individual measures had partially contradictory effects in different frequency ranges. With a suitable combination of measures, a significant reduction of the sound power can be achieved in the entire considered frequency range. In summary, it could be shown that a reduction of the maximum sound pressure level of indirect radiation via the structure is possible by up to 15 dB in the considered frequency range.

Based on these investigations, both constructive modifications and secondary solutions can be realized, which specifically contribute to noise reduction.

Acknowledgement. The IGF-project 19651 N of the European Research Association for Sheet Metal Working (EFB e. V.) is supported via the German Federation of Industrial Research Associations (AiF) within the framework of the Industrial Collective Research (IGF) program by the Federal Ministry of Economics and Technology based on a decision by the German Bundestag.

References

1. Behrens, B.-A., Marthiens, O., Werbs, M.: Electromagnetic dampening of the cutting shock on sheet metal presses. In: 8th International Conference on Technology of Plasticity, ICTP 2005, Verona (2005)
2. Behrens, B.-A., Krimm, R., Kammler, M., Schrödter, J., Wager, C.: Fatigue analysis of a mechanical press by means of the hybrid multi-body simulation. Prod. Eng. **6**, 421–430 (2012)
3. Bockel, G., Seidel, H-J.: Lärmkapselung von Pressen in Blechbearbeitungsbetrieben, Bänder Bleche Rohre (1981)

4. Doege, E., Seidel, H.-J.: Noise reduction on mechanical punch presses. CIRP Ann. **34**(1), 507–509 (1985)
5. Henn, H., Sinambari, G., Fallen, M.: Ingenieurakustik, 3rd edn. Springer Vieweg + Teubner Verlag, Wiesbaden (2001)
6. Hoffmann, H.: Maßnahmen zur Lärmminderung an mechanischen Pressen, Blech Rohre Profile 27 (1980)
7. Kollmann, F.G.: Maschinenakustik. Grundlagen, Meßtechnik, Berechnung, Beeinflussung, 2nd edn. Springer, Berlin (2000)
8. Murakawa, M., Mo, J., Wakatsuki, Y., Koga, N.: Investigation of blanking noise reduction using a hydraulic inertia damper. J. Mater. Process. Technol. **112**(2–3), 205–213 (2001)
9. Otsu, M., Yamagata, C., Osakada, K.: Reduction of blanking noise by controlling press motion. CIRP Ann. **52**(1), 245–248 (2003)
10. https://docs.plm.automation.siemens.com/tdoc/nx/1847/nx_help. Accessed 20 Apr 2020

Equal-Channel-Angular-Swaging for the Production of Medical Implants Made of Fine-Grained Titanium

L. Kluy(✉), F. Chi, and P. Groche

Institute for Production Engineering and Forming Machines, Otto-Berndt-Str. 2, 64287 Darmstadt, Germany
lukas.kluy@ptu.tu-darmstadt.de

Abstract. Demands of the aging society shift the focus of production engineering towards challenges in medical technology. An increasing number of people are equipped with implants, but incompatibilities can lead to implant loosening. These incompatibilities can be avoided by implants based on fine-grained titanium which stimulate interactions between implant and bone tissue and thus activate bone healing.

The continuous production process Equal-Channel-Angular-Swaging (ECAS) is applied to achieve grain refinement in titanium rods. Shear strains during forming are therefore numerically investigated. Processing of titanium with conventional ECAS leads to unfilled die cavity. Thus undesired bending occurs, both in simulation and experiment. The bending stresses result in material failure. Additionally, the achieved shear strains are inhomogeneously distributed in the cross-section.

It is shown that a hydrostatic stress state created by counter pressure improves the failure limit, yields in more homogenous shear strain distributions and enables a continuous production of fine-grained titanium.

Keywords: Fine-grained titanium · SPD processes · Medical implants

1 Introduction

Around 1 million dental implants with a market volume of 3.2 billion euro are implanted in Germany every year for enhancing vitality and quality of life of patients [1]. Conventional dental implants are associated with unsolved problems, particularly bone loss, implant fractures and screw loosening. Within the first ten years, 8% of all patients lose their implants [2]. This not only seriously impairs the quality of life, but also has a great economic impact due to follow-up and treatment costs. Dental implants are mostly made of titanium (Ti) due to high demands on corrosion resistance, fatigue strength and biocompatibility. Elements such as vanadium (V) and aluminum (Al) are added into current Ti alloys to fit the mentioned demands, which are, however, critical for the medical uses due to toxic reactions. Furthermore, the

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 56–65, 2021.
https://doi.org/10.1007/978-3-662-62138-7_6

Young's modulus of current Ti alloys leads to bone resorption at the contact points between implant and bone material, which leads to Periimplantitis and loosening of the implants [3].

A promising approach is the use of fine-grained Ti for medical implants. On the one hand, it was shown that the fatigue strength and corrosion resistance of fine-grained Ti can be enhanced compared to the coarse-grained Ti, so that the requirements of fatigue strength and corrosion resistance for medical applications can be fulfilled without critical elements, such as V and Al [4]. On the other hand, the fine structured surfaces stimulate interactions between the implant and the surrounding body cells. The bone cells grow to the grain boundary grooves on the surface of an implant and cross-link, if the distance between grooves is smaller than 1 µm. Hence, bone healing is expected to occur when grain sizes under 1 µm are existent [5].

One way to achieve the desired small grain sizes is to refine the microstructure by severe-plastic-deformation (SPD) processes [6]. These processes can lead to grain sizes with characteristic values below 1 µm. A special feature of SPD processes is the unaltered shape of the workpiece. This allows the process to be repeated several times with the same workpiece to achieve very large strains incrementally. Well-known SPD processes include High Pressure Torsion (HPT) [7, 8], Accumulative Roll Bonding (ARB) [9] and Equal Channel Angular Pressing (ECAP) [10, 11]. Among the SPD processes, ECAP is the most efficient one for processing rod workpieces. Stolyarov succeeds in producing ultrafine-grained (UFG) Ti alloys with an ECAP process [12]. In these studies, it is also shown that corrosion resistance and fatigue strength are enhanced.

For increasing the productivity of UFG materials, the design of ECAP process is extended with continuous process operation [11, 13, 14]. However, productions of tubular UFG material are still not realized in industrial scale up to now due to the process uncertainties caused by high process forces and tool wear. Görtan developed the process of Equal-Channel-Angular-Swaging (ECAS) [15]. The forming behavior of ECAS is similar to the ECAP process, but the die is divided into two halves, which oscillate towards each other. Consequently, friction forces and thus axial forces are significantly reduced by ECAS compared to ECAP. Previous studies demonstrated, that UFG materials can be continuously produced out of conventional copper, steel and FeCo alloy via grain refinement [15].

The materials formed by ECAS so far exhibit significantly higher formability than Ti at room temperature. Not surprisingly, fracture occurs during processing of Ti by conventional ECAS. In the paper at hand, the ECAS process is therefore extended to the production of fine-grained Ti material. Specifically, it is investigated under which process conditions the material is homogenously stressed by shear strain without failure.

2 Approach

The approach to produce fine-grained Ti via ECAS is shown in the following chapter. For this purpose, the forming behavior of Ti as well as the correlation of process and material parameters are depicted.

First, a numerical simulation is used to investigate the shear strain during the forming process with various counter pressures. Second, the results are used to derive requirements for tool modifications. Third, the necessary counter pressure device is implemented and validated in experimental investigations. The investigated ECAS process installed on a high-speed stamping press is shown in Fig. 1.

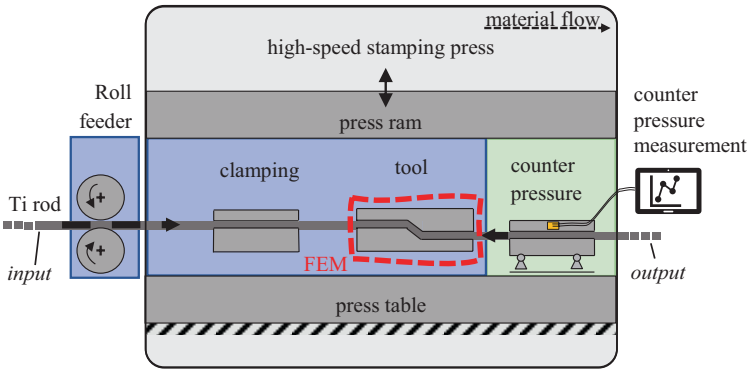


Fig. 1. ECAS Process on a high-speed stamping press for the production of fine-grained Ti

The initial material is a commercially coarse-grained pure Ti grade 1 (3.7025) rod. It is discontinuously moved forward by a roll feeder and clamped in the ECAS tool while the high-speed press ram moves to the bottom dead center. Simultaneously, the material is stressed by shear in the tool jaws and experiences a counter pressure in the direction opposite to the feeding direction. The events taking place in the tool section are modeled by FEM, since the desired grain-refinement takes place here. The counter pressure is varied in the simulations and measured in the experiments with the aid of strain gauges. The material on the output side is an ECAS processed Ti rod with the same diameter as the initial diameter.

3 Numerical and Experimental Setup

In the present paper, a continuous production process for fine-grained Ti is determined. The influence of counter pressure is evaluated to achieve grain refinement by shear strain.

To predict the process window for Ti in the ECAS process, numerical simulation by means of the finite element method with the Abaqus CAE 2019 software is used. Previous work has shown that the strain values and predicted forces are almost identical for 3D FEM and 2D FEM of ECAS [15]. Nevertheless, it should be noted that friction can lead to a three-dimensional deformation inhomogeneity, which cannot be operated in this simulation and is a focus of future research [16]. The abstraction to a 2D problem using the symmetry plane of the rod as shown in Fig. 2, leads to significantly shorter computational time. The sample is 50 mm long and the diameter is 8 mm. An isotropic elastic-plastic material behavior based on von Mises model is

used. The flow curve is derived from a cylinder compression test according to DIN 50106 and a tensile test according to DIN 6892. Because ECAS introduces strains higher than 1 into the material the flow behavior is extrapolated according to the Swifts strain hardening material model from the compression test based on the strain hardening exponent $n = 0.14$, poisson's ratio 0.34, tensile strength $R_m = 487$ MPa, yield strength $R_{p0.2} = 361$ MPa, stress amplitude $C = 737$ MPa and initial strain $\varphi_0 = 0.006$.

The dimensions of the upper and lower tool jaws correspond to reality. A coefficient of friction of 0.1 is determined by a sliding compression test for the tangential, isotropic kinematic contact model between the specimen and the top and bottom of the tools. Hard contact is selected perpendicular to the tool surface, since the forming forces are transmitted here. The lower tool jaw is fixed. The upper tool jaw performs a vertical stroke of 16 mm with a frequency of 100 stroke/min. In the simulation, this movement is implemented via a sinusoidal function. The roll feeder moves the sample 0.1 mm/stroke in the top dead center of the high-speed press, which is implemented as a displacement boundary condition via step function with acceleration and deceleration phase.

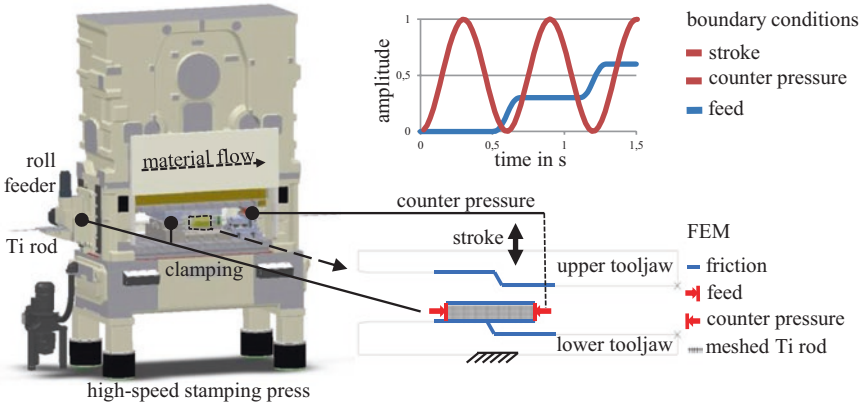


Fig. 2. Experimental (left) and numerical (right) setup for the production of fine-grained Ti via ECAS with counter pressure on a high-speed stamping press. The boundary conditions and the simplification for the FEM are shown on the right.

The counter force is varied from 0 kN to 14 kN in a step width of 2 kN. An explicit solver is used, and the forming process is simulated for a duration of 60 s to achieve conditions of a continuous process. The shear strain distribution in the forming zone is evaluated.

For validation, analytical, numerical and experimental results are compared. According to Görtan et al. Eq. 1 can be used to determine the total strain analytically. The total strain ϵ_N is based on the number of shear planes N and the angles Φ and Ψ of the tool as shown in Fig. 3 [15].

$$\epsilon_N = \frac{N}{\sqrt{3}} \cdot \left[2 \cot\left(\frac{\Phi}{2} + \frac{\psi}{2}\right) + \psi \operatorname{cosec}\left(\frac{\Phi}{2} + \frac{\psi}{2}\right) \right] \tag{1}$$

The geometry of the model with $N = 2$, $\Phi = 0$ und $\psi = \frac{2}{3}\pi$ results in an analytically determined total strain of $\epsilon_{N,a} = 1.46$. As shown in Fig. 3, angle ψ is the forming angle of the tool. Angle Φ specifies the orientation with which the rod leaves the tool. The rod has the same orientation in the initial and output state. The average simulated total strain is $\epsilon_{N,s} = 1.41$ along the path. Figure 3 shows the numerical total along the path as well as the numerical average and the analytical value. A deviation of 4% can be asserted.

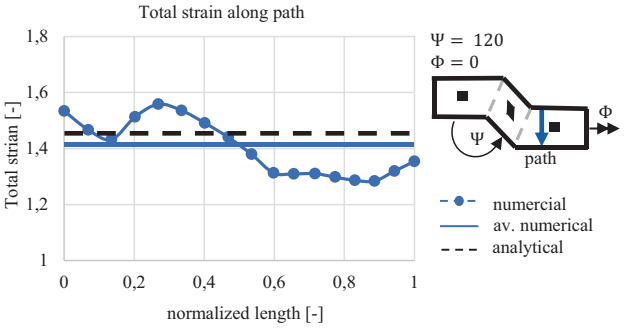


Fig. 3. The analytically and numerically determined total strain along path. A minor deviation can be asserted between analytical and numerical calculation.

In addition, geometric features are compared in simulation and experiment as seen in Fig. 4. The diameters are measured with a caliper and compared at the same points in the simulation. Simulation and experiments show an unfilled tool and bending so that the angles can be compared according to DIN 1319.

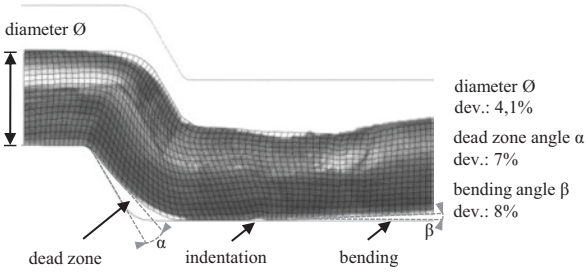


Fig. 4. Validation of the simulation by geometric features

The diameter has an average deviation between simulation and experiment of 4%, the angle in the dead zone a deviation of 7% and the bending angle a deviation of 8%.

4 Results and Discussion

Figure 5 depicts the mesh of the rod and the contour of the tool for different counter forces. The workpiece is subdivided into inlet, shear and outlet zone.

In the inlet zone material behaves in the same way independent of the level of the counter force. In the shear zone the mesh is deformed into a diamond shape. The unfilled tool volume on the underside is particularly noticeable without counter force. With increasing counter force the material fills the tools more and more. In the outlet zone, the angle at which the cells remain in the deformed status is specified. The aim is to obtain a rectangular mesh to ensure a homogeneous forming of the rod, and the ECAS process can be repeated in a following step. It has been shown that multiple passes in the SPD process are beneficial for the creation of an ultra-fine-grained structure [15]. In the outlet zone a homogeneous and an inhomogeneous strain distribution caused by running-in characteristics through the first stroke is formed. The first stroke leads to a high deformation of the material as the upper tool jaw impinges the rod and forms an indentation on the top. Without counter force, the mesh in the outlet zone is not rectangular. The change in length Δl is due to the volume constancy assumed for the plastic deformation. In the outlet zone the rod bends upwards without counter force which leads to undesired tensile stress on the surface of the rod. From 6 kN upwards the tool is filled and the mesh is approximately rectangular.

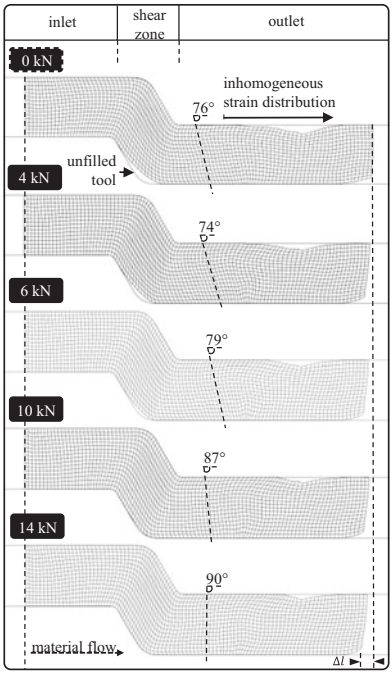


Fig. 5. Influence of counter force from 0 kN to 14 kN on ECAS process in simulation for Ti, subdivided into inlet, shear and outlet zone. The net shows a filled tool from 6 kN upwards.

Shear strains during forming are of high relevance for grain refinement and therefore numerically investigated. The shear strain over the normalized length from the top to the bottom of the sample in the shear zone along the path is shown in Fig. 6.

The shear strain without counter force is inhomogeneously distributed in the cross-section after the first shear zone. In particular, the maximum tensile stresses are determined at 609 MPa and thus lie above the tensile strength $R_m = 487$ MPa, so that material failure is predicted. In the experiments, this state of stress results in cracking. If the counter force is increased, the shear strain is homogenized. From a counter force of 6 kN, tensile stresses are below tensile strength and shift into compression stress state. The counter force prevents material failure and leads to the desired homogenous shear strain distribution in the shear zone.

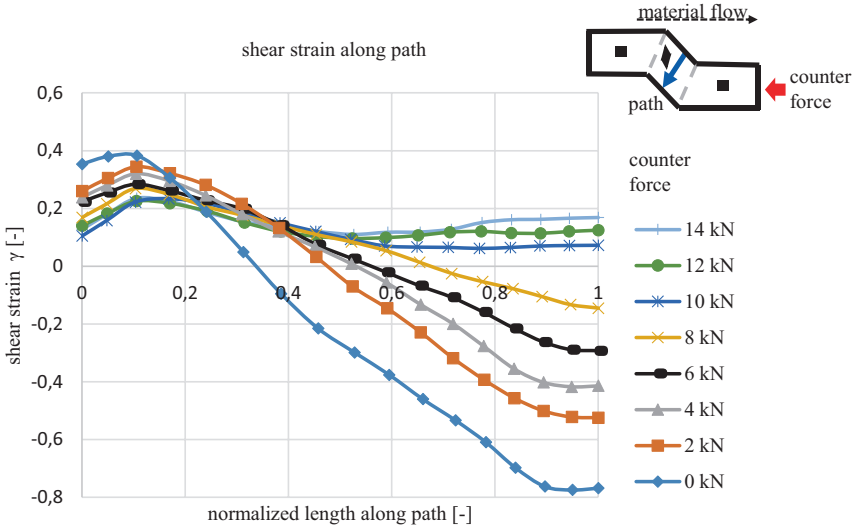


Fig. 6. Shear strain in the forming zone depending on the level of counter force. If the counter force is increased, the shear strain is homogenized.

The stress state in the direction of material flow is depicted in Fig. 7 to underline the experimental results for the processed Ti rod as shown in Fig. 8.

Without counter force, the rod is subjected to bending. The material fails and the fracture surfaces show a typical material failure due to bending. The crack starts on the underside of the sample near the unfilled tool as predicted via FEM because of high tensile stresses. When counter pressure is applied, a hydrostatic stress state increases the formability of the material and fine-grained Ti is processed successfully.

Tensile stresses on the upper surface of the rod cause the material to tear open in the surface region (shear marks). The authors assume that the shear marks can be reduced or even avoided by reduced feed speed and optimized lubrication. The upper limit of the counter force is confined by the buckling of the rod at 10 kN contrary feed direction.

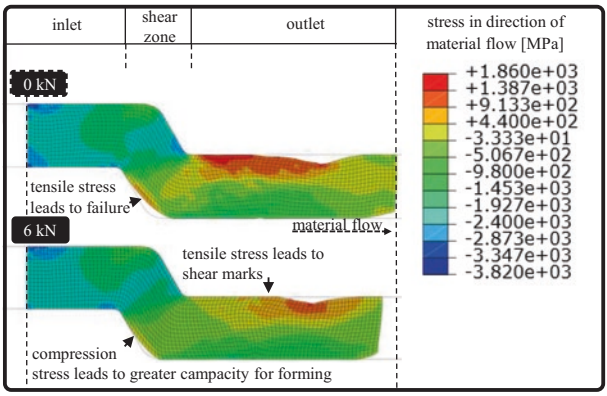


Fig. 7. Stress in the direction of flow. From a counter force of 6 kN, tensile stresses are below tensile strength and shift into compression stress state in the shear zone.

As shown in Fig. 8 in the microstructure, the grain size after ECAS is smaller than before ECAS, so that grain refinement has occurred. However, there are still some coarse grains after one path visible, so multiple paths are required to reach the targeted grain structure.

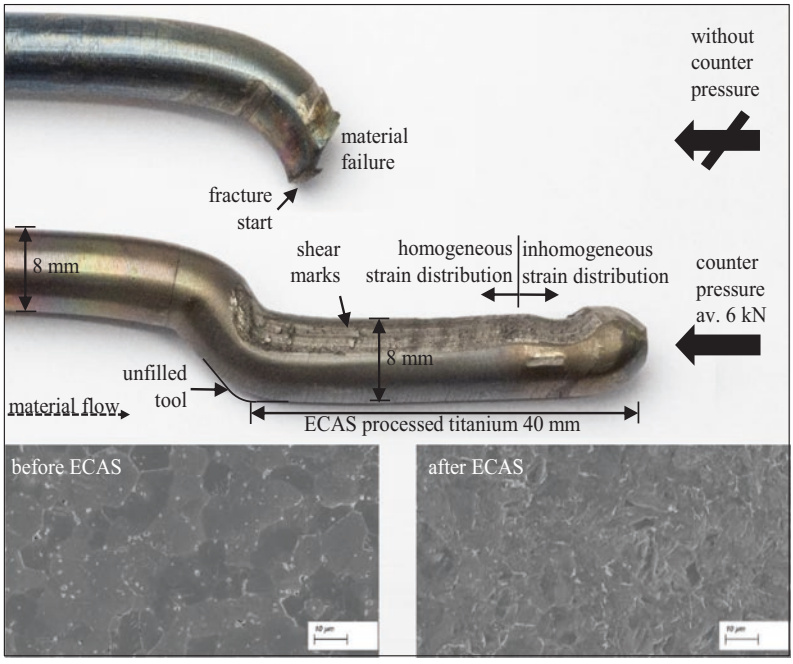


Fig. 8. Material fails without counter pressure (top). With a counter pressure of av. 6 kN, ECAS-Ti is produced successfully (middle). The microstructure shows a grain refinement.

As a result, the simulations show that a high counter force leads to a homogeneous shear strain distribution through tool filling. The height of the counter pressure is limited by the danger of buckling of the Ti rod.

From the investigations, it can be concluded that an axial counter force between 6 kN and 10 kN leads to a more homogeneous shear strain and a hydrostatic stress state which allows the continuous production of fine-grained Ti via the ECAS process.

5 Summary and Outlook

It is shown that counter pressure enables the continuous production of fine-grained Ti via Equal-Channel-Angular-Swaging. The counter pressure significantly influences the shear strain during forming, which is responsible for the mechanism of grain refinement.

An axial counter force between 6 kN and 10 kN leads to a hydrostatic stress state, filling of the tools and homogeneous shear strain that promotes the formation of a homogeneous fine-grain structure. The required parameters and stresses for the forming of fine-grained Ti are determined by FEM and are demonstrated in experiments. The counter pressure prevents bending of the rod, which reduces material stress and thus material failure. The hydrostatic stress state in the material leads to a greater capacity for forming.

Now guidelines for production of fine-grained Ti can be given for the development of new medical implants on an industrial scale. A further grain refinement can be achieved by multiple-pass processing, and parameters such as heat-treatment and feed speed to control the fine-grained structure will be investigated in ongoing projects.

Acknowledgements. The authors would like to thank the German Federal Ministry of Education and Research for funding the project 13XP5093D (IdentiTI) in which the investigations performed in this paper took place and the Institute for Materials of TU Braunschweig.

References

1. Schwarz, F.: Zahnimplantate sind für Patienten zunehmend erste Wahl. In: 32. Kongress der DGI, Wiesbaden (2018)
2. Derks, J., Schaller, D., Håkansson, J., et al.: Effectiveness of implant therapy analyzed in a Swedish population: prevalence of peri-implantitis. J. Dent. Res. **95**, 43–49 (2016)
3. Geetha, M., Singh, A., Asokamani, R., et al.: Ti based biomaterials, the ultimate choice for orthopaedic implants – a review. Prog. Mater. Sci. **54**, 397–425 (2009)
4. Polyakov, A., Semenova, I., Valiev, R.: High fatigue strength and enhanced biocompatibility of UFG CP Ti for medical innovative applications. IOP Conf. Ser. Mater. Sci. Eng. **63**, 12113 (2014)
5. Kim, T., Balakrishnan, A., Lee, B., et al.: In vitro biocompatibility of equal channel angular processed (ECAP) titanium. Biomed. Mater. **2**, 117–120 (2007)
6. Azushima, A., Kopp, R., Korhonen, A., et al.: Severe plastic deformation (SPD) processes for metals. CIRP Ann. **57**(2), 716–735 (2008)

7. Valiev, R., Krasilnikov, N., Tsenev, N.: Plastic deformation of alloys with submicron-grained structure. Mater. Sci. Eng. A **137**, 35–40 (1991)
8. Valiev, R.: Structure and mechanical properties of ultrafine-grained metals. Mater. Sci. Eng. A **234–236**, 59–66 (1997)
9. Saito, Y., Utsunomiya, H., Tsuji, N., et al.: Novel ultra-high straining process for bulk materials - development of the accumulative roll-bonding (ARB) process. Acta Mater. **47**(2), 579–583 (1999)
10. Segal, V.: Severe plastic deformation: simple shear versus pure shear. Mater. Sci. Eng. A **338**(1–2), 331–344 (2002)
11. Valiev, R., Langdon, T.: Principles of equal-channel angular pressing as a processing tool for grain refinement. Prog. Mater Sci. **51**(7), 881–981 (2006)
12. Stolyarov, V., Zhu, Y.: Microstructure and properties of pure Ti processed by ECAP and cold extrusion. Mater. Sci. Eng. A **303**(1–2), 82–89 (2001)
13. Lowe, T., Valiev, R.: The use of severe plastic deformation techniques in grain refinement. JOM **56**(10), 64–68 (2004)
14. Azushima, A., Aoki, K.: Properties of ultrafine-grained steel by repeated shear deformation of side extrusion process. Mater. Sci. Eng. A **337**(1–2), 45–49 (2002)
15. Görtan, M.: Severe plastic deformation of metallic materials by equal channel angular swaging: theory, experiment and numerical simulation. In: Dissertation, Technical University Darmstadt (2014)
16. Suo, T., Li, Y., Guo, Y., Liu, Y.: The simulation of deformation distribution during ECAP using 3D finite element method. Mater. Sci. Eng. A **432**, 269–274 (2006)

Numerical Development of a Tooling System for the Co-extrusion of Asymmetric Compound Profiles on a Laboratory Scale

N. Heimes[1(✉)], J. Uhe[1], S. E. Thürer[2], H. Wester[1], H. J. Maier[2], C. Klose[2], and B.-A. Behrens[1]

[1] Institut für Umformtechnik und Umformmaschinen, Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany
heimes@ifum.uni-hannover.de
[2] Institut für Werkstoffkunde, Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany

Abstract. In order to reduce the weight of vehicles and their CO_2-emissions and to increase driving range, new ways of producing hybrid material compounds must be developed to tailor the properties of the part exactly to the intended use. Lateral angular co-extrusion (LACE) offers the possibility to produce hybrid profiles. In the past, this process was used to extrude flat magnesium-titanium profiles and after further development, co-axial aluminium steel profiles.

In this study, the numerical development of a tooling system for the production of asymmetric hybrid semi-finished products extruded using LACE is presented. The co-extruded profiles, consisting of an L-profile made of steel (AISI 5120) that is filled with aluminium alloy EN AW-6082 on one side, will subsequently be formed to hybrid transverse control arms by die forging. The tool design is initially carried out for a laboratory scale extrusion in order to gain basic knowledge about the process and to quantify the influence of the different process variables like ram velocity or extrusion ratio.

Keywords: Co-extrusion · FEM · Tailored forming · Aluminium-steel compound

1 Introduction

Depending on the volume ratio, a hybrid component made of light metal and steel is significantly lighter than mono-material parts made of steel, while still offering sufficient stiffness. One possibility of combining light metals, e.g. aluminium and steel is co-extrusion. Co-extrusion processes can be divided into two main categories according to the type of extrusion billet [1]. On the one hand, modified extrusion billets with reinforcing elements integrated in the billet matrix material or hybrid extrusion billets can been used. In this case, both matrix material and reinforcement

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 66–75, 2021.
https://doi.org/10.1007/978-3-662-62138-7_7

are plastically deformed. For example, Foydl *et al.* investigated the extrusion of discontinuously reinforced extrusion billets of the aluminium alloy EN AW-6060 which were drilled along the symmetry axis in order to fill in cylindrical, conical or spherical steel elements [2]. The second possibility is the use of conventional extrusion billets, which are combined with reinforcing elements like steel wires or flat profiles. The reinforcing elements are fed to the deformation zone from outside the extrusion tool. Consequently, only the matrix material is plastically deformed. Chatti *et al.* used modified chamber tools to extrude aluminium profiles from EN AW-6060 together with wires made of AISI 304 steel [3]. Pietzka *et al.* succeeded in embedding steel wires in a magnesium matrix and was able to embed up to eleven wires in a co-extruded profile [4]. Grittner *et al.* investigated the lateral angular co-extrusion of flat aluminium titanium compound profiles using conventional aluminium billets. Therefore, the aluminium alloy was redirected at an angle of 90° within the tool and joint with the titanium profile in the welding chamber [5]. Based on the process described by Grittner *et al.* a LACE process for aluminium EN AW-6082 and steel AISI 5210 was developed to produce coaxial aluminium steel profiles which, are formed to hybrid bearing bushings in the further course of the process chain within the Collaborative Research Centre 1153, see Fig. 1 [6, 7].

In order to expand the range of applications for components that can be produced by LACE and to demonstrate the potential of the technology, a co-extrusion process is to be designed for asymmetrical hybrid profiles, which are subsequently formed into hybrid control arms by die forging, see Fig. 1. Based on the geometry of the transverse control arm, the intermediate stages after die forging and extrusion were designed inversely. For the production of hybrid asymmetrical profiles it is particularly challenging that there is no comprehensive form closure. The bond is mainly formed by material closure. At this, it is of great importance to consider the geometrical and process-related limits of the different technologies.

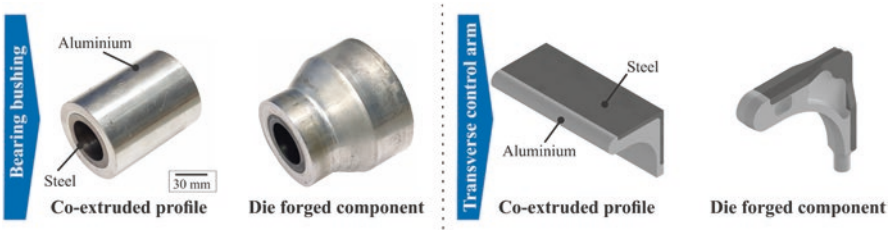


Fig. 1. Co-extruded semi-finished profile and die forged component, left: bearing bushing, right: transverse control arm

Numerical modelling has become state-of-the-art as an initial step in designing and further improving extrusion and co-extrusion processes as it enables to study the material flow and influence of process variables under various conditions and, thus reduces the number of experimental tests and saves resources [8, 9]. In addition, the actual values of local parameters like temperature or contact pressure are often difficult to obtain or not measurable during extrusion processes. In the present study, an

FE model of the LACE process to produce asymmetrical hybrid profiles is established to investigate a possible tool design for co-extrusion experiments on a laboratory scale and to gain knowledge about the influence of the tool geometry and process variables on the resulting hybrid profile.

2 Numerical Investigations

2.1 Model Setup

The commercial FE software FORGE NxT 2.1 was used to model the co-extrusion process. To limit the computation time, the existing symmetry was considered and only the half of the 3D FE model was examined, see Fig. 2a. The billet is placed in the container and pressed into the upper die *via* the press ram. At this point, the aluminium contacts the wedge, which redirects the material flow by 90° to the extrusion direction. The wedge does not only deflects the flow of aluminium, but together with the lower die it forms a feed channel, which guides the L-profile into the welding chamber. The feed occurs at an angle of 80° to the extrusion direction. The L-profile is filled with aluminium in the welding chamber and leaves the tool as a hybrid profile *via* the die opening.

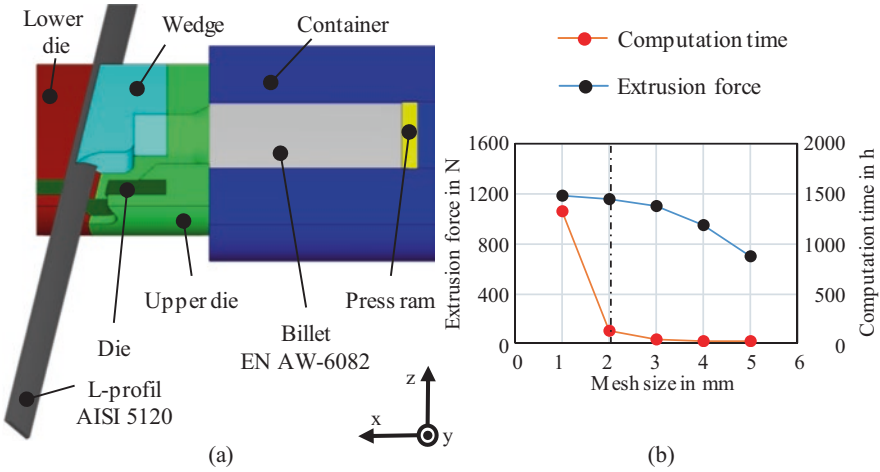


Fig. 2. Half 3D FE model of the assembly (a) Mesh study (b)

Both, the accuracy of the calculated results and the computation time are strongly dependent on the selected minimal element size. Based on the results of a preliminary mesh study, a minimum element size of 2 mm was used for the numerical investigations resulting in 177,000 tetrahedron volume elements in total. Thus, a sufficiently high accuracy and a reasonable computing time of approx. 125 h was achieved, see Fig. 2b. A further refinement of the mesh increases the simulation time by a factor of eleven, with only minimal increase in the accuracy of the results. For the mesh study,

only the mesh size of the aluminium was used, as it has a significant influence on the resulting extrusion force, which is the most important criteria for the design of the process. Since the L-profile is very thin and long compared to aluminium, a mesh size of 1.5 mm was chosen. In the area of the chamber tool and the die a refinement box was used to reduce the mesh size to 0.5 mm. With six elements over the profile thickness the results over the thickness can be represented sufficiently accurate. The tools were modelled as rigid bodies for the described investigations.

Due to extensive plastic deformation during the co-extrusion process, two remeshing criteria, a periodic initiated remeshing criterion with a fixed remeshing after 20 steps as well as an automatic size criterion to refine the mesh of the workpiece according to the curvature of the die in the contact area, were applied. Tresca's friction model was used to describe the frictional behaviour between billet and tools. According to the findings of a previous study, the friction factor was set to 0.95, which describes the high adhesion tendency of aluminium well [10]. A bilateral sticking condition was assumed for the interface between the aluminium billet and the L-profile, since the L-profile is only displaced due to the contact with the aluminium. For the heat transfer coefficient, a value of 35 kW/m^2K was chosen according to literature [11]. The ambient temperature was set to 50 °C. For an accurate prediction of thermomechanical material behaviour during the co-extrusion process, flow curves of both the aluminium alloy EN AW–6082 and the case hardening steel AISI 5120 were experimentally determined by means of uniaxial cylindrical upsetting tests and implemented in the FE software as a function of strain, strain rate, and forming temperature [10].

A numerical process design study was performed to determine a suitable tool geometry and process parameters for the experimental investigations. Therefore, different tool geometries and their influence on the material flow were analysed. After the selection of a suitable tool geometry, the extrusion ratios were varied to determine their influence on the extrusion force and the profile geometry. The extrusion ratio was calculated from the cross sectional area of the aluminium billet at the beginning of the process and the aluminium cross section of the hybrid profile. In order to vary the extrusion ratio, the contours of the die were adapted while maintaining the same L-profile thickness. By varying the profile thickness, the aluminium surface is changed in the profile cross section. To ensure a better comparability of the results of different reinforcement contents, the die contour must be adjusted accordingly so that the extrusion ratio remains the same. In addition, the influence of the ram velocity on the extrusion force requirement was investigated. Finally, the die contour was varied to improve the bond formation between aluminium and steel by clamping the L-profile. The process parameters used for the numerical-parametric study are summarised in Table 1.

Table 1. Process dimensions and parameters

Dimension	Value	Parameter	Value
Billet length	120 mm	Billet temperature	530 °C
Billet diameter	56 mm	L-profile temperature	20 °C
L-profile length	800 mm	Tool temperature	450 °C
L-profile thickness t_h	3, 5 mm	Ram velocity v	2, 4 mm/s
L-profile legs length	20 mm	Extrusion ratio ψ	3.8:1, 5.3:1, 8.1:1

2.2 Results and Discussion

First, the geometries of the chamber tool, more specifically the wedge and upper die geometry, were varied. Figure 3a shows tool variant A, where a flat redirection of the material flow was chosen for the wedge in combination with a slope at the upper die as transition to the die. In variant B, the redirection of the material flow consists of a hemisphere and the transition from the upper die to the die consists of a radius, see Fig. 3b. The evaluation of the velocity in z-direction shows that in variant A a more pronounced dead zone is formed at the wedge. A larger dead zone is desirable to ensure that the aluminium from inside the billet flows and, thus no impurities such as oxides or lubricants enter the joining zone between aluminium and steel. At the transition from the upper die to the flat die, the variant B forms a larger dead zone. However, a dead zone is not necessary there, as the material is removed in the finishing machining process.

Decoupled simulations of tool load were carried out to identify possible weak points within the components. For both variants, the 1st principal stresses do not show any extreme values that would indicate a risk of crack formation. The evaluation of the force-time curves also shows no significant differences. Due to the material flow and easier manufacturing, variant A with a flat wedge and sloping transition was selected for further investigation.

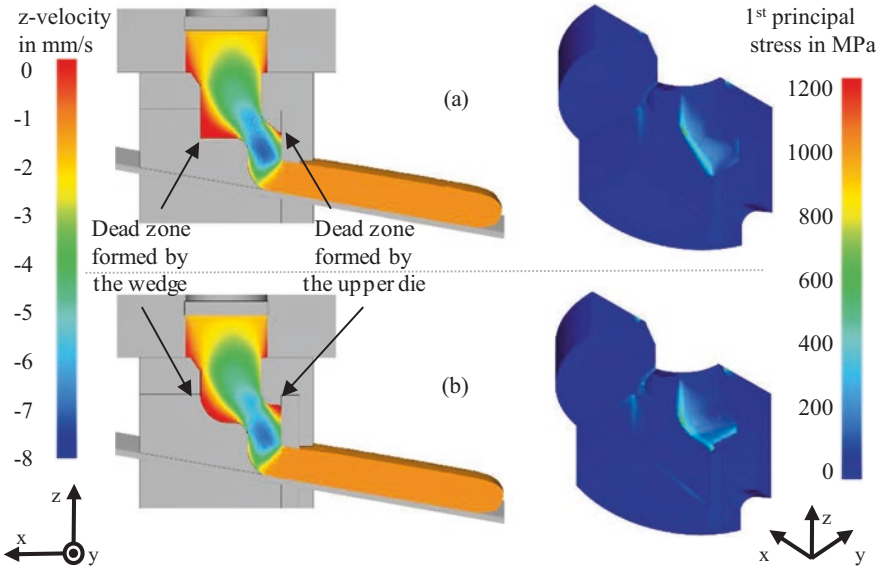


Fig. 3. Velocity in z-direction and die analysis for tool variant A (a) and variant B (b)

In the next step, the extrusion ratio was varied by changing the aluminum content while maintaining a constant L-profile thickness of 3 mm. Figure 4a–d show the extruded profiles in side view with the respective temperature field, in Fig. 4a,b,d the

ideally straight L-profile is visualised in pink to illustrate the bending. The smallest extrusion ratio of 3.8:1 shows a slight bending of the profile in negative z-direction as shown in Fig. 4a. With a higher extrusion ratio to 5.3:1 the bending of the profile increases, see Fig. 4b. At the maximum extrusion ratio of 8.1:1, the L-profile was deformed until necking occurs which leads to an abortion of the computation. The contact temperature is the equal for all variants at approx. 460 °C. The L-profile heats up very quickly due to the small profile thickness. With an extrusion ratio of 3.8:1, the profile heats up faster in the feed channel than with an extrusion ratio of 5.3:1. The varying heating is due to the different exit speeds. As the extrusion ratio increases, the exit speed of the profile increases and, thus reduces the contact time in the feed channel. The increased temperature of the profile in the channel in Fig. 4c is due to the fact that the temperature is shown at a press height H_p of 60 mm. Due to the tool contact in the channel, the L-profile always heats up in the same way, regardless of the extrusion ratio, until the aluminium fills the welding chamber. Subsequently, the L-profile is pressed out of the die together with the aluminium so that the following L-profile has less time in the channel to heat up. The higher deformation of the L-profile for larger extrusion ratios correlates with the increased contact normal stress. Contact normal stresses of up to 600 MPa were determined for the extrusion ratio ψ of 8.1:1, for ψ of 5.3:1 stresses up to 460 MPa and for the smallest extrusion ratio ψ of 3.8:1 a contact normal stress of 370 MPa was determined. The flow stress of AISI 5120 at the present contact temperature of 460 °C is 550 ± 30 MPa [10]. The deformation of the L-profile at the extrusion ratio of 8.1:1 can therefore be explained by exceeding the flow stress of the material. Increasing the ram velocity will not result in an increased bend of the profile, see Fig. 4a,d. Only the heating of the L-profile in the feed channel is reduced by increasing the ram velocity.

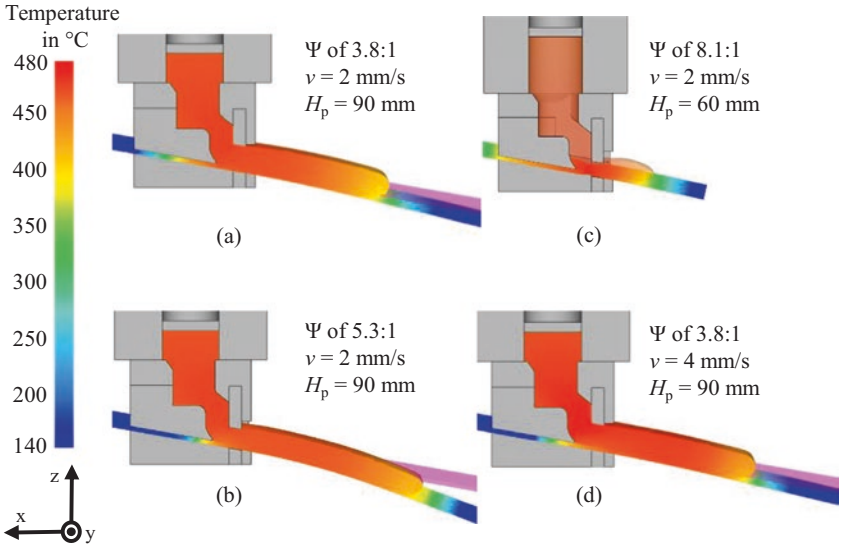


Fig. 4. Profile geometry and temperature distribution for varying extrusion ratios (a–c) and extrusion velocity (d) for an L-profile with a thickness of 3 mm

One possibility to increase the flow stress would be to reduce the contact temperature by decreasing the process temperatures of the aluminium and the heated tools. However, a reduction of the temperature is not recommended for two reasons. Firstly, the resulting extrusion force would increase and thus, possibly exceed the maximum force of the extrusion press. On the other hand, the formation of a material compound between aluminium and steel is impeded at lower temperatures. However, this is of crucial significance for the strength of the compound profile. The calculated extrusion force-time curves are plotted in Fig 5a. Dashed lines represent the force curves with an L-profile thickness of 5 mm, all the other curves represent results with a profile thickness of 3 mm. All curves show a similar characteristic course and differ only in the time shift due to varying extrusion velocity and the maximum extrusion force resulting from the variation of the extrusion ratios. Due to the design of the tooling system, different steps are visible in the extrusion force curve. First, the aluminium billet is upset to the container diameter, whereby the force increases only slightly. The subsequent first increase in force to 400 N marks the diameter reduction in the upper die. The aluminium is now pressed further into the tool, whereby the force remains constant until the aluminium reaches the wedge's redirection. As the wedge is deflected towards the die, the force increases slightly to approx. 800 N. The last steep rise marks the contact phase between the aluminium and the L-profile up to the stationary force plateau, which describes the extrusion. In the stationary phase, the force requirement decreases continuously due to the decreasing friction surfaces of the aluminium billet in the container. An increase in the extrusion ratio leads to an increase of the required extrusion force, as shown in the comparison of the courses with a ram velocity of 2 mm/s in Fig. 5a.

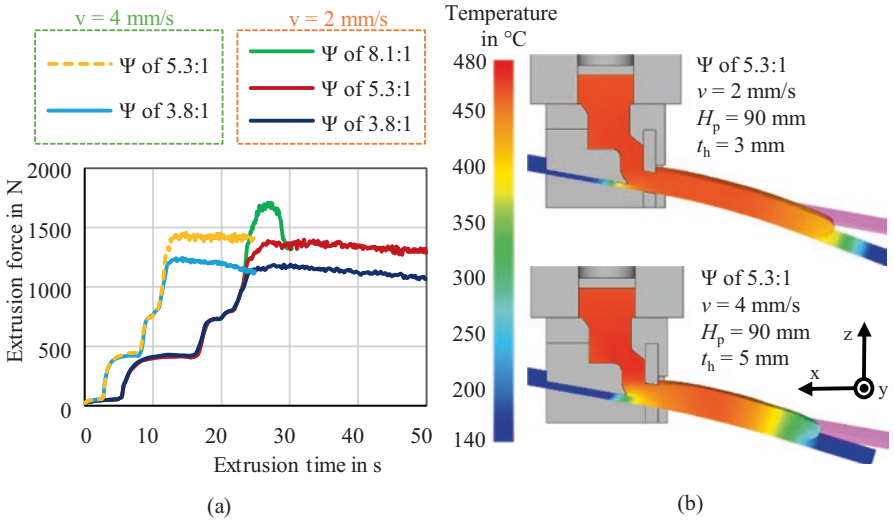


Fig. 5. Force-time curves for the variation of extrusion ratio, ram velocity and profile thickness (a) profile geometry and temperature distribution for a profile thickness of 5 mm and 3 mm (b)

The drop in extrusion force after an extrusion time of 28 s at an extrusion ratio of 8.1:1, is caused by the necking of the L-profile due to the large extrusion ratio. However, an increase in ram velocity only leads to a slight increase in the maximum extrusion force, comparing the two curves for an extrusion ratio of 3.8:1. To increase the profile straightness, the profile thickness was increased to 5 mm and the mean extrusion ratio of 5.3:1 was simulated with a press velocity of 4 mm/s. The thicker L-profile in combination with the increased ram velocity is intended to compensate for the profile bending, as the profile heats up more slowly and the bending stiffness is increased. When using the thicker L-profile, the die geometry has to be adjusted to achieve the same extrusion ratio as with the 3 mm profile. The resulting extrusion force for a profile thickness of 5 mm is slightly higher, but still on a comparable level to the force requirement for a 3 mm profile. The slightly higher force level is due to the higher ram velocity. Figure 5b shows the extruded profile for an L-profile thickness of 5 mm and 3 mm. By increasing the profile thickness, the bending could be significantly reduced. However, a slight curvature can be seen, so that if the extrusion ratio is increased to 8.1:1, no necking is to be expected, but rather a strong curvature of the profile. The contact temperature dropped to 420 °C due to the thicker L-profile, since the larger volume results in the L-profile extracting more heat from the aluminium. Simulated cross-section of the hybrid profile with an extrusion ratio of 5.3:1 with a profile thickness of 5 mm is shown in Fig. 6a. The section was taken 30 mm after exiting the tool at the final press height H_p of 100 mm. The extruded profile deviates only slightly from the target geometry, which can be explained by the expansion of

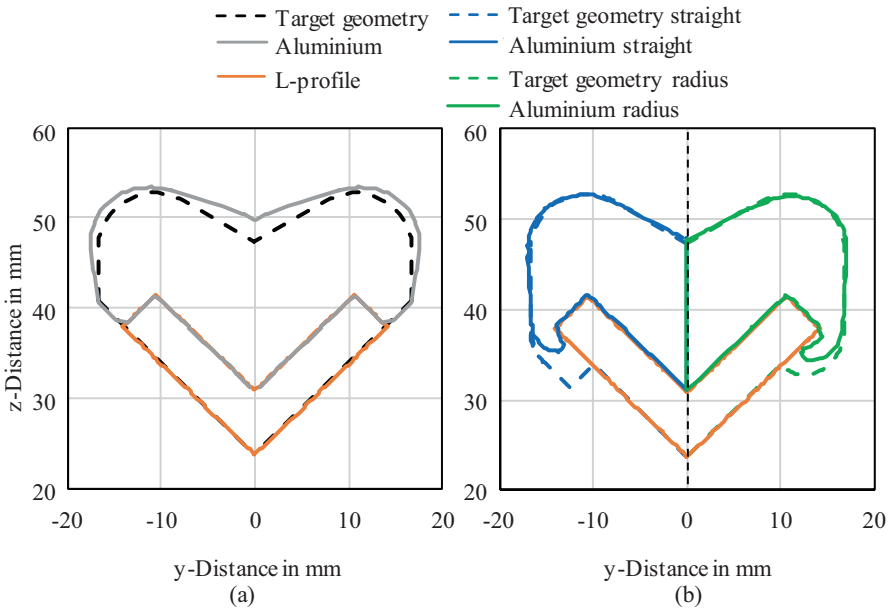


Fig. 6. Simulated profile cross-section ψ of 5.3:1, $v = 4$ mm/s, $t_h = 5$ mm (a), profile cross-section for variation of die geometry (b)

the aluminium after exiting the die. The aluminium completely fills the cavities of the L-profile, so that no hollow spaces are created in the profile.

The clamps are not completely filled in both variants. The reason for the incomplete filling might be that the lower die, which guides the L-profile, is aligned flush with the die. This allows the aluminium to flow directly into the die without filling the form completely. The incomplete filling reduces the pressure in the welding chamber and thus, the expansion of the aluminium after leaving the die is decreased. In order to fill the clamp, the contour of the lower die should not be flush with the die. By offsetting the lower die contour, a pronounced welding chamber is formed, so that enough pressure is generated for complete forming.

3 Conclusions and Outlook

In the scope of this work, the numerical design of a lateral angular co-extrusion (LACE) process for the production of asymmetrical aluminium-steel compound profiles was presented. An FE model was built up and used to examine the influence of different tool geometries and process parameters on the developed LACE process. It was established that the extrusion ratio has a strong influence on both the maximum extrusion force and the profile straightness. Furthermore, it was shown that with an L-profile thickness of 5 mm the straightness is increased.

Based on these results, an extrusion tool system will be manufactured for experimental investigations on a 2.5 MN extrusion press on a laboratory scale. The numerical model will be validated on the basis of the experimental results. For better filling of the clamp geometries, a larger welding chamber will be added to improve the composite quality.

Acknowledgements. The results presented in this paper were obtained within the Collaborative Research Centre 1153/2 “Process chain to produce hybrid high performance components by Tailored Forming” in the subproject A1, funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation)—252662854. The authors thank the German Research Foundation (DFG) for their financial support of this project.

References

1. Kleiner, M., Schomäcker, M., Schikorra, M., Klaus, A.: Manufacture of extruded and continuously reinforced aluminum profiles for ultra-lightweight constructions. Mater. Werkst. **35**, 431–439 (2004)
2. Foydl, A., Pfeiffer, I., Kammler, M., Pietzka, D., Matthias, T., Jäger, J., Tekkaya, A.E., Behrens, B.-A.: Manufacturing of steel-reinforced aluminum products by combining hot extrusion and closed-die forging. Key Eng. Mater. **504–506**, 481–486 (2012)
3. Chatti, S., Pietzka, D., Ben Khalifa, N., Jäger, A., Selvaggio, A., Tekkaya, A.E.: Lightweight construction by means of profiles. Key Eng. Mater. **504–506**, 369–374 (2012)
4. Pietzka, D., Ben Khalifa, N., Gerke. S., Tekkaya, A.E.: Composite extrusion of thin aluminum profiles with high reinforcing volume. Key Eng. Mater. **554**, 801–808 (2013)

5. Grittner, N., Striewe, B., Hehl, A.V., Bormann, D., Hunkerl, M., Zoch, H.W., Bach, F.W.: Co-extrusion of aluminium-titanium-compounds. Key Eng. Mater. **491**, 67–74 (2012)
6. Behrens, B.-A., Klose, C., Chugreev, A., Thürer, S.E., Uhe, J.: Numerical investigations on the lateral angular co-extrusion of aluminium and steel. AIP Conf. Proc. **1960**, 030001 (2018)
7. Behrens, B.-A., Uhe, J., Thürer, S.E., Klose, C., Heimes, N.: Development of a modified tool system for lateral angular co-extrusion to improve the quality of hybrid profiles. Procedia Manuf. **47**, 224–330 (2020)
8. Lu, X., Zhang, C., Zhao, G., Guan, Y., Chen, L., Gao, A.: State-of-the-art of extrusion welding and proposal of a method to evaluate quantitatively welding quality during three-dimensional extrusion process. Mater. Des. **89**, 737–748 (2016)
9. Priel, E., Ungarish, Z., Navi, N.U.: Co-extrusion of a Mg/Al composite billet: a computational study validated by experiments. J. Mater. Process. Technol. **236**, 103–113 (2016)
10. Behrens, B.-A., Klose, C., Chugreev, A., Heimes, N., Thürer, S.E., Uhe, J.: A numerical study on co-extrusion to produce coaxial aluminum-steel compounds with longitudinal weld seams. Metals **8**, 717 (2018)
11. Duan, X., Sheppard, T.: Simulation and control of microstructure evolution during hot extrusion of hard aluminium alloys. Mater. Sci. Eng. A **351**, 282–292 (2003)

Investigation of the Phase Transformation in Hot Stamping Processes with Regard to the Testing Facility

A. Horn[1(✉)], T. Hart-Rawung[2], J. Buhl[2], M. Bambach[2], and M. Merklein[1]

[1] Institute of Manufacturing Technology, Friedrich-Alexander-Universität Erlangen-Nürnberg, Egerlandstraße 11-13, 91058 Erlangen, Germany
alexander.horn@fau.de

[2] Chair of Mechanical Design and Manufacturing, Brandenburg University of Technology Cottbus - Senftenberg, Konrad - Wachsmann - Allee 17, 03046 Cottbus, Germany

Abstract. Hot stamping of boron manganese steel is a state of the art process for manufacturing safety-relevant automotive components with regard to lightweight design. In terms of the mechanical properties, the microstructural evolution is of particular interest. Besides conventional hot stamping processes, this is especially apparent for the manufacture of hot stamped parts with tailored properties. To reduce costs and scrap parts, numerical process design requires appropriate material models considering transformation kinetics during in-die quenching. The phase transformation behavior can be characterized through experiments on various testing facilities, which can lead to discrepancies in the results and therefore in the predicted mechanical properties. In this paper, the results from two testing facilities are compared to each other through dilatometry experiments on the boron-manganese steel 22MnB5. For further characterization of the differences, secondary samples are taken to analyze the hardness as well as the microstructure by optical microscopy.

Keywords: Hot stamping · Dilatometry test · Phase transformation

1 Introduction

Hot stamping of ultra-high strength steels is a key technology in terms of lightweight constructions in the automotive industry. Especially safety-relevant components, such as B-pillars are manufactured with this process [1]. Its importance can be traced back to the excellent strength-to-weight ratio of the processed boron-manganese steels, such as 22MnB5. Due to various controlled phase transformations, an ultimate tensile strength of at least 1500 MPa is achievable. In the first process step, the sheets are heat-treated above AC_3 temperature with the aim of a full austenitization of the initial ferritic-pearlitic microstructure. Subsequently to austenitization, the sheets are

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 76–85, 2021.
https://doi.org/10.1007/978-3-662-62138-7_8

transferred to a forming-press, where forming and quenching is done simultaneously. As long as a critical cooling rate is ensured, the martensitic transformation takes place. The control of phase transformation is also highly relevant in process variations such as partial hot stamping. Unlike conventional hot stamping, the martensitic transformation is locally prevented to manufacture components with tailored properties [2]. By maintaining structural integrity and increasing ductility in some areas of the parts, passenger safety can be improved. Regarding process control, heat transfer between workpiece and tool needs to be reduced. By applying grooved tools [3] or tool material with low thermal conductivity [4], the quenching rate is reduced to receive a microstructure composed of ferrite, pearlite and bainite. Another approach is the utilization of heated tool segments [5] to decrease the temperature gradient between tool and workpiece. In this case, the sheets are not quenched to room temperature but held isothermally until the desired phase composition is achieved [6].

To be able to predict mechanical properties accurately as well as to ensure process stability extensive knowledge on phase transformation is necessary. Without previous deformation of the austenite, the critical quenching rate for martensitic transformation is typically specified between 20 K/s [7] and 27 K/s [8]. Additional pre-strain results in an increase in the critical quenching rate [9] and it shifts the transformation rate. Hence, the deformation of the austenitic phase decreases thermal expansion and reduces the density difference between austenite and martensite [10]. In this context, further investigations on the influence of non-isothermal deformation of the austenitic phase were done inter alia by Naderi et al. [11] and Nikravesh et al. [12].

Since the investigation of phase transformation is highly time-consuming research effort is directed towards developing a virtual dilatometer [13]. By referring to a trained model, which may be also a neuronal network [14], based on real dilatometry data, the testing effort will be significantly reduced. For this purpose, a database with appropriate grid points is required. However, to enable the combination of virtual and additional experimental data, the transferability of the results has to be ensured. In this context, processing of measured data from various testing facilities has to be taken into account. Thus, this study aims to investigate the phase transformation of hot stamping steel under various quenching and load conditions with regard to two different testing rigs.

2 Material and Methodology

Within this study, the transformation behavior is analyzed using two different testing facilities. One of them is a deformation dilatometer DIL805 A/D/T by TA Instruments and the other one is a thermomechanical simulator Gleeble 3500-GTC from Dynamic Systems Inc. A more detailed description of both testing apparatuses is given in the two subsequent subchapters. As sheet material, the conventional boron-manganese hot stamping steel 22MnB5 is used. The sheets have a thickness of 1.5 mm and an Al-Si coating in the case of the specimens for Gleeble. The same thermal treatment is applied in both facilities. The specimens are first heated to the austenitization

temperature of 950 °C with a heating rate of 5 K/s and held for 300 s. Afterwards, the samples are cooled to 800 °C with 30 K/s, where potential pre-strain is applied. This heating cycle is valid for samples being continuously quenched, as well as for those being isothermally held. Regarding the process route of continuous quenching to room temperature, the experiments are conducted with and without additional deformation. Subsequently to pre-straining, five different quenching rates are applied. In the case of isothermal holding, pre-strain is not considered. The samples are cooled to 550 °C with 50 K/s and held for 60 s. An overview of the thermal treatment is given in Fig. 1. The tests were repeated two times to improve the statistical quality. For the comparison of the transformation kinetics, the transformation rate is calculated. The objective is to make the measurement data independent from the testing apparatus. For this purpose, the lever rule is used to determine the transformed fraction from dilatometric data in the first step. Based on these results, the transformation rate is computed by derivation. Furthermore, the respective phase fractions are determined from the transformation rate. After the dilatometric experiments, secondary samples are taken from the tested specimens to analyze the resulting hardness on a Fischerscope HM 2000.

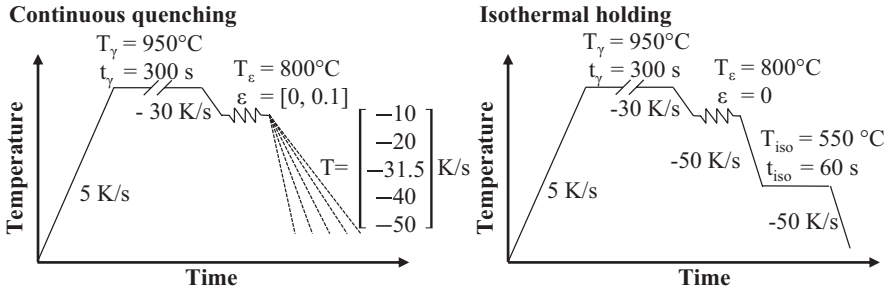


Fig. 1. Thermal treatment in the physical simulator and the deformation dilatometer

2.1 Deformation Dilatometer

The hot deformation dilatometer is an instrument, which finds many applications in material characterization, for example, the determination of flow stresses, the characterization of phase transformation kinetics, and the approximation of the coefficient of thermal expansion at high temperatures. In this work, the experiments are performed with a quenching/deformation dilatometer DIL805A/D as depicted in Fig. 2. Inside the dilatometer, there is an induction coil used for heating the specimens to the desired temperature. The instrument is also equipped with a cooling system using gaseous media to quench the samples with a defined cooling rate. To measure the temperature during experiments, thermocouples of type S were spot welded on the center of the gauge length. The strain is measured tactile at the edge of the gauge length l. The dimensions of the specimens used in this investigation are shown in Fig. 2.

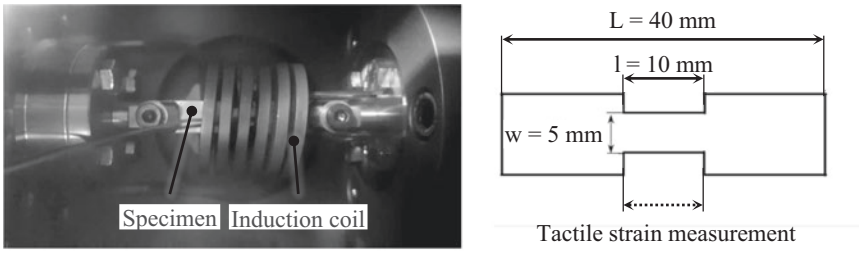


Fig. 2. Testing setup of the Dilatometer with respective sample geometry

2.2 Thermomechanical Simulator

The thermomechanical simulator Gleeble 3500-GTC is equipped with an optical extensometer Keyence LS-7601. Contrary to the deformation dilatometer, the Gleeble uses electric resistance for heating. Since the samples are only heated in their center, a temperature gradient develops in width and length direction. To minimize temperature loss on the fringe, especially in the measuring area, the sample is modified compared to conventional tensile specimens. The adjustment is oriented towards the geometry for dilatometer experiments recommended by Dynamic Systems Inc. The temperature and phase transformation dependent strain is measured crosswise over the central bridge, as depicted with the dotted line in Fig. 3. For the temperature control, spot-welded thermocouples of type K are used, while quenching of the sheets is realized by utilizing a compressed air nozzle system.

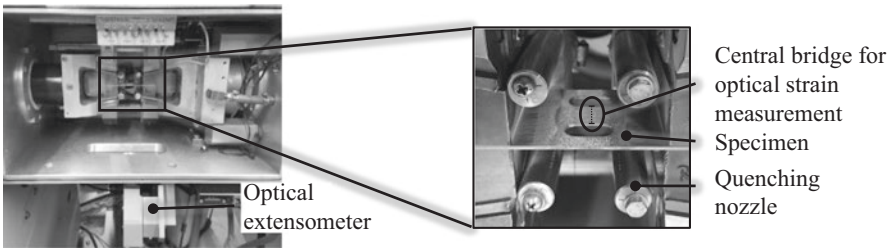


Fig. 3. Experimental setup on the thermomechanical simulator

3 Results and Discussion

Figure 4 shows exemplarily the results from dilatometric testing in the thermomechanical simulator Gleeble 3500-GTC and the deformation dilatometer DIL805A/D. The samples were quenched with 40 K/s, whereas no pre-strain was applied. On the left side, the respective true strain development is depicted. It can be seen, that there are distinct differences between both testing facilities. To make the results

independent of the testing rig and improve comparability, the transformation rate is calculated. After applying the lever rule, it is computed by deriving the transformed fraction after temperature.

Three different findings can be deduced from the evaluation of the various transformation rates. As shown on the right side in Fig. 4, the onset of phase transformation as well as the maximum of the transformation rate is different for both testing facilities. The results indicate that the whole calculated phase transformation is shifted to lower temperatures in the Gleeble. In the case of the presented results in Fig. 4, the martensite start temperature M_s varies by around 40 °C, with $M_s = 404$ °C in case of the dilatometer and $M_s = 362$ °C in the case of the Gleeble. By using a dilatometer Baehr DIL 805 as well, Nikravesh et al. [15] identify a martensite start temperature of 401 °C at a quenching rate of 40 K/s. This corresponds well with the specifications of $M_s = 400$ °C specified by various suppliers as seen in [8] or [16]. On the contrary Chang et al. [17] identify a martensite start temperature of 380 °C within their experiments on a Gleeble 1500. However, these studies relate to various austenitization parameters, which also influence the transformation behavior as stated by Naderi et al. [11]. The effect of shifted start temperatures is especially apparent for martensitic transformation, but it can is also be seen for other phases and occurs independently of pre-strain and quenching rate. The reason for the difference between both tests can originate from the geometry of specimens, and the method employed to measure the strain from both experiments. Thus, further analysis should be performed to determine the root cause of this deviation.

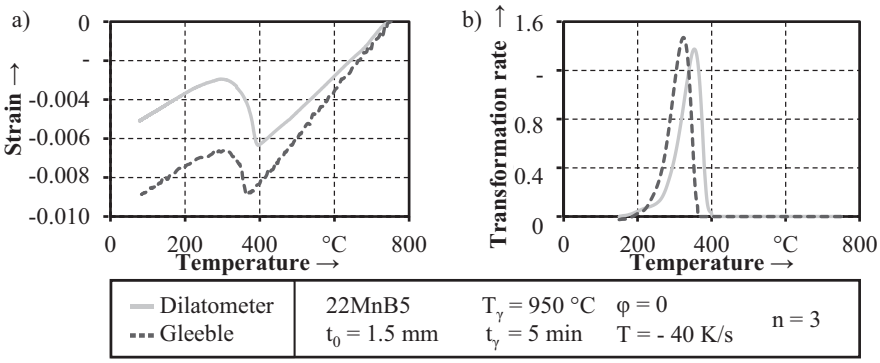


Fig. 4. a) Strain curve and b) calculated transformation rate at a quenching rate of 40 K/s

Besides a shift in start temperatures, also differences regarding the development of the transformation rate are detectable, as exemplarily depicted on the left side of Fig. 5. The samples were continuously quenched to room temperature with 10 K/s. As already seen for a cooling rate of 40 K/s, the onset of phase transformation is at a higher temperature in the case of the dilatometer. Moreover, it can be seen, that the transformation rate rapidly increases in the dilatometer, while the growth is less

pronounced in the Gleeble. Regarding 40 K/s, the temperature difference between the beginning of the phase transformation and the maximum of the transformation rate remains constant, whereas it is increased for 10 K/s. This indicates not only a shift in the transformation kinetics but also a variation concerning the phase composition. Further discrepancy can be seen in the martensitic region, where the Gleeble specimens exhibit an increased transformation rate.

Contrary to the findings regarding continuous quenching to room temperature, both testing facilities show good agreement regarding the transformation kinetics in the case of isothermal holding, as presented on the right side in Fig. 5. The undeformed samples were quenched to $T_{iso} = 550$ °C and held for 60 s. To demonstrate the repeatability of the experiments, all three test runs are depicted. It can be seen, that the development of the transformation rate is nearly identical for both testing setups. In the case of the Dilatometer, the results also indicate a good repeat accuracy, while the Gleeble tests exhibit a temporal variation of around five seconds.

The evaluation of the transformation rates pointed out, that the transformation kinetics are very similar in the case of isothermal holding, while differences can be seen when continuous quenching is applied. This relates to a shift of the start of the transformation temperatures as well as to the transformation kinetics, which indicates a varying phase composition. Thus, to enable further analysis of the differences, the final phase fractions are calculated from the transformation rates. As mentioned above, the transformation rate is calculated from the change of the phase fraction respect to temperature. Therefore, the phase fraction for each phase can be obtained by integrating the transformation rate within the temperature regions where ferrite, bainite, and martensite are formed.

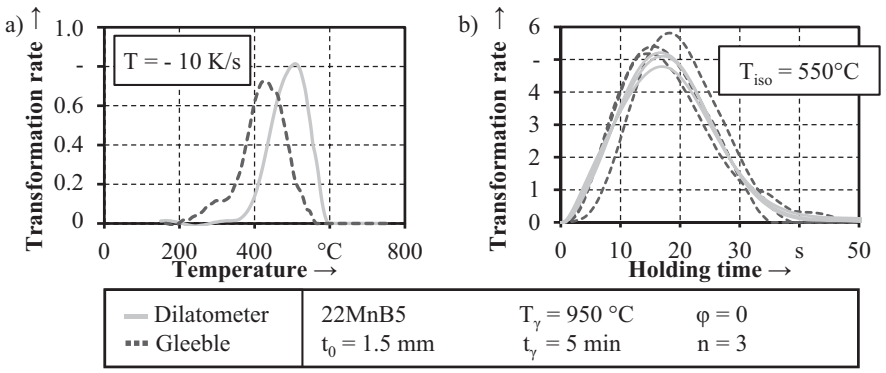


Fig. 5. Transformation rate of selected parameter combinations during a) continuous quenching and b) isothermal holding

The respective phase fractions are presented in Fig. 6. Generally, it can be seen, that as expected, the amount of martensite decreases at lower cooling rates, mainly in favor of bainite. Small fractions of ferrite occur in the case of the lowest quenching rate investigated within this study. Regarding the influence of pre-strain, the amount

of martensite decreases at higher degrees of deformation, as already seen in the work of Nikravesh et al. [12]. Comparing both testing facilities, the results show good agreement for quenching rates between 31.5 and 50 K/s.

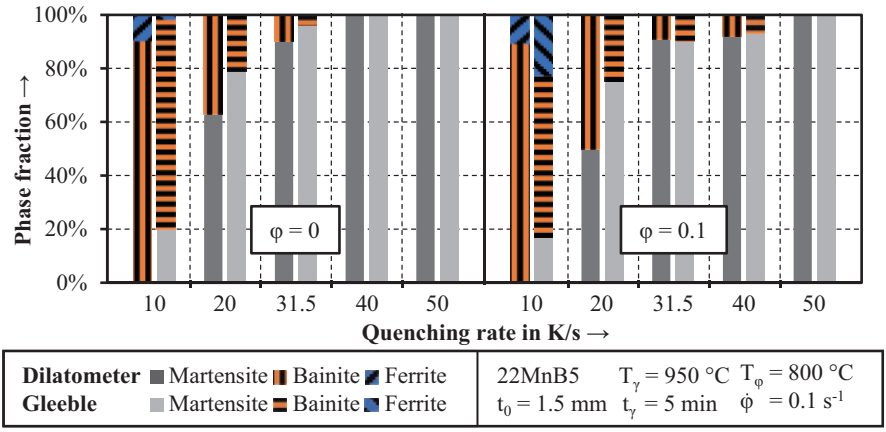


Fig. 6. Calculated phase fraction from the transformation rate

In the undeformed condition, the results from the Gleeble and the dilatometer indicate a fully martensitic structure for 40 K/s and 50 K/s. At 31.5 K/s, the bainite fraction is higher in the case of the dilatometer but lies within 5%. However, it can be seen, that in the case of applied pre-strain of 0.1, the phase fractions are identical within 1%. At a quenching rate of 20 K/s, the differences in the predicted phase composition increase, while the influence of pre-strain is visible for both facilities. By contrast, in the case of 10 K/s, no influence of deformation in the austenitic region is observable from the results of the dilatometer. Furthermore, only bainitic and ferritic fractions are identifiably from the transformation rate, whereas the measurement data of the Gleeble also shows a martensitic transformation. Regarding the thermomechanical simulator, pre-strain results in a decrease of bainite in favor of ferrite, whereas the amount of martensitic fraction remains approximately constant.

Since these values of phase fractions are predictions from the transformation curves, further investigations are necessary to verify them. Thus, secondary samples were taken to analyze the hardness in the cross-section of the quenched specimens. Regarding the hardness after heat treatment in one of both testing facilities, no significant difference can be seen between the Gleeble and the dilatometer as shown in Fig. 7. However, there is a trend of slightly higher values for specimens tested in the thermomechanical simulator. Concerning the predicted phase fractions, the hardness measurements correspond well to the previous findings. At quenching rates of 31.5 K/s, 40 K/s and 50 K/s, the hardness remains nearly constant with a maximum of 10 HV0.2 difference between undeformed and pre-strained samples.

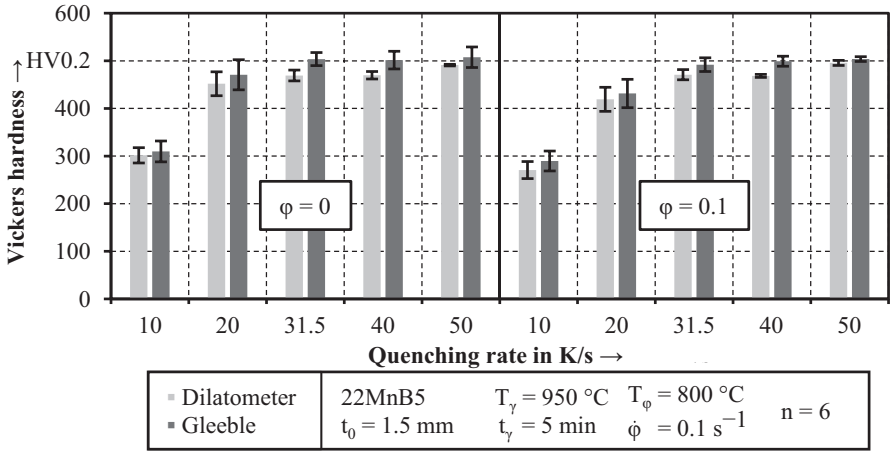


Fig. 7. The hardness of quenched samples

For 20 K/s, a deformation of 0.1 at a temperature of 800 °C results in a drop of 30 HV0.2 and 40 HV0.2 respectively, which can be traced back to the decreased amount of martensite in favor of bainite. While the evaluation of the measurement data from the dilatometer indicated, that there is no influence of pre-strain for 10 K/s, the hardness values in Fig. 7 show a decrease of 30 HV0.2. However, due to the standard deviation, the effect is not significant. Regarding the difference between 20 K/s and 10 K/s the hardness measurements confirm the predicted decrease of martensite in favor of bainite and ferrite. This was also seen in the investigations of Nikravesh et al. [12], were a significant increase in hardness was observed while rising the quenching rate from 10 K/s to 20 K/s. Furthermore, while the calculated phase fractions from transformation rates indicated a deviation of both facilities in case of low quenching rates, the hardness values were within a range of 20 HV0.2. Especially under consideration of the standard deviation, it is not possible to verify the differences in the predicted phase fractions from hardness measurements.

4 Summary and Outlook

Within this study, the phase transformation of hot-stamping steel 22MnB5 was analyzed under different load conditions and with regard to the transferability of the data from different testing facilities to set up a database for a virtual dilatometer. The evaluation of the results revealed, that the transformation kinetics differ by using either the dilatometer DIL805 A/D/T and the thermophysical simulator Gleeble 3500-GTC, although the same thermal cycle was applied. This refers to the beginning of the phase transformation and the predicted phase fraction. Contrary to that, the evaluated transformation kinetics of both testing rigs showed a good agreement regarding isothermal holding. Moreover, the mechanical properties of secondary samples were in

accordance with the predicted phase fractions and exhibited only minor differences when comparing the dilatometer and the thermophysical simulator. In the case of low quenching rates of 10 K/s and 20 K/s it was not possible to verify the predicted differences through hardness measurements. However, since isothermal holding is of particular interest in terms of modeling of the transformation kinetics, fundamental transferability of the results can be ensured for the development of a virtual dilatometer, especially with regard to the processing of simulated and experimental data.

Though, further research is necessary concerning the handling of measurement data from different testing facilities in terms of continuous quenching. In this context, prospective investigations have to focus inter alia on the thermomechanical field of the whole specimens. Currently, the influence of a temperature gradient in both framing bridges of the Gleeble specimen, which may cause a superposition of tensile or even compressive stress, is unknown. Moreover, one of the next steps is the validation of the predicted microstructure by using quantitative metallography.

Acknowledgment. The authors would like to thank the Research Association for Steel Application (Forschungsvereinigung Stahlanwendung e. V. FOSTA) for funding the project P1305 “Development of a learning phase transformation and dilatometer model for the virtual process design of press hardening processes”. Further thanks go to all cooperating industrial partners for their support and cooperation in this project.

References

1. Mori, K., Bariani, P.F., Behrens, B.-A., Brosius, A., Bruschi, S., Maeno, T., Merklein, M., Yanagimoto, J.: Hot stamping of ultra-high strength steel parts. CIRP Ann. **66**, 755–777 (2017)
2. Merklein, M., Wieland, M., Lechner, M., Bruschi, S., Ghiotti, A.: Hot stamping of boron steel sheets with tailored properties: a review. J. Mater. Process. Technol. **228**, 11–24 (2016)
3. Mori, K., Okuda, Y.: Tailor die quenching in hot stamping for producing ultra-high strength steel formed parts having strength distribution. CIRP Ann. **59**, 291–294 (2010)
4. Casas, B., Latre, D., Rodríguez, N., Valls, I.: Tailor made tool materials for thepresent and upcoming tooling solutions in hot sheet metal forming. In: Steinhoff, K., Oldenburg, M., P., B. (eds.) 1st International Conference on Hot Sheet Metal Forming of High-Performance Steel - CHS^2, pp. 22–36 (2008)
5. Bardelcik, A., Vowles, C.J., Worswick, M.J.: A mechanical, microstructural, and damage study of various tailor hot stamped material conditions consisting of Martensite, Bainite, Ferrite, and Pearlite. Metall. Mater. Trans. A **49**, 1102–1120 (2018)
6. Svec, T., Grüner, M., Merklein, M.: FE-simulation of the heat transfer by defined cooling conditions during the hot stamping process. Key Eng. Mater. **473**, 699–706 (2011)
7. Chen, X., Xiao, N., Li, D., Li, G., Sun, G.: The finite element analysis of austenite decomposition during continuous cooling in 22MnB5 steel. Modell. Simul. Mater. Sci. Eng. **22**, 065005 (2014)
8. Aranda, L.G., Chastel, Y., Pascual, J.F. and Negro, T.D.: Experiments and simulation of hot stamping of quenchable steels. In: 7th International Conference on Technology of Plasticity, pp. 1135–1140 (2002)

9. Barcellona, A., Palmeri, D.: Effect of plastic hot deformation on the hardness and continuous cooling transformations of 22MnB5 microalloyed boron steel. Metall. Trans. A **40**, 1160–1174 (2009)
10. Hart-Rawung, T., Buhl, J., Bambach, M.: Extension of a phase transformation model for partial hardening in hot stamping. J. Mach. Eng. **18**, 88–98 (2018)
11. Naderi, M., Saeed-Akbari, A., Bleck, W.: The effects of non-isothermal deformation on martensitic transformation in 22MnB5 steel. Mater. Sci. Eng. A **487**, 445–455 (2008)
12. Nikravesh, M., Naderi, M., Akbari, G.H., Bleck, W.: Phase transformations in a simulated hot stamping process of the boron bearing steel. Mater. Des. **84**, 18–24 (2015)
13. Bambach, M., Buhl, J., Hart-Rawung, T., Lechner, M., Merklein, M.: Towards virtual deformation dilatometry for the design of hot stamping process. Procedia Eng. **207**, 1821–1826 (2017)
14. Hart-Rawung, T., Buhl, J., Bambach, M.: A fast approach for optimization of hot stamping based on machine learning of phase transformation kinetics. Procedia Manuf. **47**, 707–712 (2020)
15. Nikravesh, M., Naderi, M., Akbari, G.H.: Influence of hot plastic deformation and cooling rate on martensite and bainite start temperatures in 22MnB5 steel. Mater. Sci. Eng. A **540**, 24–29 (2012)
16. thyssenkrupp: Product information for manganese-boron steels for hot forming (2019). https://www.thyssenkrupp-steel.com/en/products/sheet-coated-products/manganese-boron-steel-for-hot-forming/mbw-manganese-boron-steel-for-hot-forming/. Accessed 05 Mar 2020
17. Chang, Y., Li, X.D., Zhao, K.M., Wang, C.Y., Zheng, G.J., Hu, P., Dong, H.: Influence of stress on martensitic transformation and mechanical properties of hot stamped AHSS parts. Mater. Sci. Eng. A **629**, 1–7 (2015)

Simulation of an Electromagnetic Foil-Feeding Device

O. Commichau(✉), A. Höber, B.-A. Behrens, and R. Krimm

Institute of Forming Technology and Machines, Leibniz Universität Hannover,
An der Universität 2, 30823 Garbsen, Germany
commichau@ifum.uni-hannover.de

Abstract. Very thin foil material is usually fed by roll-feeding devices. However, these devices can limit the feeding speed and cause damage to the foil surface finish. A promising alternative are electromagnetic systems. This technology is based on the working principle of a linear induction machine.

In this paper, a study is presented which examines the potential and limitations of this approach for foils. For this purpose, a parametric finite element analysis (FEA) model was developed using different ANSYS element types. With this model, the physical limits of the system, e.g. heating by the excitation of the windings and induction of eddy currents in the foil, the model is also used to optimise the ratio between feeding force and material heating.

The feeding technology presented here provides for a more efficient production and thus significantly contributes to shaping the change in the production industry.

Keywords: Contactless feeding · Electromagnetic simulation · Finite element analysis

1 Introduction

Many branches in electronic and automotive production technology use very thin metallic foils as an engineering material for parts of their products. The thickness of these foils is less than 500 µm and can possibly be as thin as 15 µm. In modern production processes, the material is moved in highly automated processes. Hence, the material needs highly dynamic handling to be transported within the manufacturing process. Forming technologies such as deep drawing and high-speed cutting are highly repetitive processes and therefore very economic. Modern forming presses can reach stroke rates between 200 and 1200 per minute [1]. Accordingly, the foil material needs to be moved fast. For example, the aforementioned stroke rates would require feeding accelerations between 0.89 and 32 m/s^2. Between every stroke of the press, the foil needs to be moved to its next position, quickly and highly accurate. This movement is typically achieved by mechanical feeding systems [2]. These types of feeding devices work accurately, however they are limited in terms of dynamics and quality assurance. The inertia of their

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 86–95, 2021.
https://doi.org/10.1007/978-3-662-62138-7_9

components impairs the ability to keep up with the stroke rate of modern presses. Also, the mechanical force applied to the foil can cause damage to the surface of the foil [3]. A promising alternative to mechanical feeding systems are systems based on electromagnetic forces. This method was initially developed for the purpose of feeding sheet-metal, however it can also be applied to the feeding of thin metallic foils. The electromagnetic feeding of foil material bears several challenges though: The fragile shape requires a homogeneous feeding force acting on the surface, otherwise creases can occur. The flat shape and low mass result in a low heat capacity, so that the thin material overheats quickly and must not be exposed to high-temperature environments. To sum up, transferring the concept of electromagnetic feeding from sheet metal to metallic foils bears high requirements for the technology and must be investigated accordingly.

2 State of the Art

In many cases, roller and clamp feeders are the chosen solution for material feed. These devices are based on friction: By applying the rollers or clamps to the feed material with contact pressure, frictional contact allows the material to be guided with a movement of the actuators. The contact force between rollers/clamps and the component must be higher than the feeding force to prevent slippage. The contact is usually applied to the component by means of pneumatic or hydraulic actuators while the feeding force is applied by servo electric motors coupled via gearing to the rollers. The application of contact pressure may cause damage to the foils' surface finish [3], which is undesired, especially for refined surfaces. Further, the absolute positioning accuracy of these systems is usually limited from ± 0.05 mm (rollers) to ± 0.03 mm (clamps) [4]. The components of these feed systems have to accelerate the weight of the rolls/clamps for every stroke, which limits the dynamics due to their inertia.

A promising alternative to the feeding technologies mentioned above has been investigated in prior research projects [5, 6]. This concept is based on the operating principle of a linear induction machine in which eddy currents are induced in the conductive material (Fig. 1). Two stators are used in a double-stator arrangement underneath and on top of the sheet metal. The stators contain electric conductors that are arranged in a three-phase winding to generate a moving magnetic field in the air gap. The magnetic field induces eddy currents in the sheet metal.

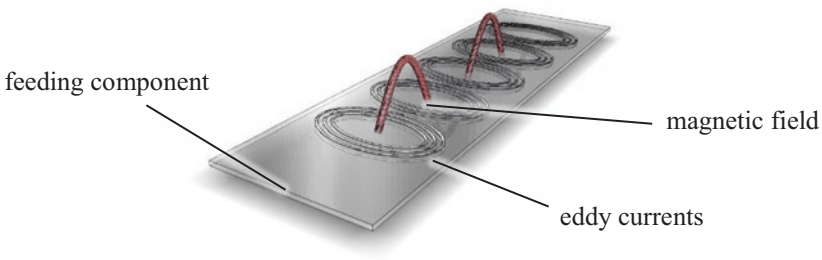


Fig. 1. Occurrence of eddy currents due to magnetic induction [3].

This leads to a mutual reaction applying Lorentz force to the sheet metal. Due to this application of Lorentz force, the material accelerates to the required speed. A major advantage of this approach is that no drive elements such as rollers/clamps, gear elements, rotors etc. need to be accelerated, thus ensuring high dynamics.

However, these investigations are limited to sheet metal [5] and profile bars [6] rather than foil material with a thickness of 500 µm and below. The fineness of foils leads to several challenges. Foils have limited heat capacity due to their layer-like appearance. As described in [5], this feeding method leads to high eddy currents of over 40 A/mm^2 in the sheet metal. These currents could lead to overheating and damage of the foil within milliseconds. With the simulation model used in [5], the resulting eddy currents rise up to 187 A/mm^2 while using the same input current. The reason for that is the reduced distance between the upper and the lower stator in comparison to the sheet-metal feeding case. The total distance between the stators is *air gap* + *metal thickness*. While the air gap remains constant, the thickness of the metal is reduced. With lower volume between the stators and constant current, the energy density increases. This load would lead to a current-induced temperature enhancement of 399 K after one second. As mentioned above, another challenge is the low stiffness of the foil. While sheet metal provides sufficient stiffness to be handled without the appearance of creases, this is not the case with foils. The force impact must be homogenous along the foil rather than having a sectional force impact to avoid creasing.

None of the existing publications deal with these challenges. In [5], sheet-metal with a thickness between 1 mm and 3 mm was investigated. In [6], the feeding method was applied to circular rods with a wall thickness of 5 mm by means of field forming elements. An FEA model designed in ANSYS was conducted for a simulation study. Investigations were carried out with aluminium, copper and steel as sheet-metal material. Heat generation was neglected in these publications, which is inappropriate when investigating thin metallic foils.

The state of the art depicts that the application of electromagnetic feeding to metallic foils is both, promising and challenging. In this paper, simulation studies are presented to gain better knowledge regarding the handling of thin foils. This paper aims at closing the above-mentioned research gap by investigating the applicability of this approach to metallic foils. The main objective to apply the principle of electromagnetic feeding to metallic foils has mandatory requirements. Overheating of the foil must be avoided as well as the appearance of creases. This paper aims to lay the foundation to verify the adherence of these requirements by executing simulation studies regarding the heat and force development in the foil.

3 Simulation

The simulation approach to the existing feeding devices mentioned above, provides the basis for the studies carried out here. However, the analysis has been carried out with ANSYS element type SOLID117 which has meanwhile become outdated and has been replaced with SOLID236 [7]. For this reason, it is necessary to evaluate the

difference between the two elements and the influence of their usage on the simulation results.

3.1 Comparison of Element Types

In order to model thin metal foil, the choice of an appropriate ANSYS element type is of major importance for the calculation results, since this defines the required degrees of freedom (DOF). As potential elements for modelling thin structures, ANSYS provides so-called SOLIDs and SHELLs [8]. In the application addressed here, the foil can be modelled either as a "surface" using a SHELL element or as a very thin three-dimensional body using a SOLID.

SHELL elements allow efficient simulation of mechanical stress and heating processes for thin structures, but are only of limited use for the consideration of electromagnetic forces. SOLID elements offer the exact definition of boundary conditions and the calculation of internal effects, such as electromagnetic forces or eddy currents for instance, but require much more computing resources and are subject to an effect known as "locking" [9, 10] which could lead to calculation errors under certain conditions. With the SOLSH190 (Solid-Shell) element type, ANSYS software provides a compromise solution for thin layered objects. However, since SHELL and SOLSH elements cannot be used for the calculation of electromagnetic effects at all, they are left out of further consideration.

Additionally, ANSYS internal programming language APDL allows user-defined USER300 elements with up to ten DOF including nodal temperature (TEMP), nodal magnetic vector potential (AZ) and nodal electric potential (VOLT) [11]. By using the provided API, the USER300 element's behaviour can be set to SOLID, SHELL or SOLSH.

The element types considered in this paper are SOLID117, SOLID236 and SOLID90. The properties of these elements are listed in Table 1.

Table 1. Properties of element types utilized.

Element Type	Nodes	DOF	Relevant Output
SOLID117	20	AZ, VOLT	electromagnetic force, magnetic flux density, joule heating
SOLID236	20	AZ,VOLT	electromagnetic force, magnetic flux density, joule heating
SOLID90	20	TEMP	nodal temperature

A feasible approach could be using SOLID117 or SOLID236 elements to calculate electromagnetic forces and eddy currents and using SOLID90 for calculating heating effects due to induction. The use of a combined model with electromagnetic and thermal solids does not lead to a disadvantage in calculation time and usability. Test applications have even indicated that the calculation time for user-defined elements increases. The additional effort in setting up a user-defined element is therefore not justified and discarded. The SOLID117 elements have been used in prior

research. In order to motivate the investigations carried out in this work, a comparative study will be held to reveal the differences between the SOLID117 and the more accurate SOLID236 elements. In the comparative study, a ferromagnetic plate with a height-edge-length-ratio of 0.01 is placed in a magnetic field, generated by an electric coil. In the study, the following parameters are kept constant: ambient temperature ($T_a = 22$ °C), excitation current of the cylindric coil (20 kA rms), frequency (f = 50 Hz), dimensions of the coil ($r_{inner} = 50$ mm, $r_{outer} = 60$ mm, $h_{coil} = 100$ mm), dimensions of the plate ($A_{plate} = 150$ mm x 150 mm, $h_{plate} = 1.25$ mm) and distance between plate and coil (d = 0.5 mm).

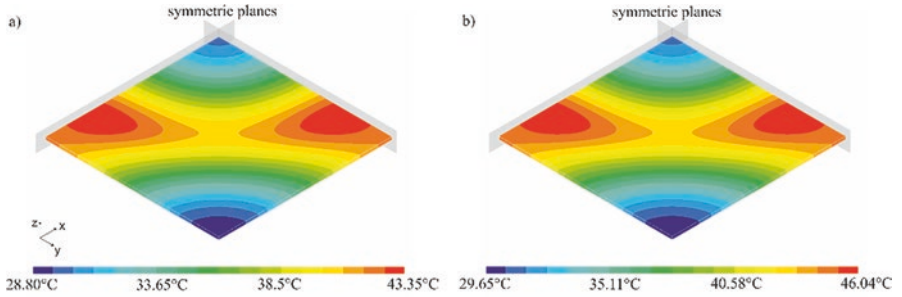


Fig. 2. Heating due to eddy currents in a ferromagnetic plate calculated for element types SOLID117 (a) and SOLID236 (b).

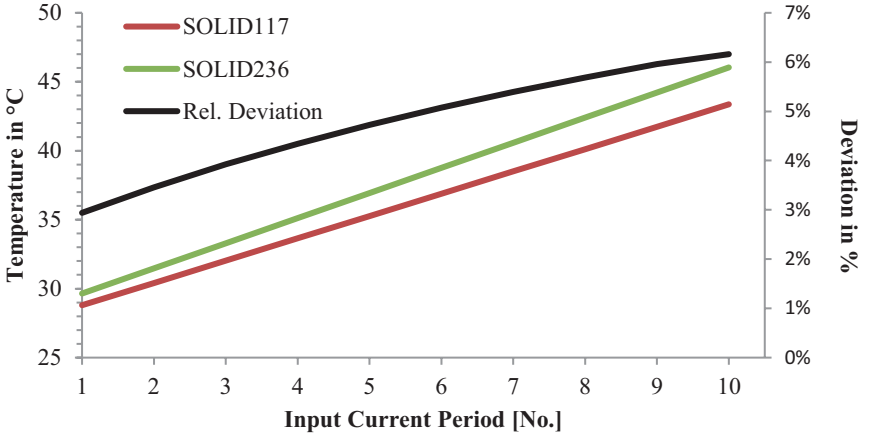


Fig. 3. Peak temperatures for the element types SOLID117 and SOLID236 in the plate and relative deviation of the simulation results.

To reduce simulation time, a quarter model was used. The results of the temperature distribution are illustrated in Fig. 2. It shows that the electromagnetic solid elements generate slightly divergent results: The post-processing results for the power-loss calculation in the plate are 18.39 W for SOLID117 and 20.71 W for

SOLID236, which is also indicated by the temperature values. The total energy loss was calculated by the total energy that is required for the heating of the foil. Further, Fig. 3 clearly shows that the deviations in the results of the different element types even increase with rising temperature. Thus, the selection of the optimal element has a significant influence on the accuracy of the results. Possible reasons for deviations are the ability to simulate a stranded coil, which comes with SOLID236, or the extended calculation of SOLID236 with a larger number of equations to enable more DOF than SOLID117. Moreover, experimental evaluations confirm that the results of SOLID236 are more accurate [7, 12]. On the contrary, the simulation time with SOLID236 is significantly increased.

Since the heating of the foil might be a critical factor, accuracy is more important than simulation time. Hence, the element type SOLID236 was chosen for the study to optimize the design of the foil-feeding device.

3.2 Simulation Study

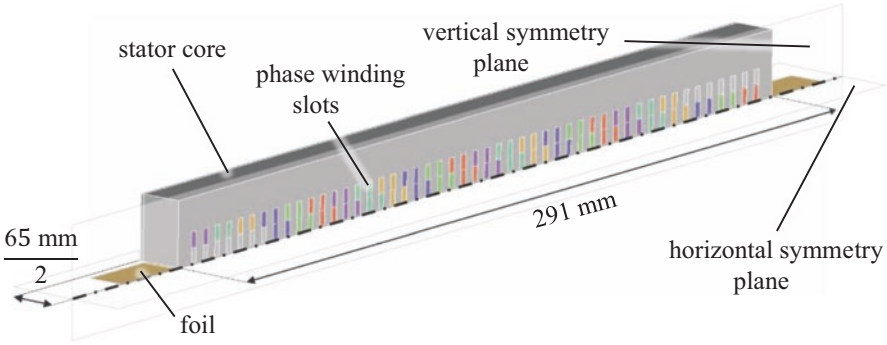


Fig. 4. Three-dimensional FEA-quarter-model of the foil-feeding device in ANSYS.

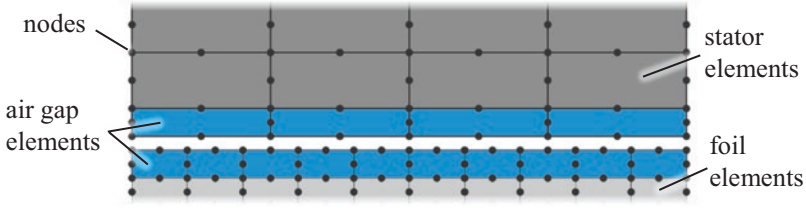


Fig. 5. Splitting of air gap elements between stator and foil.

It has been mentioned above that temperature might be a critical factor in the design of the foil-feeding device. In order to evaluate this, a three-dimensional FEA-model based on the previous work was set up with the addition of a thermal simulation,

consisting of a stator core with a length of 291 mm and a width of 65 mm, the air gap and foil (Fig. 4). The elements used for electromagnetic simulation were SOLID236 and for thermal simulation SOLID90. In order to reduce the calculation time, again a quarter model was set up while using the symmetry of the setup and appropriate boundary conditions. The symmetry planes are shown in Fig. 4. The elements, that represent the air gap between the foil and the stator core have been split into two halves. The upper half is tied to the stator part and the other more dense half is tied to the foil (Fig. 5). By this measure, the simulation of the foil moving relative to the stator becomes available. The meshing of the model produced a total of 20748 elements. The resulting element size for the foil is $5 \times 5\ \text{mm}^2 \times 0.5$ mm. The size of the stator elements is $5 \times 5 \times 5\ \text{mm}^3$. A simulation study has been carried out in which an aluminium foil with a thickness of 500 µm moves through the feeding device with a velocity of 0.03 m/s. Hence, the section of foil is moved through the entire feeding device in 9.7 s. With this parameter choice, an investigation of the foil heating, isolated from application purposes, can be executed. The force necessary to speed up the foil section in the feeding device within 0.1 s is based on the foil mass that is located between the stators during feeding. The force calculates to

$$F = m \cdot a = m \cdot \frac{v}{\Delta t} = 0.0255\ \text{kg} \cdot \frac{0.03\ \text{m/s}}{0.1\ \text{s}} = 7.65\ \text{mN}$$

As parameters for the excitation in the coil, a peak current of 1.1 A and a frequency of 90 Hz have been set. The calculation was carried out as a harmonic simulation, which calculated the effects during the passing of one period of the excitation current. During the movement, the temperature of a fixed point in the foil is observed. In this way, the heating of the section of foil can be calculated as it passes through and possible overheating of the material can be predicted. The initial temperature was set to 22 °C. The study in ANSYS consists of two parts. First, in an electromagnetic analysis with element type SOLID236 the thermal energy due to eddy currents as well as the resulting feeding force is calculated. In a following transient thermal analysis with element type SOLID90 the temperature distribution due to the generated thermal energy is calculated. The simulation model has calculated, that the set parameters lead to a feeding force of 45 mN, this ensures that the chosen input delivers sufficient feeding force. The results of the simulation are shown in Fig. 6. It displays the max. Temperature at any point of the foil at the given timesteps. It heats up to a temperature of about 22.28 °C. The thermal energy loss in the entire foil is about 78 mW per current period which is a direct output from the Ansys calculation. It is calculated by summing up the heating of every section of the foil. The total loss due to the heating of the foil therefore amounts to

$$P_{\text{L,H}} = 4 \cdot 78\ \text{mW} \cdot 9.7\ \text{s} = 3.0264\ \text{J}.$$

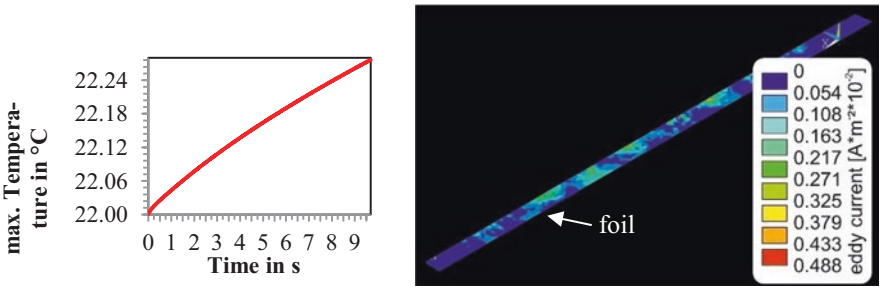


Fig. 6. Maximum temperature in the foil for the proposed load case and corresponding contour illustration of eddy-currents.

The distribution of force in the foil is displayed in Fig. 7. It illustrates the force in feeding direction (x-direction) in every volume element of the foil. It can be seen that the electromagnetic force is applied rather inhomogeneously along the foil. This will be discussed in the next section.

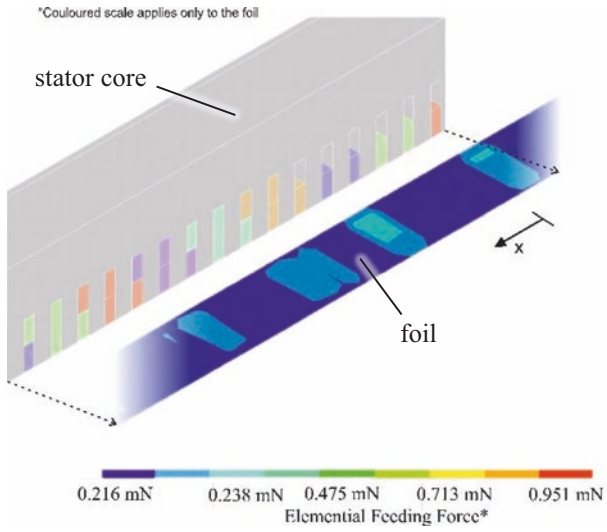


Fig. 7. Inhomogeneous application of force in the foil.

4 Discussion and Future Work

Foils are fragile and sensitive workpieces, thus the simulation needs to be highly accurate. Therefore, an FEA model which uses state-of-the-art element types was developed. The simulation results show that the foil material heats up passing through the feeding device due to eddy currents. However, the low mass of the foil is the salient point that allows to work with extremely low currents of 1.1 A for excitation and

thus reduce heat generation in coil and foil due to eddy currents. This is just a small fraction of the exciting currents that were mentioned in [4] and [5]. As a result, no critical temperatures are reached, which could lead to melting of the foil or damage to surface coatings, for example. This leads to the conclusion that the heating of the foil is not as critical as initially assumed since the heated foil section leaves the feeding device before it reaches damaging temperatures. The simulation results indicate that high feeding rates for foils can be achieved by electromagnetic feeding devices without causing overheating. Nevertheless, this aspect still needs to be confirmed in practical experiments. The simulation method will also be applied to thinner foil material to determine whether there are limitations when an unacceptable feeding force – heat loss ratio sets in.

The total feeding force of the foil is determined by summing up the feeding force of every volume element. It calculates to a total feeding force of over 45 mN, which exceeds the calculated necessary force from the example mentioned above. However, the simulations also show that the force applied to the foil is not homogeneous. As pointed out in the introduction, inhomogeneous forces lead to crease, fold or even tear due to the fragile structure of the foil. This problem is an important issue for future research. However, the illustration in Fig. 7 shows the feeding-force distribution in one single moment rather than over time. Future simulations will show the force distribution over time by means of transient simulations. The results will show how the impact of force moves along the foil and whether or not this leads to creases. A possible design shift for the unequally distributed force could be a higher pole number with a finer tooth pitch or pole pieces at each tooth to distribute the magnetic field in the air gap more evenly.

Acknowledgement. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number: 388386231 (KR3718/7-1)

References

1. Behrens, B.-A., Doege, E.: Handbuch der Umformtechnik. Springer, Berlin (2016). ISBN 978-3-662-43890-9
2. Razali, A.R., Qin, Y.: A review on micro-manufacturing, micro-forming and their key issues. Procedia Eng. **53**, 665–672 (2013)
3. Behrens, B.-A., Krimm, R., Teichrib, S.: Contactless feeder for electrically conductive sheet metals. Prod. Eng. Res. Dev. **11**, 1–8 (2016)
4. Kolbe, M., Hellwig, W.: Spanlose Fertigung – Stanzen, 12th edn. Springer Vieweg, Wiesbaden (2018)
5. Teichrib, S., Krimm, R., Behrens, B.-A.: Contactless sheet metal feeder for metal-forming machines. In: Proceedings of the 3rd International Conference on Mechanical Engineering and Mechatronics, , Prague, Czech Republic, pp. 107.1–107.8 (2014)
6. Commichau, O., Krimm, R., Behrens, B.-A.: Electromagnetic feeding device for conductive material. In: Proceedings of the 8th Congress of the German Academic Association for Production Technology (WGP), Aachen, 19–20 November 2018
7. Hanke, M., Otto, J.: Elektromechanische Antriebe, simuliert mit ANSYS. CADFEM Infoplaner **1**, 30–32 (2009)

8. ANSYS, Inc.: ANSYS Mechanical APDL Element Reference. SAS IP, Inc., Canonsburg (2012)
9. Zienkiewicz, O.C., Taylor, R.L.: The Finite Element Method, vol. 1, 5th edn. The Basis, Butterworth Heinemann, Oxford (2000)
10. Banerjee, B., Chen, Z.J., Kathirgamanathan, A., Das, R.: Comparison of ANSYS elements SHELL181 and SOLSH190 (2011)
11. ANSYS Inc.: ANSYS Mechanical APDL Command Reference. SAS IP, Inc., Canonsburg (2012)
12. Hämmerle, F.: Simulation and Characterization of a Miniaturized Planar Coil. Master's Thesis, Vorarlberg University of Applied Sciences, Dornbirn (2009)

Extension of Process Limits with Bidirectional Deep Drawing

S. Kriechenbauer(✉), P. Müller, R. Mauermann, and W.-G. Drossel

Fraunhofer IWU, Nöthnitzer Straße 44, 01187 Dresden, Germany
sebastian.kriechenbauer@iwu.fraunhofer.de

Abstract. Bidirectional deep drawing is a process for sheet metal forming on servo screw presses in order to influence part properties in a targeted manner and to extend process limits. With servo drives, it is possible to freely change the motion and force during the forming process. Experimental and numerical investigations show the extension of the process limits. In deep drawing, the process forces are indirectly transmitted from the flange via the part frame to the bottom area. In this process, a crack may occur near to the bottom of the part due to excessive thinning. In bidirectional deep drawing, the weak point can be strengthened with a suitable motion profile. Thus, the drawing depth can be deeper, because higher process forces can be applied. The design of the motion profile in bidirectional deep drawing is supported by a variance-based sensitivity analysis. This allows the potential of servo screw presses for sheet metal forming to be better exploited.

Keywords: Servo press · Bidirectional deep drawing · Free motion

1 Process Description

With deep drawing of sheet metal parts, the local weak point is mostly located at the outlet of the bottom radius (Fig. 1). This is where work hardening is at its lowest and the stress from the process load is at its highest level [2]. The sidewall of a cup indirectly transmits the drawing force from the flange to the bottom of the part. The sheet metal becomes thinner due to the tensile load in the sidewall. If the stressability of the deep drawn part exceeds the limits, then e.g. a bottom crack will occur at the weak point due to excessive thinning [2].

Bidirectional deep drawing starts before conventional deep drawing with constant blankholder force. The aim of bidirectional deep drawing is to increase work hardening at the weak point by bending back and forth without damaging the deep drawn part. As a result, a higher yield stress is necessary to plastically deform the material. The increased stressability of the part allows using higher process loads. This allows the process limits for deep drawing to be extended, e.g. to produce deeper components.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 96–104, 2021.
https://doi.org/10.1007/978-3-662-62138-7_10

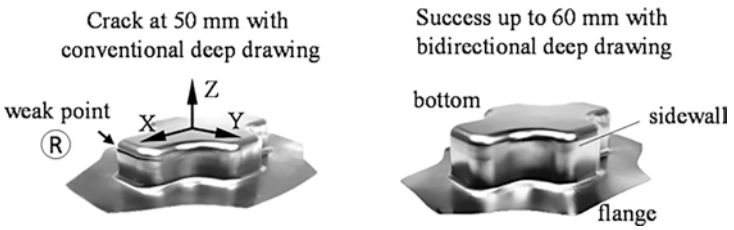


Fig. 1. Parts with conventional (left) and bidirectional (right) deep drawing

Figure 1 shows a comparison for a cross cup made out of 1.4301 (X5CrNi18-10) with a sheet thickness of 1 mm. The dimensions of the blank and the tool setup show Figs. 2 and 3. For bidirectional deep drawing, a counter punch is required to fix the bottom of the drawn part during reverse bending.

In contrast to the complex motion function described in [1], here BDD uses a simpler motion function (Fig. 2). During bidirectional drawing the sheet is bent around the drawing radii (Fig. 3 middle) with an amplitude $\hat{z} = 10$ mm and a cycle $N = 1$ against the drawing direction Z.

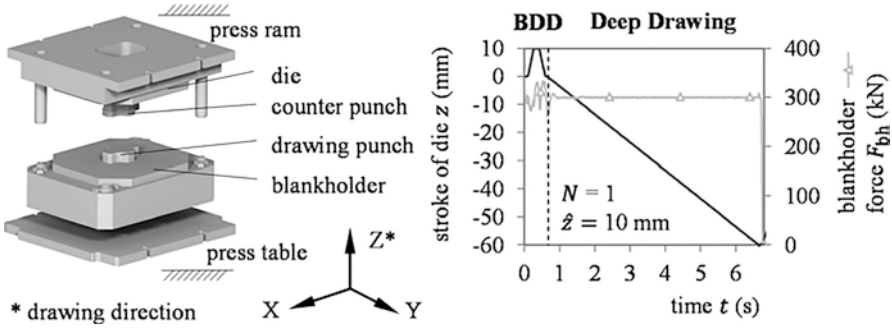


Fig. 2. Tool setup (left) and process functions (right) for bidirectional deep drawing

In bidirectional deep drawing, inertial forces influences the blankholder force due to the dynamics of the process. Therefore, the blankholder force fluctuates around the constant nominal value of 300 kN during bidirectional deep drawing (Fig. 2). The process simulation does not consider this effect.

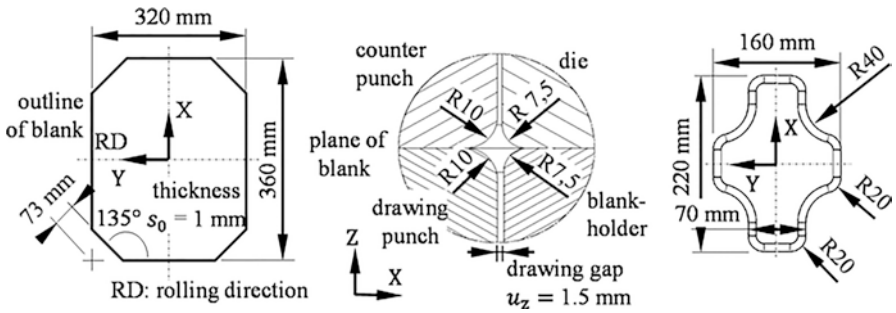


Fig. 3. Geometry of blank (left), tool radii (middle) and drawing punch (right)

2 Material Model

The flow curve describes the hardening behavior of the material. The examined material 1.4301 with metastable austenite forms deformation-induced martensite during deep drawing [3]. The simulation software LS-DYNA implements the non-isothermal material model in MAT_133 with the parameter HARD = 3.0. With a thermo-mechanically coupled deep-drawing simulation, the formation of martensite can be calculated as a function of temperature and the already formed martensite content. The process simulation in LS-DYNA uses the material parameters from Table 1. The material model was calibrated on the measured martensite volume evolution from adiabatic tensile tests with different strain rates [4].

Table 1. Isotropic hardening parameters of 1.4301 for MAT_133 in LS-DYNA determined from measurements in [4]

Hansel Hardening Card 1							
A	B	C	D	P	Q	E0MART	VM0
1.39	0.168	−47.892	0.0	7.895	1376.15	0.13	0.3
Hansel Hardening Card 2							
AHS	BHS	M	N	EPS0	HMART	K1	K2
0.3508455	7.90974625	0.25782	0.80353	0.02	0,4925	2.49723	−0.00511

The alternating motion during bidirectional deep drawing results in tensile and compressive stresses in the part. Therefore, the simulation must consider the kinematic hardening behavior of the material. In LS-DYNA the material model according to Chaboche-Rousselier in MAT_133 is suitable for the description of the examined material 1.4301. The parameters are determined from tension-compression tests. Table 2 shows the parameters of the material model MAT_133 in LS-DYNA.

Table 2. Cinematic hardening parameters of 1.4301 for MAT_133 in LS-DYNA

Chaboche-Rousselier Card							
CRC1	CRA1	CRC2	CRA2	CRC3	CRA3	CRC4	CRA4
14.95600	0.197128	293.6710	0.091742	0	0	0	0

In sheet metal forming, high strains can result in failure e.g. a crack. For conventional deep drawing with constant blankholder force, the forming limit curve has proven to be a good method for failure prediction [2]. The forming limit curve (FLC) is determined experimentally with the Nakajima test [2]. The specimen geometries provides forming with linear strain path until failure.

Non-linear strain paths due to alternating bending, however, characterize bidirectional deep drawing. Figure 4 shows examples of typical strain paths on the component surface. Therefore, failure prediction in bidirectional deep drawing is not possible with forming limit curve, which is based on linear strain paths.

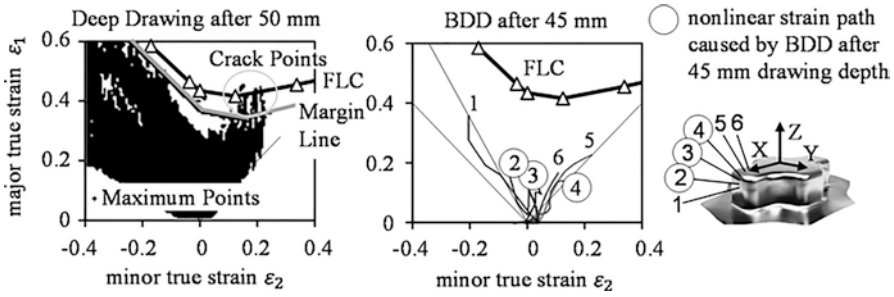


Fig. 4. Forming limit curve (FLC) of 1.4301, formability of deep drawing and nonlinear strain paths caused by BDD

3 Process Model

In LS-DYNA, a thermo-mechanically coupled process simulation was performed for the deep-drawn parts from Fig. 1. The process model is symmetrical with respect to the XZ and YZ plane. Surfaces build the model for deep drawing tool. Figure 5 shows the process model with the boundary conditions. During conventional deep drawing a constant speed and during bidirectional deep drawing a motion function are specified for the die. Shell elements in the dimensions 1 mm by 1 mm define the blank mesh. In LS-DNYA, Belytschko-Lin-Tsay shells of type 2 use 10 integration points distributed over the sheet thickness according to Gauss integration rule. The thermic boundary conditions are chosen from [4] and summarized in Fig. 5.

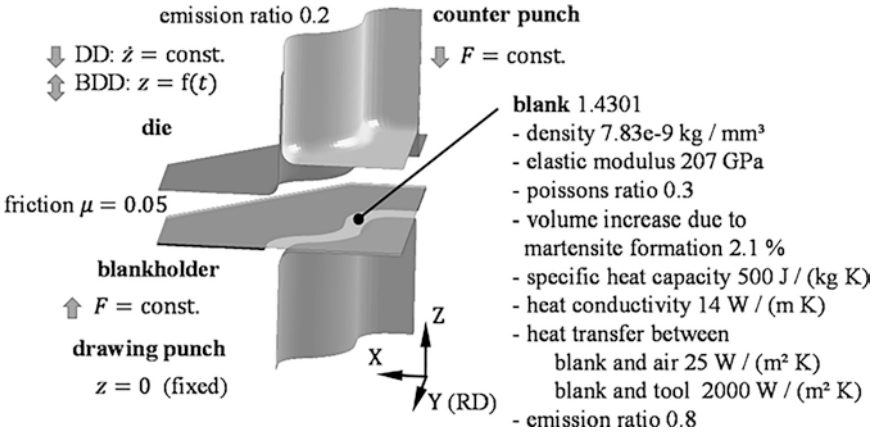


Fig. 5. Process model for thermo-mechanical simulation of symmetric cross die

In the simulation, bidirectional deep drawing can increase the work hardening and the formation of martensite at the weak point. Figure 6 shows a comparison between conventional and bidirectional deep drawing at a depth of 45 mm for the cross cup. It shows the maximum values for the integration points over sheet thickness. With bidirectional deep drawing, a depth of 60 mm is achievable Fig. 1.

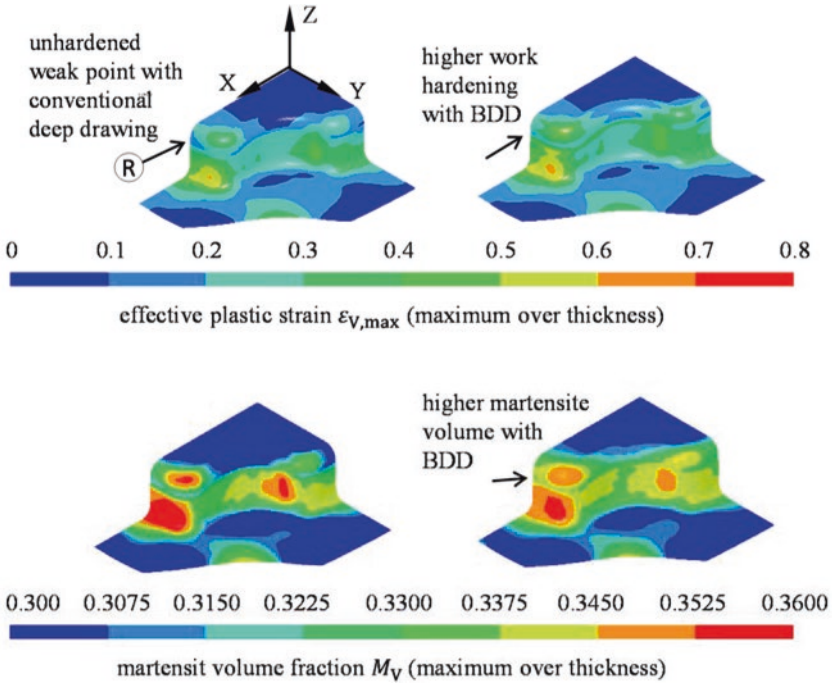


Fig. 6. Simulation results at 45 mm drawing depth

At the beginning of the process, the bidirectional deep drawing operation forms a circumferential fold. The subsequent deep drawing process smoothes out this fold. The Fig. 7 shows the sheet cross section in the deep drawing tool for different motion functions. The cross section consists of all through thickness integration points.

In bidirectional deep drawing, small amplitudes lead to low work hardening. Large amplitudes cause excessive folding, which results in a failure during deep drawing.

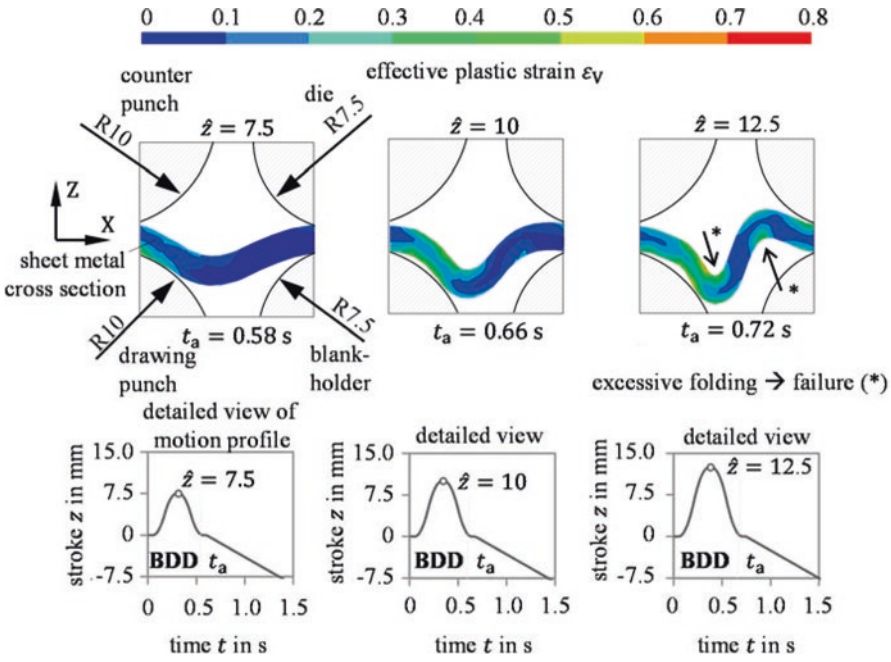


Fig. 7. Simulation result at the end of the bidirectional deep drawing with various motion profiles

4 Sensitivity Analysis

The influence of the process parameters on the deep drawn part can be evaluated with a sensitivity analysis [5]. The Fig. 8 shows the variable stroke and the constant blankholder force. In a first step, the sheet is bent downwards or upwards around Δz. From this starting position, the sheet is bent alternately upwards and downwards with the amplitude $\hat{z}$. After N cycles, the component is fully deep-drawn to a depth of 45 mm. The blankholder force remains constant during the whole process.

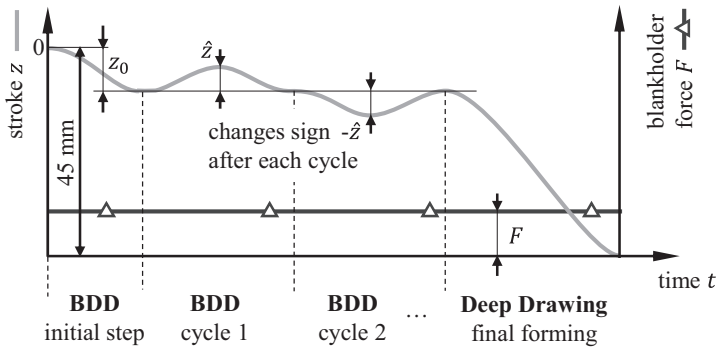


Fig. 8. Variable stroke and constant blankholder force evaluated in sensitivity analysis

Table 3 summarizes all influence parameters, which are varied within the technological limits. The reference values belongs to the validated process model from Sect. 3. The sensitivity analysis consists of 160 simulated designs generated with Latin Hypercube Sampling.

In addition to the process parameters, friction and the flow curve are also investigated. With a simple scaling factor for the yield stress, the strength of the material continuously varies in the sensitivity analysis. As a result, an engineer must identify a suitable material with a measured flow curve near the calculated one.

Table 3. Setup of the senstivity analysis of the bidirectional deep drawing process

	Parameter		Reference value	Range	Unit
Influence Parameters	Amplitude	$\hat{z}$	10	0 … 10	mm
	Cycles	N	1	1; 2 … 10	
	Initial motion	Δz	0	−10 … 10	mm
	Blankholder force	F	300	50 … 500	kN
	Friction coefficient	μ	0.05	0.05 … 0.15	
	Scale factor of flow curve	S	1.0[a]	0.8 … 1.2	
Results	Sheet thickness	s			mm
	Effective plastic strain (Maximum over Thickness)	$\varepsilon_{V,max}$			

[a]corresponds to the material model in Sect. 2 based on 1.4301

According to [5], the field metamodel describes the relationships between the influence parameters and the result variables distributed over the whole component. That field metamodel predicts the maximum achievable degree of deformation and the resulting sheet thickness. Both field metamodels show a high model accuracy (Fig. 9) in the critical area.

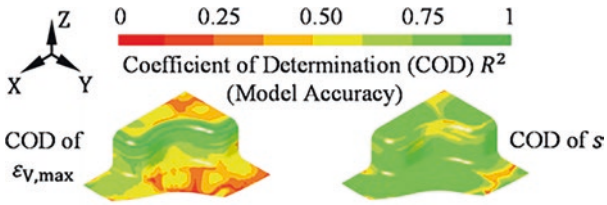


Fig. 9. Model accuracy of field metamodel for maximum of effective plastic strain (left) and for sheet thickness (right) at 45 mm drawing depth

The Fig. 10 shows different results generated from the field metamodels for selected process parameters, other influence parameters are hold constant at the reference values (Table 3). The effective plastic strain varies with different motion functions. The parameters in Fig. 10a build the conventional deep drawing process, which is shown in Fig. 6. The effective plastic strain increases with a larger amplitude and further cycles, as the predicted values from the field metamodel in Fig. 10b–f demonstrate.

The corresponding sheet thickness in Fig. 10b–f decreases in the critical area (R) with BDD. That is why larger drawing depths can be reached because higher process loads can be transmitted over the weak point before excessive thinning will start. Nevertheless, failure cannot be evaluated with FLC due to the nonlinear strain paths in bidirectional deep drawing. LS-DYNA has no option in MAT_133 to calculate a formability index for failure detection with FLC.

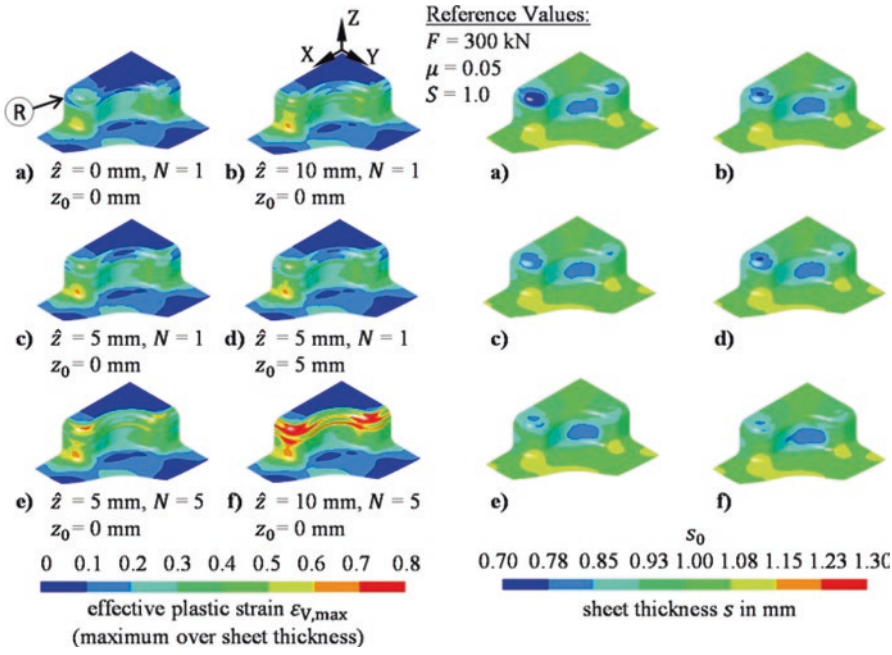


Fig. 10. Result of the sensitivity analysis with different motion functions at 45 mm drawing depth

The sensitivity analysis is suitable for the development of the bidirectional deep drawing process. The results provide suitable parameters, which can be used to harden the weak point in the workpiece. The amplitude must not be too large. The initial position, before the alternating bending operation starts, must be adapted to the selected amplitude. The strength can be additionally increased by additional movement cycles. The sensitivity analysis reduces the experimental effort required for the process design.

5 Summary

The experiment and the simulation confirm the extension of the process limits with bidirectional deep drawing. The additional bending motion increases the effective plastic strain at the weak point of a cross cup.

Realistic results for bidirectional deep drawing depends on the exact description of the material behavior with isotropic and kinematic hardening. For the examined material 1.4301, the process simulation must calculate the martensite formation. Without the calculation, the failure cannot be predicted with FLC for the conventional deep drawing process.

During bidirectional deep drawing, nonlinear strain paths occur due to forward and backward bending around the drawing radii. Therefore, failure cannot be evaluated with FLC based on linear strain paths.

A sensitivity analysis, for example, allows determining the influence of different motion functions on the effective plastic strain. Larger amplitudes and further cycles increases the effective plastic strain and the formation of martensite. Nevertheless, the failure cannot be evaluated due to the nonlinear strain paths for bidirectional deep drawing.

References

1. Kriechenbauer, S., Mauermann, S., Landgrebe, D.: Deep drawing with local hardening on digital multi-axis servo press. In: International Deep Drawing Research Group (IDDRG Conference), Shanghai (2015)
2. Altan, T., et al.: Sheet Metal Forming - Fundamentals. ASM International, Ohio (2012)
3. Krauer, J., Hora, P.: Enhanced material models for the process design of the temperature dependent forming behavior of metastable steels. Int. J. Mater. Form. **5**(4), 361–370 (2012)
4. Heinemann, G.: Virtuelle Bestimmung des Verfestigungsverhaltens von Bändern und Ble-chen durch verformungsinduzierte Martensitbildung bei metastabilen rostfreien austenitischen Stählen. Dissertation ETH Zürich (2004)
5. Schwarz, C., Ackert, P., Mauermann, R.: Principal component analysis and singular value decomposition used for a numerical sensitivity analysis of a complex drawn part. Int. J. Adv. Manuf. Technol. **94**, 2255–2265 (2018)

Further Development of a Hydraulically Operated Oscillation Device for Application to an Industrial Forming Process

P. Müller[1(✉)], D. Rosenbusch[1], N. Missal[2], H. Vogt[1], S. Hübner[1], and B.-A. Behrens[1]

[1] Institute of Forming Technology and Machines, Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany
mueller@ifum.uni-hannover.de
[2] Felss Systems GmbH, Dieselstraße 2, 75203 Königsbach-Stein, Germany

Abstract. A superimposed oscillation in the main force flow of a forming process is investigated. Therefore a tool system was developed. The oscillation in this system is generated by pulsating oil streams. As part of this research, the oscillation device is equipped with additional storage tanks, to generate a higher oil volume flow. The upgraded oscillation system is installed in a forming process. In this research, the influence of the superimposed oscillation on the necessary plastic work to form a complex gear geometry is examined. An analysis of the influence of higher forming forces, higher operating pressures and longer process times on the operating behavior of the oscillating device also takes place. The identified optimal process parameters are used in a model process based on the industrial forming process of the company Felss Systems GmbH. Furthermore, an experimental test setup to carry out these investigations is developed in this research. Forming experiments to iron an external gear geometry oscillation-free and superimposed oscillated are conducted with the new test stand.

Keywords: Sheet-bulk metal forming · Oscillation superimposed forming · Dynamic process forces

1 Introduction

A new production technology combines the techniques of sheet-metal forming and bulk-metal forming to sheet-bulk metal forming (SBMF) [1]. At the Institute of Forming Technology and Machines (IFUM) a SBMF process was developed to extend the technical boundaries of this new technology [2]. Thereby, a component with an internal and external gearing is produced from a sheet ronde in three production steps. An essential aspect is the investigation of a superimposed oscillation during the forming of the internal gearing in the second production step [2]. For this purpose, a hydraulically working oscillation device is used. The device is able to generate

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 105–115, 2021.
https://doi.org/10.1007/978-3-662-62138-7_11

frequencies of up to $f = 600$ Hz and displacement amplitudes of up to $A = 50$ µm [3]. In the first production step of the SBMF process, a cup is deep-drawn from a sheet round blank. After that, an outer gearing is formed by compressing the cup wall into the cavity of the tool die. During the compression, forming forces of up to 600 kN occur when the form filling is almost completely achieved [4]. This made a previous use of the oscillation device for this process step impossible because necessary high system pressures and process times could not be achieved to handle such high forming forces. The positive influence of an oscillation overlay on the parameters forming force and surface quality could be proven in previous studies. BLAHA et al. discovered the phenomenon of yield stress reduction during a superimposed oscillation in the ultrasonic range back in 1959 [5]. This is partly attributed to material-related effects such as oscillation-induced temporary acoustic softening of the material [6]. In addition, experiments on smoothing of rough surfaces during superimposed oscillation forming, which lead to a reduced friction between sample and die has been investigated in previous studies [7]. Furthermore, the influence of superimposed oscillation in combination with lubricants were part of past researches. MAENO et al. observed optimized lubrication conditions by the installation of a superimposed oscillation in a plage forging process [8]. A low friction due to improved lubrication conditions induced by pulse ram motions was also demonstrated by MATSUMOTO et al.. Nevertheless, the authors also found increased tool wear due to the applied oscillating movements [9]. However, most researches took place in the low frequency range of a few hertz [10] by applying the superimposed oscillation with the press drive [11] or in the very high frequency range of several kilohertz by applying the superimposed oscillation with ultrasonic oscillation devices [12]. Experiments in the middle frequency range (of 100–600 Hz) were mostly conducted as basic investigations [13]. An example of an industrial application is the use of superimposed oscillation adapted on a deep drawing process by JIMMA et al. [14]. In order to make the oscillation device at the IFUM more applicable for industrial implementation, the existing hydraulics will be upgraded with additional storage tanks. Thus, a provision of a higher system pressure is ensured. The device is installed in the first stage of the SBMF process. Compression tests are conducted both statically and superimposed oscillated. The oscillation parameters frequency and amplitude are varied. The influence of a superimposed oscillation on the plastic work necessary for compressing a complex gear geometry is investigated. In addition, the influence of the high system forces and longer process run times on the operating behavior of the oscillating device is considered. Concerning the forming behavior in combination with the process parameters, the optimal operation point is identified. An experimental test stand for the reconstruction of an industrial ironing process to form an axial gear geometry superimposed oscillated is developed and put into operation. The effects of oscillation-induced forming force reduction and form filling enhancement are investigated in this process. Therefore, the quality of the components is improved by an oscillation induced increase of the form filling.

2 Experimental Setup

The experiments were conducted at the hydraulic press Wanzke GTA 250. The hydraulic oscillation device is positioned under the tool system of the first stage of the SBMF-process. The tool assembly consists of a tool die, a deep drawing stamp and a compression stamp (Fig. 1). The oscillation device consists of a rotary piston valve, an electric motor and a high-frequency hydraulic cylinder.

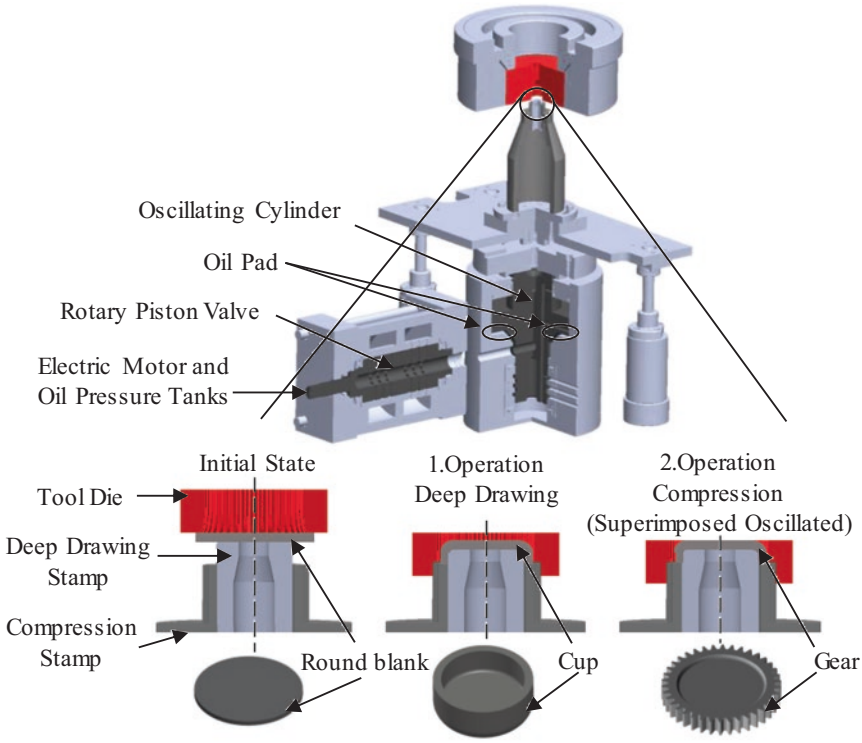


Fig. 1. Enhanced tool system for superimposed oscillated compression experiments.

The rotary piston valve is supplied with a volume flow of hydraulic oil by pressure tanks. The tanks contain a volume of 200 l at a pressure of 120 bar. The pressure in the oil pad under the high frequency cylinder is adjusted by controlling the volume flow in front of the rotary piston valve. The frequency of the oscillation is set by the rotation speed of the electric motor connected to the rotary piston valve. The experiments are conducted on a tool assembly, which compresses a gear. Therefore, a tool die, which is installed under the pestle of the press, moves down to a sheet metal blank positioned on a deep-drawing stamp. First, a cup is deep-drawn. Subsequently, the oscillating device installed under the tool structure is switched on and an outer gearing is superimposed oscillated compressed by displacing the cup frame into the

gear cavity of the die. As wrought material, blanks made of the deep-drawing steel DC04 with a diameter of 42 mm are cut out of a 2 mm thick sheet by waterjet cutting. The used lubricant is Beruforge 152 DL. Both the pressure in the oil pad and the frequency of the oscillation are varied. The amplitude results indirectly from the setting of the frequency and can therefore not be varied separately. Several gears are compressed to different depths to analyze individual process parameters. As a reference, gears are compressed without superimposed oscillation to different depths. In order to avoid an overloading of the oscillating device, a maximal forming force of 360 kN is used. The achievable tooth flanks height is 4.35 mm. During the superimposed oscillation experiments, the high-frequency cylinder is displaced downwards by the stamp. The displacement is a function of frequency, pressure and increasing forming force. When the piston reaches its lowest stopping point, the hydraulic pad collapses and the superimposed oscillation stops abruptly. This effect is to be avoided. Therefore, the displacement of the piston is measured in the process. The displacement of the piston depends on the pressure, frequency and forming force. The lower the displacement of the piston, the better are the process parameters for the oscillating device. The plastic work required to form the components is directly dependent on the oscillation parameters. The forming behavior of the samples is evaluated with regards to the reduction of the plastic work necessary for forming. Therefore, the static and superimposed oscillation process parameters that leads to the same height of the tooth flanks, are compared.

The press speed is constantly set at 16 mm/s. The displacement of the high-frequency cylinder is continuously measured in addition to the force pathways. Table 1 shows the operating parameters of the oscillation device that are applied during the experiments. From previous investigations, the settings in Table 1 have emerged as promising parameters for the operation of the oscillation device [13].

Table 1. Experimental parameters for the superimposed oscillation compression tests.

Oscillation Parameters						
Frequency [Hz]	100	200	200	100	200	300
System Pressure [bar]	50	50	65	100	100	100

3 Experimental Results

In the following chapter the compression tests are evaluated. One aspect is the analysis of the operating behavior, which is evaluated on the basis of the displacement of the high-frequency cylinder into the hydraulic oil pad. Another aspect is the analysis of the forming behavior, which is evaluated on the basis of the plastic work required for forming.

3.1 Operating Behavior of the Oscillation Device

Figure 2 shows the displacement of the high-frequency cylinder in dependence of the system pressure, the oscillation parameters and the forming force. It can be seen that

the high frequency cylinder is already highly displaced by relatively low forces in the range of 100 kN as well as at low system pressures and low frequencies of the oscillation device.

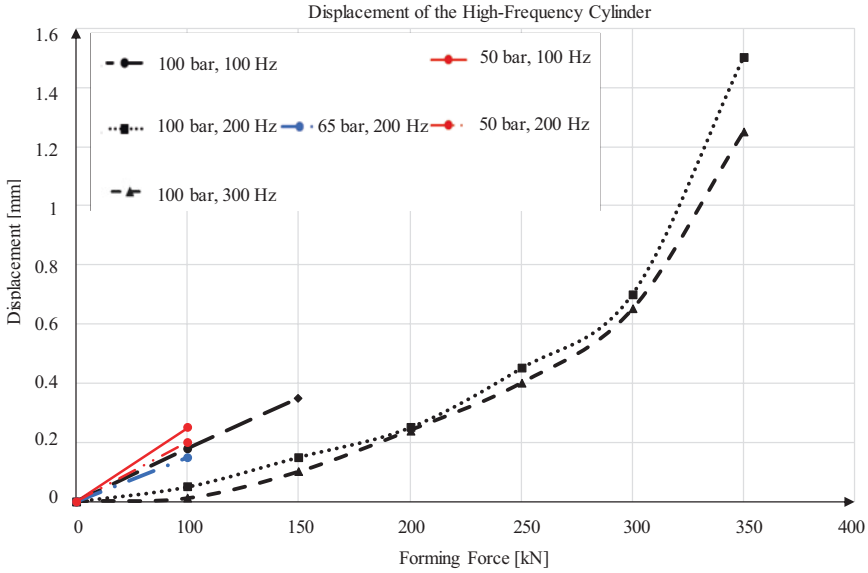


Fig. 2. Displacement of the high-frequency cylinder in dependence of the forming force and the oscillation parameters.

In order to reduce the risk of an abruptly collapsing oil pad, the experiments in these operating modes are already interrupted in a force range of about 100 kN. When the frequency and system pressure is increased, the pressure pad does not collapse rapidely and higher forming forces can be achieved. At a pressure of $p = 100$ bar and a frequency of $f = 200$ Hz, the cylinder is only slightly displaced (in the range of about 0.42 mm) when the forming force achieves a value of 250 kN. As the forming force continues to increase, the displacement increases exponentially until it reaches a value of around 1.45 mm at a forming force of 350 kN. When the frequency is increased to 300 Hz while maintaining the system pressure at $p = 100$ bar, the characteristic curve measured at 200 Hz shifts to higher forces for 300 Hz (Fig. 2). So a cylinder displacement of 1.22 mm is achieved at a forming force of 350 kN. With regard to the stability of the operating behavior, a frequency of 300 Hz and a system pressure of 100 bar is the optimal of the investigated operating parameters.

3.2 Influencing the Plastic Work

In addition to the investigations of the operating behavior of the oscillation device during the compression tests, the forming behavior is also analysed. The used evaluation criterion is the plastic work necessary for forming. For this purpose, the force and

the displacement are recorded during forming with a laser sensor and a load cell. The plastic work can be determined from the force path diagram using the mathematical context according to Eq. 1 [15].

$$W = \int_{s_o}^{s} F \, ds \tag{1}$$

With:	W = plastic work	s_0 = initial state
	F = forming force	s = forming path

The results of the evaluation for an operating pressure of p = 100 bar with a maximum forming force of 380 ± 20 kN are shown in Fig. 3, once without a superimposed oscillation and once with a superimposed oscillation frequency of $f = 200$ Hz and $f = 300$ Hz. The components were all compressed to an equal tooth height around 4.35 ± 0.04 mm.

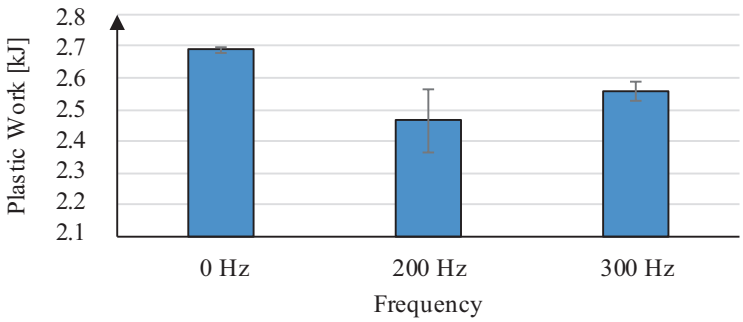


Fig. 3. Plastic work necessary for forming in dependence to the oscillation parameters.

As can be seen in Fig. 3, the plastic work decreases by about 220 J at a superimposed oscillation of $f = 200$ Hz compared to the static reference tests. This effect of plastic work reduction due to a superimposed oscillation was also observed in previous investigations [10]. Figure 3 shows also that with a further increase in frequency (up to $f = 300$ Hz) an increasing of the plastic work occurs. This is explained by the different oscillation profiles generated by the system at constant pressure and different rotatory speeds of the rotary piston valve. In the following, the results are not evaluated according to the displacement amplitude but according to the force amplitude since the plastic work is calculated via the force progression (see Eq. 1). Figure 4 shows that the force amplitude of the oscillation decreases with an increase of the frequency from $f = 200$ Hz to $f = 300$ Hz. According to BLAHA et al. [5], the plastic work necessary for forming is primarily dependent on the amplitude of the applied superimposed oscillation. This statement is confirmed by the results shown here. The

observed operating behavior can be explained by the fact that an increased frequency leads to a decrease of the pressure pad decomposition, and therefore a decrease of the amplitude. With regard to the forming behavior, the operating point with an operating

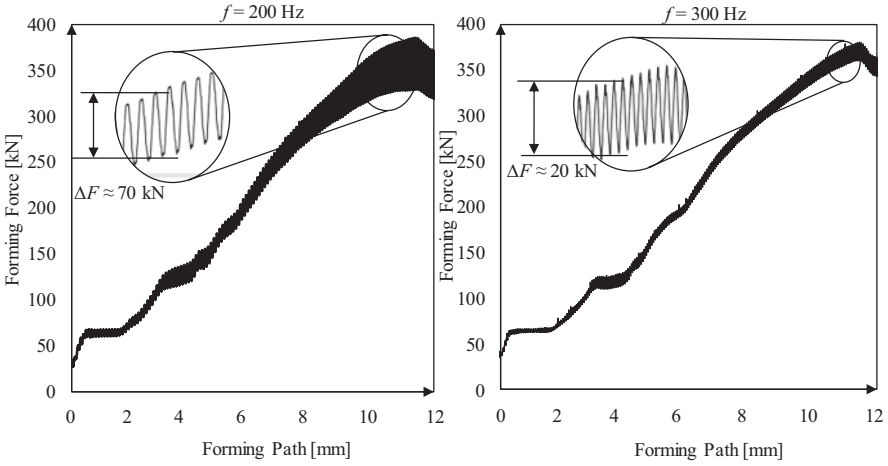


Fig. 4. Force path of the superimposed oscillation compression experiments once with $f = 200$ Hz (left) and once with $f = 300$ Hz (right).

pressure of $p = 100$ bar with a frequency of $f = 200$ Hz therefore turns out to be the optimal state.

3.3 Development of the Experimental Test Stand

By identifying the optimal process parameters, the results can be transferred to an industrial-like process for the superimposed oscillation forming of an axial gear geometry of the company Felss Systems GmbH. Therefore, an experimental tool system is developed, which will be installed on the hydraulically working oscillation device. The tool system consists of a tool die, which is used by Felss Systems. The die is positioned over the oscillating device. A clamping device in which a hollow shaft is fixed (Fig. 5) is installed under the pestle. In the process, the hollow shaft is positioned in the clamping device and afterwards inserted into the oscillating die to form an external gear geometry. The CAD-model of the tool system is shown in Fig. 5.

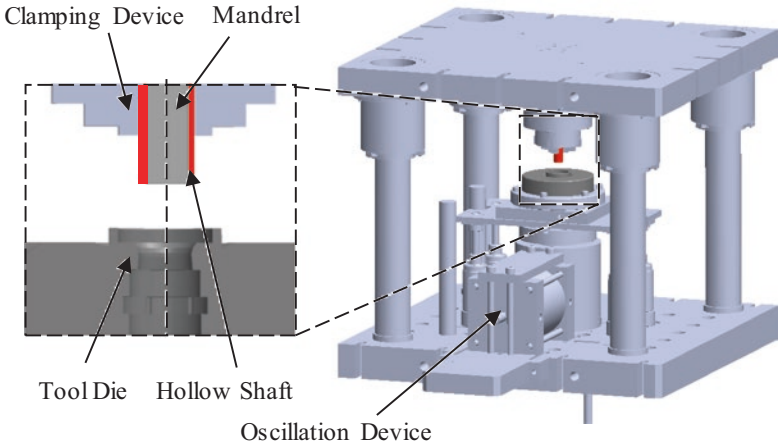


Fig. 5. Tool system for superimposed oscillation experiments for the forming of an axial geometry on a hollow shaft

First forming experiments are conducted with the new test stand. The steel 1.7139 is used as sample material for the hollow shaft (see Fig. 5). The lubricant Multidraw of the company Zeller and Gmelin is used as lubricant. The experiments are conducted at the hydraulic press Schirma und Plate at the IFUM. The press speed is set constant at 30 mm/s. Due to the completely different test conditions chosen in comparison to the previous chapter; it can be determined whether the identified optimum operating parameters of the oscillation device is applicable across processes. The operating parameters 200 Hz, 100 bar and 300 Hz, 100 bar are selected, which were identified as the most promising in Fig. 2. In addition, oscillation-free reference tests are performed. The evaluation of the force curves and the final component geometry is shown in Fig. 6. The results confirm that for 200 Hz and 100 bar the largest reduction of the average forming force can be achieved (in the range of 15% compared to the oscillation-free forming). Therefore, it is stated that 200 Hz and 100 bar represents the best operating point across all processes with regard to a reduction in plastic work.

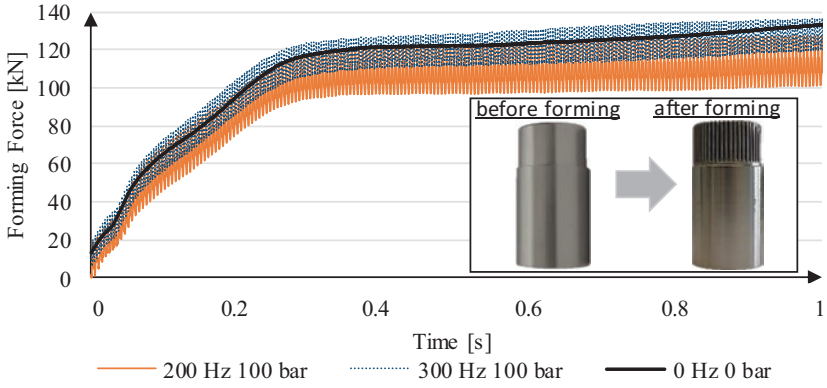


Fig. 6. Force curves for oscillation-free and superimposed oscillated extrusion experiments for the forming of an axial gear geometry.

In addition, the formed components are scanned on the optical measuring macroscope VR3200 of the company Keyence. The determined teeth heights of the gearing elements are shown in Fig. 7 as a function of the oscillation parameters.

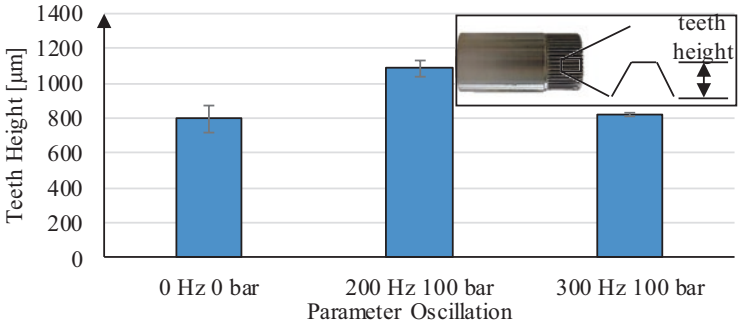


Fig. 7. Teeth heights of the gearing elements.

Considering the form filling, the process parameter with $f = 200$ Hz and $p = 100$ bar also shows the strongest influence. The teeth height increases in a range of 34% compared to oscillation-free forming. The operating point at $f = 300$ Hz and $p = 100$ bar demonstrates only a minor influence with regard to the form filling. This can be explained by a more pronounced oscillation at 200 Hz (see Fig. 6). Therefore, oscillation-induced effects for improving the material flow as described by BLAHA et al. [5] are increasingly introduced into the material.

4 Summary and Outlook

In this research, a sheet-bulk metal forming process for producing an external gearing on a demonstrator component was equipped with a hydraulically working oscillation device. The aim of this research was to expand the limits of superimposed oscillation forming with regards to the achievable process forces. In addition, an optimal operating point for the performance and forming behavior of the oscillation device was aimed. Regarding the operational performance, the displacement of the high-frequency cylinder was considered. Here, an operating pressure of $p = 100$ bar with a frequency of $f = 300$ Hz was identified as the optimal parameter. This is explained by the fact that the pressure pad under the high-frequency cylinder cannot collapse rapidly when the system pressure and the frequency is set high. Considering the forming behavior, forming forces around 380 ± 20 kN could be realized. At a pressure of $p = 100$ bar and a frequency of $f = 200$ Hz the optimal operating point has be proven with regards to a reduction of the plastic work necessary for forming. This is because the pressure pad has more time to drop when the frequency is reduced. This resulted in a larger amplitude, which favors the forming behavior. Since the operating behavior changes only slightly between $f = 200$ Hz and $f = 300$ Hz at a pressure of $p = 100$ bar, it can be concluded that a frequency of $f = 200$ Hz and a pressure of $p = 100$ bar represents to be the optimal operating parameter. This is confirmed by the development of a new experimental test stand for oscillation-free and superimposed oscillated forming experiments. Ironing experiments with this test stand show that the identified optimal operating point of the oscillation device is the most promising across processes. In addition, the form filling of the gearing elements produced in the ironing process is increased most with this operating point. The scope of application of the hydraulic oscillating device could be extended in the current research. An adaptation of the oscillation device to an industrial process can thus be conducted. A deeper analysis of the influence of the superimposed oscillation on the quality of the components produced in the ironing process as well as on the occurring tool wear will be executed in following investigations.

Acknowledgments. Funded by the Deutsche Forschungsgemeinschaft (DFG; German Research Foundation) – 417860324 in the project T05 and 116817829 in the project A7 within the scope of the transregional collaborative research centre 73 “Sheet Bulk Metal Forming” (SFB/TCRC 73).

References

1. Merklein, M., Behrens, B.A., Brosius, A., Hagenah, H., Tekkaya, A.E., Weckenmann, A.: Bulk forming of sheet metal. Ann. CIRP **61**, 725–745 (2012)
2. Behrens, B.A., Bouguecha, A., Vucetic, M., Hübner, S., Rosenbusch, D., Koch, S.: Numerical and experimental investigations of multistage sheet-bulk metal forming process with compound press tools. Key Eng. Mater. **651–653**, 1153–1158 (2015)
3. Behrens, B.A., Hübner, S., Vucetic, M.: Influence of superimposed oscillation on sheet-bulk metal forming. Key Eng. Mater. **554–557**, 1484–1489 (2013)

4. Behrens, B.A., Tillmann, W., Biermann, D., Hübner, S., Stangier, D., Freiburg, D., Meijer, A., Koch, S., Rosenbusch, D., Müller, P.: Influence of tailored surfaces and superimposed-oscillation on sheet-bulk metal forming operations. J. Manuf. Mater. Process. **4**, 41 (2020)
5. Blaha, F., Langenecker, B.: Plastizitätsuntersuchungen von Metallkristallen in Ultraschallfeld. Acta Metall. **7**(2), 93–100 (1959)
6. Hu, J., Shimizu, T., Yang, M.: Investigation on ultrasonic volume effects: Stress superposition, acoustic softening and dynamic impact. Ultrason. Sonochem. **48**, 240–248 (2018)
7. Ulmer, J.: Beitrag zur Berechnung der Reibungskraftreduktion beim ultraschallüberlagerten Streifenziehversuch. In: DGM Informationsgesllschaft (Hrsg.): Beiträge zur Umformtechnik, Bd. 37. Universität Stuttgart, Institut für Umformtechnik - Dissertation (2003)
8. Maeno, T., Mori, K., Hori, A.: Application of load pulsation using servo press to plate forging of stainless steel parts. J. Mater. Process. Technol. **214**, 1379–1387 (2014)
9. Matsumoto, R., Utsunomiya, H.: Punch wear in the forming of deep holes with pulse ram motion on a servo press. Key Eng. Mater. **611–612**, 127–133 (2014)
10. Kirchner, H.O.K., Kromp, W.K., Prinz, F.B., Trimmel, P.: Plastic deformation under simultaneous cyclic and unidirectional loading at low and ultrasonic frequencies. Mater. Sci. Eng. **68**(2), 197–206 (1985)
11. Osakada, K., Wang, X., Hanami, S.: Precision Forging of Spline by Flashless Die Forging with Axially Driven Die. Ann. CIRP **46**(1), 209–212 (1997)
12. Michalski, M., Leicht, U., Engel, S., Merklein, M.: Tool system for ultrasonic-assisted forming and material characterisation with 15 kHz oscillation frequency. Appl. Mech. Mater. **794**, 427–434 (2015)
13. Müller, P., Rosenbusch, D., Wehmeyer, J., Hübner, S., Behrens, B. A.: Investigations of forming force, friction values and surface qualities in ring compression tests using oscillating tools. In: Proceedings of the 9th Congress of the German Academic Association for Production Technology (WGP), pp. 73–81 (2019)
14. Jimma, T., Kasuga, Y., Iwaki, N., Miyazawa, O., Mori, E., Ito, K., Hatano, H.: An application of ultrasonic vibration to the deep drawing process. J. Mater. Process. Technol. **80–81**, 406–412 (1998)
15. Resnick, R., Halliday, D.: Physics, Section 1–3 (Vol I and II, Combined edition), Wiley International Edition, Library of Congress Catalog Card No. 66-11527 (1966)

Investigation of Clinched Joints – A Finite Element Simulation of a Non-destructive Approach

B. Sadeghian(✉), C. Guilleaume, R. Lafarge, and A. Brosius

Chair of Forming and Machining Processes, Technische Universität Dresden, 01062 Dresden, Germany
behdad.sadeghian@tu-dresden.de

Abstract. Clinching is a cost-effective mechanical joining process used for metallic as well as nonmetallic materials. Destructive testing methods such as peel and tensile tests can be used for strength investigations of such joints. Additionally, by measuring the geometrical properties such as undercut, neck thickness, and final bottom thickness, the joint quality can be estimated. These methods are cumbersome and do not meet the time and cost efficiency requirements of industrial production. The harmonic analysis benefits from changes in vibration characteristics. In this paper, the structural response of a clinched joint, subjected to a dynamic displacement in order to introduce an acoustic wave testing method is studied. This is accomplished by using simulated joining processes. The approach of this paper is based on a transient dynamic finite element analysis, followed by a fast Fourier transform. The results are presented, together with an analysis of sensitivity to different process parameters.

Keywords: Mechanical joints · Lightweight design · Wave energy dissipation · Harmonic response simulation · Non-destructive testing

1 Introduction

The idea of clinching was originated in Germany in the 18th century. It is a bulk sheet metal forming process. The parts are joined using a punch and a die set and bonded by a unique form. Automotive manufacturers are joining different materials such as advanced high strength steel, aluminum sheet, cast, and extruded profiles or fiber-reinforced plastics [2]. The goal is to minimize the weight of automobiles, reduce overall fleet consumption, and thereby achieve cost reductions. Clinching is one of the possible solutions to increasing demand for the integration of various materials in lightweight parts. The traditional methodology for assessing the clinched joint uses metallurgical investigations to measure the neck thickness, undercut, and bottom thickness [3] (see Fig. 1). Other testing methods, such as peel and tensile tests, are borrowed from spot-welding. These are used for strength investigations of such joints in quasi-static loads. In addition to these methods, fatigue tests are performed to study

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 116–124, 2021.
https://doi.org/10.1007/978-3-662-62138-7_12

the break criteria in dynamic loads [4]. The predefined force-displacement curve for process monitoring is the latest and most effective method to ensure a safe connection [5]. However, this method does not make sure if the geometrical characteristics of the joint in cross-section are appropriately formed.

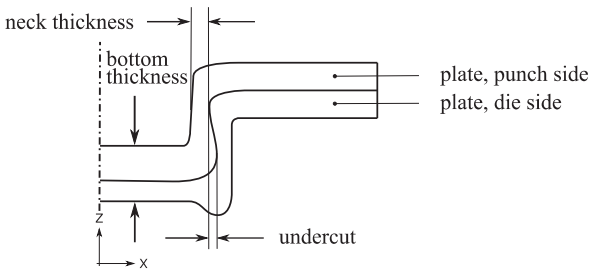


Fig. 1. Clinch joint geometry and common characteristics in the traditional assessment methodology.

To take advantage of the potential of clinching in mass production, a suitable and cost-efficient examination method must be developed. Non-destructive testing (NDT) is a technique to evaluate, measure, or detect discontinuities in raw materials or finished parts without causing any damage. The objective of this paper is to investigate the sensitivity of an NDT method to changes in input parameters of the clinching process.

2 State of the Art

2.1 Non-destructive Testing Methods

According to [6], the NDT methods can be classified based on their physical fundamentals into the following: Penetrant testing (PT), Magnetic particle testing (MT), Eddy current testing (ET), Radiographic testing (RT), Ultrasonic testing (UT), Acoustic emission testing (AE), Visual testing (VT), and Thermal infrared testing (TIR).

PT is based on capillary action and is used for the identification of surface discontinuities [7]. It is an inexpensive method. VT uses electromagnetic radiation at visible frequencies [8]. TIR takes advantage of variations in heat propagation and weakening in solid material [9]. PT, VT, and TIR are limited to the accessible surface of the test object. MT uses ferromagnetic particles to visualize the magnetic leakage field. It is easy to use. However, it applies only to ferromagnetic test objects [10]. The material must be electrically conductive to use ET [8]. It is convenient to characterize volumetric discontinuity using RT. Despite the safety hazards, the accessibility to both sides of the test object is needed [11]. AE uses acoustically generated waves to detect

rapidly growing cracks. However, it is noise sensitive. Therefore, the background noise should be known [12]. UT takes advantage of elastic wave propagation in solid materials. Almost all types of discontinuities can be detected. Among these methods, UT is swift and can be applied to a combination of metallic and nonmetallic materials [6]. These properties make it attractive for testing the clinched joints in body structures connected with different materials. However, the orientation of discontinuities and material grain structure can affect the detectability in UT. Furthermore, a complex shape can reduce the effectiveness of testing. This work investigates the suitability of UT method for clinched joints.

2.2 Ultrasonic Testing (UT)

When mechanical characteristics of a structure change, its dynamic behavior alters. Vibration-based damage detection and wave propagation-based damage detection utilize these changes to detect and localize the damage. Vibration-based damage detection uses long wavelengths to investigate the changes in modal frequencies, mode shapes, and modal damping. Low frequency modes capture the global response of the structure and are less sensitive to local alterations [13]. Wave propagation-based damage detection is more sensitive to local changes. However, measurement and signal processing require substantial effort [14]. Van Den Abeele et al. [15] proved the sensitivity of non-linear methods by focusing on the interaction of low and high-frequency signals. Wolf et al. [16] measured the impact strain of the propagation of an elastic strain wave through a bolted joint. The loss of energy is associated with the different torque moments applied to the bolt. The connection quality is estimated by analyzing the resulting energy loss in an insufficiently performed joint. Lafarge et al. [1] investigate the impact response of a clinched joint using FE method. By varying the process parameters, the stress wave propagation has been studied. This work aims to assess the frequency response of a clinched joint using FE-method. The FE-models from [1] is used to simulate the time-domain response of the joint to a time-harmonic displacement. The simulation makes use of a transient FE calculation, followed by an analysis of the structural response using fast Fourier transform (FFT).

3 Modeling

In this paper, models from Lafarge et al. [1] are used (see Fig. 2). Each simulation has three major steps: forming simulation, springback simulation, and transient dynamic simulation.

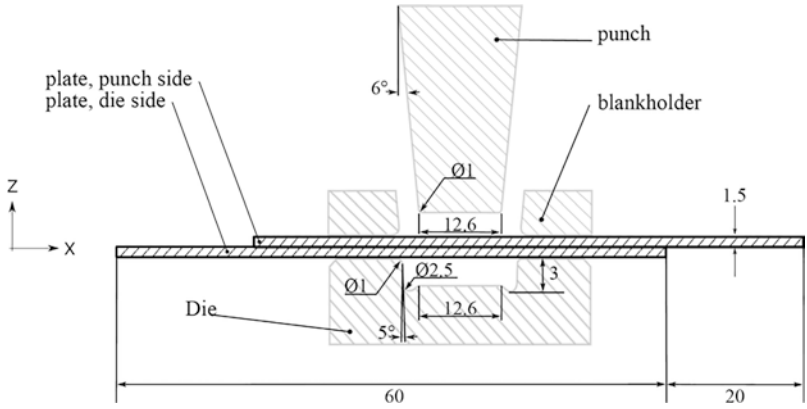


Fig. 2. Dimensions of the forming simulation model used in [1].

3.1 Forming and Springback Simulation

The dimensions of the tool and the workpiece are shown in Fig. 2. The forming simulation is performed in three dimensions using hexahedral elements for plates. Shell elements are used for tools. Hourglassing issues are reduced by using a fully integrated element formulation. A single symmetry plane is used to reduce the simulation time. An elastic-plastic material model with isotropic strain hardening is used. The strain rate dependency is not considered. An associated flow rule describes the changes in the plastic strain. The penalty method is used to define the contact. The flow cuve is in accordance with DC04(1.0338) steel. The program LS-DYNA is used in MPP execution with 16 Intel Haswell processors. Each simulation ran for about seven hours.

The springback simulation is initiated with the output of the forming simulation. Both simulations benefit from the implicit scheme to solve the governing equilibrium equation. Out of the five simulations described in [1], the first simulation is assumed to be an ideal joint. The subsequent simulations are performed with altered process parameters, as shown in Table 1.

Table 1. The altered process parameters in each simulation.

Simulation	Altered Parameters
1	Ideal joint
2	Coefficient of friction at the interface is increased from 0.1 to 0.3
3	Punch stroke is decreased from 5.1 to 4.1 mm
4	Punch is displaced 0.2 mm in the x-direction
5	Punch is turned 1° around the y-axis

3.2 Transient Dynamic Simulation

The harmonic response of the structural components subjected to a harmonic vibration is usually calculated in a steady-state. Due to contact between the surfaces and the complex geometry of the joint, the problem is non-linear. Therefore, to understand the behavior of the joint, it is beneficial to employ a transient analysis and then perform an FFT. In this work, transient analysis using a time-dependent harmonic excitation is performed, followed by an FFT. The model is initialized by using the deformations and the history variables, such as stress and plastic strain, of the spring-back simulation. The body structure is stationary at the beginning of the simulation; i.e., its field variables, such as velocity, acceleration, and displacement, are zero. The system is free to move in the x-direction since it is planned to perform the experimental analysis in this way.

In structural analysis with FE-methods, many factors can influence the results. The mesh size, the time step, the material model, and the time integration method can have a notable impact on the accuracy of the calculation. The equation of equilibrium governing the dynamic response is solved using the central difference method. It is very efficient in the individual step. However, it is disadvantageous that the information at the time, t_{n+1} is extracted from the data at the time, t_n. Therefore, the equilibrium condition may not precisely be met at t_{n+1}. A minimal time step is required to overcome this problem. For stability reasons, the time step cannot be greater than a critical time step [17].

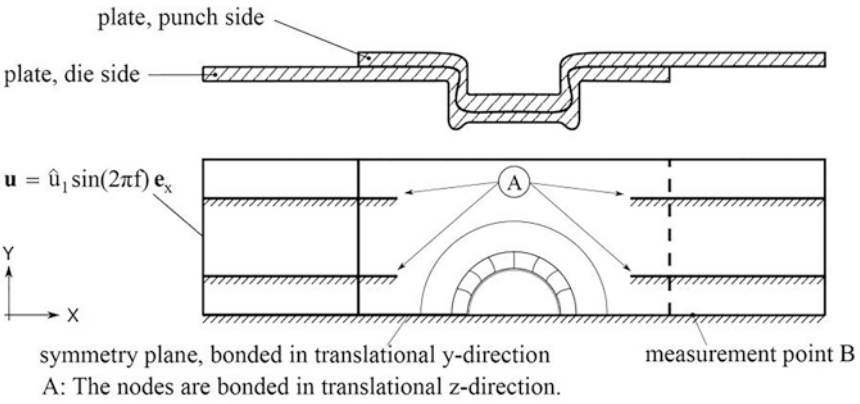


Fig. 3. The model of a clinch joint with a symmetry plane. The left plate has been exited with a time-harmonic displacement. The response is measured on point B. The translational freedom of the nodes on both surfaces of the plates shown with the letter "A" are bonded in the translational z-direction. The translational freedom in the y-direction in the symmetry plane is bonded. The model is free in translational x-direction.

The stability condition will be satisfied on the conservative side if the minimum element time step is used (see Eq. 1).

$$\Delta t = min\frac{L_e}{c_e} \tag{1}$$

where c_e is the wave propagation speed of a Longitudinal wave, and L_e is the characteristic length of the element. The wave propagation speed is calculated for a three-dimensional case using Eq. 2 [17]:

$$c_e = \sqrt{\frac{E(1-\nu)}{\rho_e(1+\nu)(1-2\nu)}} \tag{2}$$

The changes in process parameters alter the geometrical characteristics of the joint. For analytical reasons, it is desirable to have a clear distinction between the effects of the changes in the geometrical characteristics of the joint on the longitudinal and transversal deformations. Hence, the Poission ratio is set to zero. Hence, for steel with E = 207000 N/mm^2 and ρ = 7830 kg/m^3, the wave propagation speed is equal to 5141.7 m/s^2. The smallest characteristic length is equal to 0.03 mm. Thus for stability reasons, the time step size should be smaller than 5.83 ns. With a total calculation time of 500 µs, the smallest frequency expected in the frequency domain is 2 kHz.

Due to the joining process, at the beginning of the transient simulation, both plates in the contact interface are pre-stressed. The contact model disables the nodal interpenetration check, inorder to allow the contact forces to remove the initial interpenetrations. The contact model is penalty based, since the program has the above option just for penalty based contacts.

The left plate in Fig. 3 has been excited by a time-harmonic displacement, $\boldsymbol{u} = \hat{u}_1 \sin(2\pi f)\boldsymbol{e}_x$, with a frequency, f = 120 kHz, and an amplitude, $\hat{u}_1$ = 0.05 mm. The displacement has been measured at point B in Fig. 3 in the x-direction. A high excitation frequency of 120 kHz has been chosen to avoid deflections. We used the program LS-DYNA to solve the equation of equilibrium using MPP execution with 16 Intel Haswell processors family. Each simulation ran 72 min on average.

4 Results and Discussion

The objective of the simulations is to prove the sensitivity of the wave propagation-based testing method to the changes in process parameters. If changes in process parameters cause changes in the characteristics (amplitude and phase) of the transmitted wave, this method could be used to evaluate clinch joints.

The time-domain response is shown in Fig. 4. The results of the simulation correspond to decreasing the punch stroke shows higher amplitudes than others. The single-sided amplitude spectrum shown in Fig. 5 demonstrates that the transmitted wave consists of different frequencies, which include the base signal (120 kHz) and noise. Table 2 shows the amplitude and phase information of the base signal obtained from FFT. The transmitted wave characteristics of the different simulations

are compared with the first simulation in Table 1, which represents an ideal joint. A decrease in the punch stroke caused an increase in the amplitude and a significant negative phase shift of the base signal. Turning the punch around the y-axis resulted in an increase in the amplitude. The displacement of the punch along the positive x-axis resulted in a decrease in the amplitude of the base signal.

Increasing the friction coefficient resulted in a slight decrease in the base signal amplitude. It appears that friction has most likely less impact on the quality of the joint. This is similar to the results obtained in [1]. However, it differs from the bolted joint investigated in [16], where the friction force transmitted most of the energy.

To summarize, changing the geometrical process parameters alters the geometrical characteristics of the joint, which leads to a quality deficiency. This is reflected in the transmitted wave characteristics shown in Table 2.

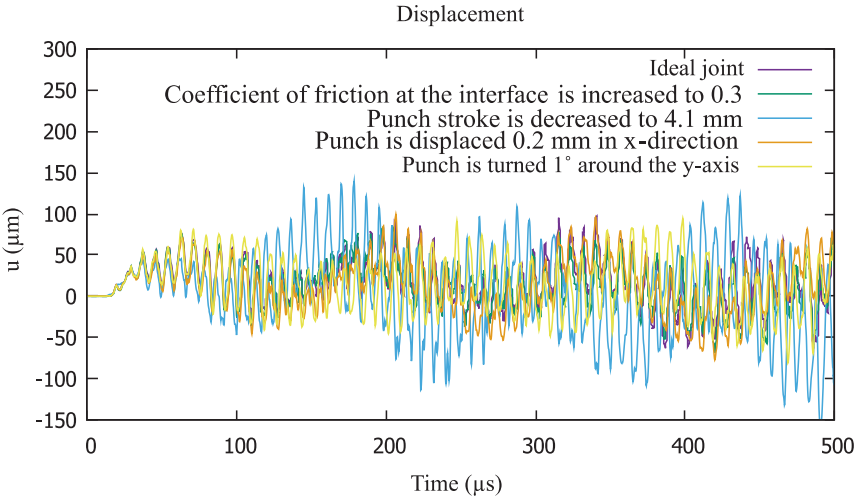


Fig. 4. The simulation results in the time domain at point B.

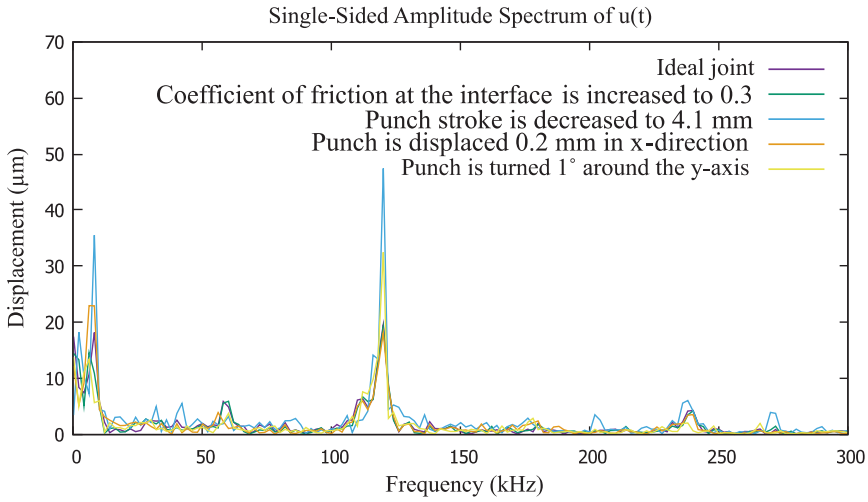


Fig. 5. The single-sided amplitude spectrum.

Table 2. Characteristics of the base signal

Simulation	Base frequency		
	Frequency (kHz)	Amplitude (µm)	Phase/π
1	120	20.4	0.9
2		19.8	1.0
3		48.3	−0.4
4		18.9	0.8
5		37.0	−0.9

5 Conclusions

The ultrasonic testing method shows great promise in smart manufacturing applications of clinching. It is also possible to use this method for the maintenance of such joints. It is shown through FE simulations that the changes in geometrical process parameters, such as punch stroke and displaced or eccentric punch result in amplitude and phase changes of the transmitted base signal. This could be used for the identification of weak joints. Changes in the coefficient of friction most likely do not have a significant impact on the quality of the clinched joint. Further studies will be conducted on the impact of these variables using a combination of different excitation frequencies.

Acknowledgments. This research has been Founded by the Deutsche Forschungsgemeinschaft (DFG, German research foundation) – Project-ID 418701707-TRR 285, subproject C04. We thank the Center for Information Services and High Performance computing (ZIH) at TU Dresden for generous allocations of computer time.

References

1. Lafarge, R., Wolf, A., Guilleaume, C., Brosius, A.: A new non-destructive testing method applied to clinching submitted to ICTP (2020)
2. Meschut, G., Janzen, V., Olfermann, T.: Innovative and highly productive joining technologies for multi-material lightweight car body structures. J. Mater. Eng. Perform. **23**(5), 1515–1523 (2014). https://doi.org/10.1007/s11665-014-0962-3
3. DVS, Merkblatt DVS-EFB 3420, Clinchen – Überblick (2012)
4. Doege, E., Behrens, B.-A.: Handbuch Umformtechnik, 3rd edn. Springer, Heidelberg (2016)
5. Khrebtov, P.: Neuartiges Verfahren zur Online-Prozessüberwachung und -Fehlerklassifizierung beim Durchsetzfügeverbinden von Blechen. Technische Universität Clausthal (2011)
6. Brook, M.V.: Ultrasonic Inspection Technology Development and Search Units Design. Examples of Practical Applications. IEEE Press, Oxford: Wiley-Blackwell, Piscataway (2012)
7. NDT Resouece Center. https://www.nde-ed.org/EducationResources/CommunityCollege/PenetrantTest/cc_pt_index.php. Accessed 15 May 2020
8. Hellier, C.J.: Handbook of Nondestructive Evaluation. McGraw-Hill Companies Inc., USA (2003)
9. DIN 19277, Zustandsüberwachung und -diagnostik von Maschinen - Thermografie – Allgemeine Methoden (2015)
10. DIN EN ISO 9934-1 Zerstörungsfreie Prüfung - Magnetpulverprüfung – Teil 1: Allgemeine Grundlagen (ISO 9934-1:2016); Deutsche Fassung EN ISO 9934-1:2016 (2016)
11. Diederichs R.: Nondestructive Testing (NDT). https://www.ndt.net/ndtaz/ndtaz.php. Accessed 15 May 2020
12. Vahaviolos, S.J.: Acoustic Emission: Standards and Technology Updates. ASTM, USA (1999)
13. Doebling, S.W., Farrar, C.R., Prime, M.B.: A summary review of vibration-based damage identification methods. Shock Vib. Digest **30**(2), 91–105 (1998)
14. Kögl, M., Hurlebaus, S., Gaul, L.: Finite element simulation of non-destructive damage detection with higher harmonics. NDT E Int. (2004). https://doi.org/10.1016/j.ndteint.2003.09.003
15. Van Den Abeele, K.E.-A., Johnson, P.A., Sutin, A.: Nonlinear Elastic Wave Spectroscopy (NEWS) techniques to discern material damage, part I: Nonlinear Wave Modulation Spectroscopy (NWMS). Res. Nondestr. Eval. (2000). https://doi.org/10.1080/09349840009409646
16. Wolf, A., Lafarge, R., Brosius, A.: A non-destructive testing method for joints by the measurement of the energy dissipation. Prod. Eng. Res. Dev. **13**(1), 99–106 (2018)
17. Wagner, M.: Lineare und nichtlineare FEM. Springer, Wiesbaden (2019)

Experimental Process Design for Reclamation of Geared Components

P. Kuwert(✉), T. Petersen, K. Brunotte, and B.-A. Behrens

Institute of Forming Technology and Machines, Leibniz Universität Hannover,
An der Universität 2, 30823 Garbsen, Germany
kuwert@ifum.uni-hannover.de

Abstract. From a metal forming perspective, warm forming technologies can be used to reclaim the shape and functionality of worn engineering steel parts such as case-hardened geared components. Since forming temperatures are beneath the material-transformation temperature, no loss in hardness is expected and a new heat treatment can be omitted. In order to determine a process window for reclaiming geared components, upsetting tests were carried out on case-hardened workpieces made of AISI 5115.

The workpieces were heated to between 700 °C and 900 °C by induction or furnace heating and upset to a true strain between 0.1 and 0.4. The upset workpieces were cooled in two media, air and water, and metallographically investigated. The combination of 900 °C furnace heating, upsetting ($\varphi_h = 0.4$) and quenching led to an increase of 60% in case-hardening depth. This procedure can be used for reclaiming geared components by means of forming technology.

Keywords: Bulk metal forming · Product recycling · CO_2-reduction

1 Introduction

Iron, mostly in the form of steel, is by far the most frequently used metal due to its mechanical and chemical properties [1]. In combination with the resulting high demand and energy-intensive production, steel is responsible for a high amount of CO_2 emissions. In order to reduce CO_2 emissions, the amount of crude steel produced from remelted steel scrap by material recycling has risen to 55.5% in 2017 in Europe [2]. One way to reduce the environmental impact even further is to recycle steel components by product recycling. Due to this approach, energy-intensive processes in the manufacturing of steel and steel components can be omitted. Product recycling is particularly suitable for recycling mass products such as worn-out gears.

In general, the failure of geared components results from the high mechanical and tribological loads leading to the predominant failure modes of wear, micropitting and pitting [3, 4]. Depending on hardness, lubricating film, height and surface roughness, the surface of the flank changes significantly from first tooth contact [5]. Due to

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 125–133, 2021.
https://doi.org/10.1007/978-3-662-62138-7_13

additional wear, dimensional deviation increases during the component's service life. Consequently, the load distribution on the tooth flank changes and leads to a local concentration of stress and thus different predominant surface fatigue failures (pitting, micropitting and wear) occur [6, 7]. Through continuously increasing load, the loss of material in the pitting and worn areas increases from the tooth root. As a result, the component shape changes negatively by a stochastic amount and finally leads to the failure of the geared component [8, 9].

From a metal forming perspective, bulk metal forming technologies at elevated temperatures can be used to reclaim geared components as illustrated in Fig. 1. The worn tooth areas can be refilled by the forming of grooves in a closed die. Behrens et al. investigated conceptually reclaiming geared components by precision forging [10]. By means of heating up gears to 900 °C and the movement of an inner stamp ring, material could be pressed into the worn out areas. Resulting from this forming strategy, the shape and function of the geared component could be restored [10].

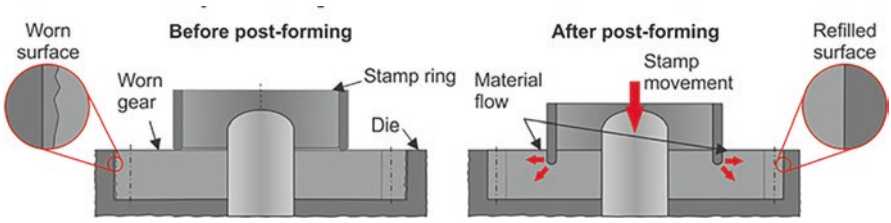


Fig. 1. Forming strategy for the re-manufacturing of spur gears [10]

Depending on the extent of damage, an increase in forming temperature is necessary to ensure higher formability for better die filling. However, if forming temperatures exceed the material's transformation temperature (A_{c1}), microstructural transformation and the loss of the required hardness are inevitable. In this case, the geared components have to be heat-treated and subsequently machined anew, which offsets the economic advantage of this recycling concept. To avoid these additional efforts, it is necessary to identify a suitable process control which can handle the heating method, forming temperature and cooling method. By this means, the case hardness, case hardening depth (CHD) and microstructure in the case and core can be maintained.

This paper addresses the experimental determination of a proper process control for reclaiming geared components based on upsetting experiments.

2 Materials and Method

The method of reclaiming geared components by metal forming is to be demonstrated using an FZG-C-PT spur gear, which is usually made of AISI 5115 and case hardened to CHD: 52,3 HRC 0.6–0.9 mm [11]. In order to avoid an additional heat

treatment after reclaiming, forming temperatures must not exceed the material's transformation temperature (A_{c1} temperature). Forming below A_{c1} temperature is a limiting factor in terms of the limited formability of case hardened material which might result in surface damage during forming. Therefore, the workability and hardness evolution of case hardened AISI 5115 was characterised by basic cylinder upsetting tests. Induction heating and furnace heating were chosen as heating strategies to represent different heating rates; the same applies to chosen cooling methods (water/air), as both influence the resulting microstructure. Before the experiments, the cylindrical workpieces (dimensions: Ø30 × 40 mm) were case hardened at a service company in order to reproduce the same case hardness and case hardening depth as the FZG-C-PT spur gears (CHD: 52,3 HRC 0.6–0.9 mm) [11]. The samples were initially carburised at 830 °C in a closed container with additives against scale formation, cooled, anew reheated in a closed container with additives to 830 °C, quenched and then tempered at 180 °C for 2 h. Subsequently, the state of microstructure and hardness of the case-hardened workpieces were analysed in case and core. Figure 2 shows the classic microstructure in case and core after case hardening. Taking into account the measured hardness values and [12], the microstructure of the case can be described as a composition of retained austenite and coarse-needled martensite. The core microstructure shows a typical dual-phase grain structure of ferrite and martensite. After quenching, the austenite is transformed to martensite and occurs together with ferrite. Compared to the case, the carbon content here should only be 0.16% (according to [12]) resulting in an austenite-ferrite phase region, which is also confirmed by hardness measurements. In summary, the hardness properties and microstructure comply with the test conditions, although the CHD is slightly lower in comparison to the FZG-C-PT spur gears (see Table 1).

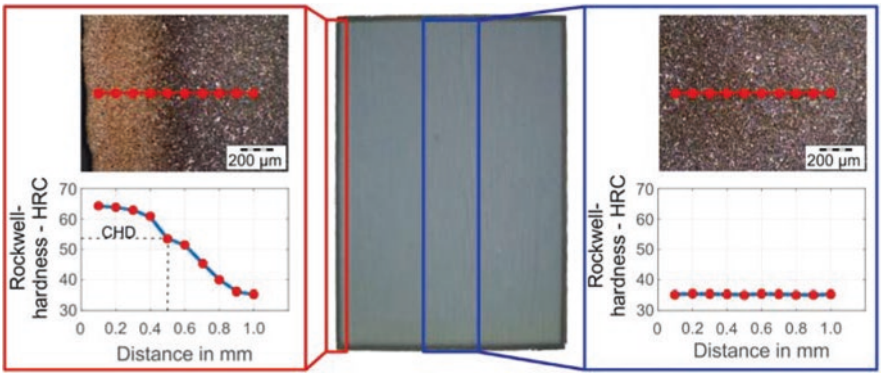


Fig. 2. Hardness and microstructure of case-hardened AISI 5115 samples

Table 1. Hardness profile of case-hardened AISI 5115 workpieces and FZG-C-PT gear [11].

	Case hardness in HRC	CHD in mm
Workpieces	63.8 ± 0.5	0.5 ± 0.05
FZG-C-PT gear	60–62	0.6–0.9

Before upsetting test, induction heating experiments were carried out in order to determine proper time-power combinations to obtain homogenous temperature distributions in the hardened workpieces. The target temperatures were chosen below A_{c1}, namely 700 °C, and based on preliminary investigations above A_{c1}, namely 800 °C and 900 °C [10]. The higher temperatures can lead to an unfavorable microstructure and hardness, which can however be compensated by appropriate cooling [12]. For inductive heating, a mid-frequency generator (Hüttinger TruHeat MF 3040) was used with an output voltage of about 300 V and maximum power output of 40 kW. Thermocouples were positioned in the core and near the case of the workpiece to monitor the evolution of workpiece temperature during and after induction heating. As Fig. 3 shows, the targeted forming temperature of 700 °C was achieved without exceeding A_{c1}. The higher temperatures (800 °C and 900 °C) were also adjusted, although overheating was necessary to achieve a homogenous temperature distribution in the same heating time.

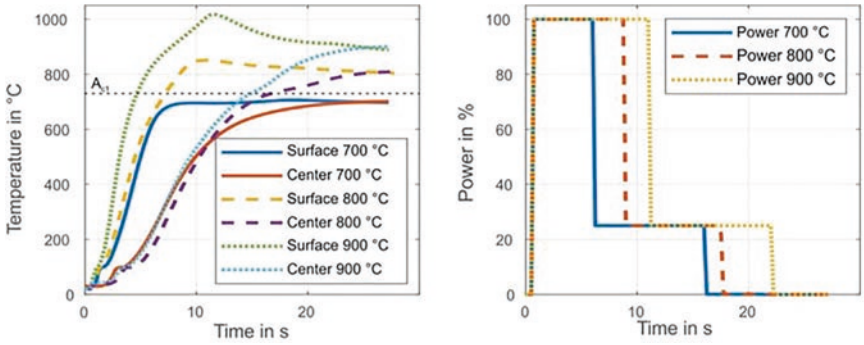


Fig. 3. Time-power and time-temperature curve for inductive heating

After determining proper parameters for induction heating, upsetting experiments were carried out with the parameters shown in Table 2, using a full-factorial approach. Regarding reclaiming geared components, it was proven in preliminary theoretical work that a maximum true strain of 0.4 is necessary to refill the worn out gaps of the analysed FZG-C-PT spur gear. In order to investigate a possible damage evolution, the experiments start from a true strain of 0.1 and were increased up to 0.4.

Table 2. Plan of upsetting experiments.

Heating	Forming temperatures in °C	True strain	Cooling
Furnace, inductive	700, 800, 900	0.1, 0.2, 0.4	Air, water

For this, a fully automated forging cell consisting of a screw press with a nominal capacity of 5,000 kN, an industrial robot and an induction-heating rig were used.

The hardened workpieces were heated to their forming temperature by induction heating or by using a chamber furnace (20 min in a closed container with additives

against scale formation). Then they were transferred to the forming tool by the robot, upset and cooled in air or water. Afterwards, the upset parts were analysed optically and metallographically (microstructure and hardness measurements).

3 Results and Discussion

As shown in Fig. 4, it was possible to upset the case-hardened workpieces without any damage at all forming temperatures. Accordingly, only samples upset to a true strain of 0.4 will be analysed in the metallographic investigations.

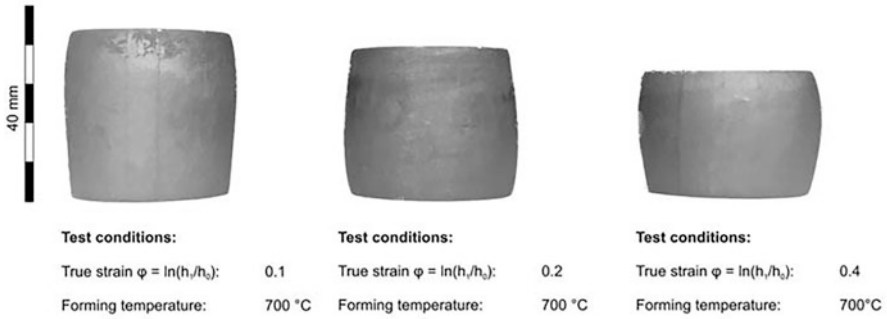


Fig. 4. Overview of upset samples

Light-microscope scans of the air-cooled samples showed a complete ferritic-pearlitic microstructure transformation at all forming temperatures, with an average hardness of approximately 40 HRC. Due to this finding, air-cooling after post-forming would generally require a new case hardening of the formed component. In contrast, quenching maintains or increases the hardness profile for both heating methods. With regard to the respective hardness profiles, furnace and induction heating show characteristic differences in case hardness and CHD (see Fig. 5).

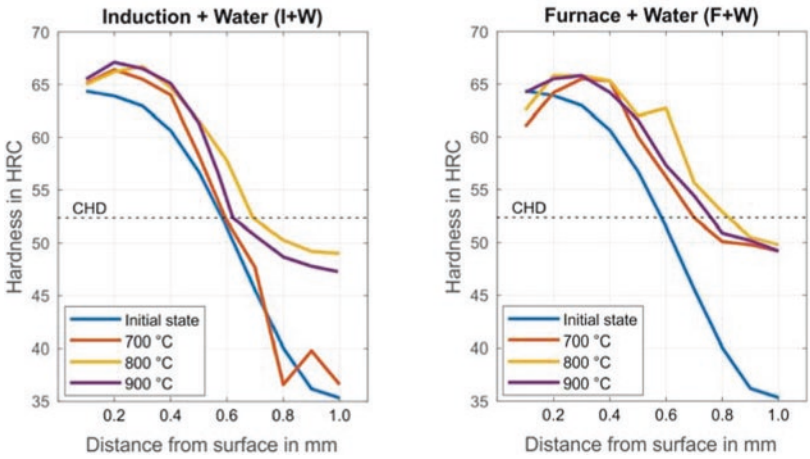


Fig. 5. Hardness evolution after upsetting and cooling of AISI 5115 samples

It can be seen, that the case hardness of upset parts, which were induction-heated and water-cooled, is characterised by a general increase for all forming temperatures (see Fig. 5, left). This increase can be described as the result of a transformation of the retained austenite into martensite due to heating followed by forming and quenching. Similar effects of phase transformation are well-known in the thermomechanical treatment of this material [13]. This would explain the difference in the hardness profiles when the forming temperature 700 °C is compared with 800 °C and 900 °C, since the latter result in enhanced hardness profiles at a depth of 0.6 mm. In contrast, heating in a chamber furnace leads to a slight decrease in case hardness (directly on the surface), except for 900 °C forming temperature. Basically the same explanation approach, i.e. the transformation of retained austenite into martensite by renewed thermomechanical processing (heating, forming, quenching), can be applied here. However, due to the longer heating times, diffusion processes are more pronounced. In an effort to achieve equilibrium, a carbon diffusion from the surface to the core takes place during furnace heating, which would explain the hump-shaped hardness profiles. Another explanation approach could be a brief soft annealing, where hard cementite in the boundary areas decomposes leading to reduced hardness. However, this would not explain the humpy shape of the hardness profiles. Based on these findings, induction heating with quenching and furnace heating with quenching at 900 °C are appropriate methods for improving the hardness profile, as case hardness as well as CHD increase (see Table 3).

Table 3. Hardness of upset AISI 5115 samples.

	Initial state	Furnace + Water (F + W)			Induction + Water (I + W)		
Temp. in °C	–	700	800	900	700	800	900
Case hardness in HRC	63.8	62.0	62.3	64.2	65.0	65.1	65.5
CHD in mm	0.5	0.71	0.74	0.81	0.61	0.63	0.68

Figure 6 shows the corresponding microstructure in the case (left) and core (right) of these upset samples. In the case, coarse-needled martensite is present together with retained austenite, which results in a case hardness of 64.5 ± 0.3 HRC (F+W) and 65.2 ± 0.4 HRC (I+W). At the same time, an intermediate stage microstructure is present in the core. Even after case hardening, carbon content in the core is still 0.16%, where the iron-carbon diagram indicates a pure ferritic-perlitic region. To this end, a fine-grained ferritic-martensitic structure with a hardness of 35 HRC ± 1.2 HRC is present in both samples, which is almost identical to the initial state (see Fig. 2). Therefore, a ductile core was adjusted with the described process sequences, which is an important criterion for geared components [14].

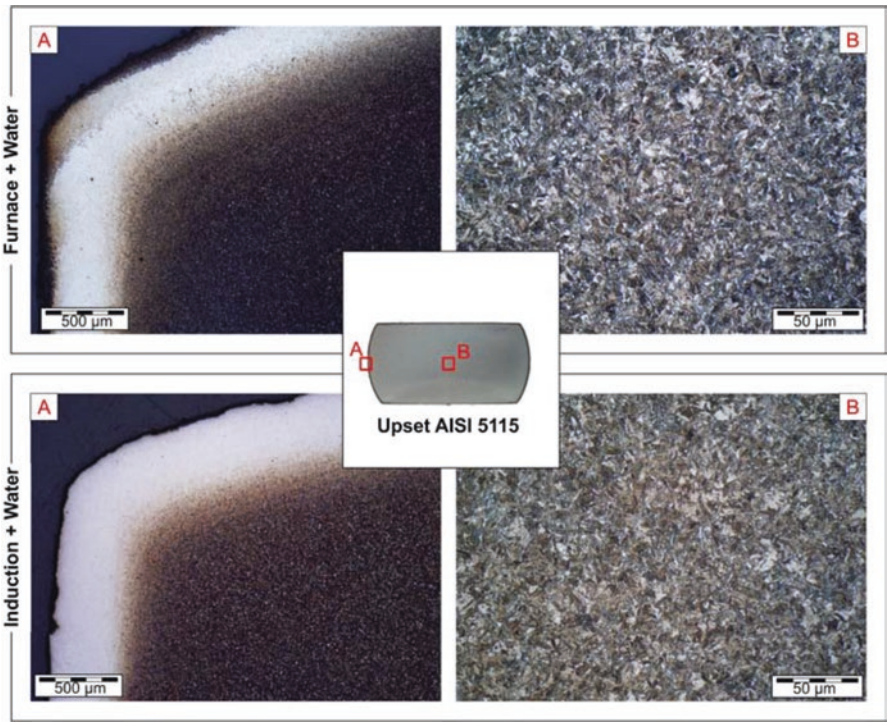


Fig. 6. Microstructure of 900 °C heated, upset and quenched samples ($\varphi = 0.4$)

In summary, the investigations proved that microstructure and hardness profile in case and core can be maintained and even improved by induction or furnace heating with a subsequent cooling in water after forming. With regard to reclaiming geared components, an increased CHD benefits service life [14, 15]. As a result, post-forming of damaged geared components should be performed by heating in a furnace with subsequent cooling in water. Based on the findings, this procedure of product recycling would even improve the properties of the geared component.

4 Summary and Outlook

Within the scope of the investigations, suitable process sequences as well as a process control for the remanufacturing of geared components by forming technology were investigated. For this purpose, workpieces made of AISI 5115, with an identical hardness profile and microstructure as FZG-C-PT spur gears, were upset by varying heating strategy, forming temperature, true strain and cooling method. The metallographic investigations showed an increase in CHD of the heated, formed and water-quenched samples, which can contribute to the service life of gears. Furthermore, with 60% furnace heating shows the greatest increase in CHD resulting from the longer heating times and diffusion of carbon from the case into the core. These effects can be transferred to the recycling of spur gears by using warm forming technologies and will be analysed in future forging experiments to reclaim FZG-C-PT spur gears.

Acknowledgements. The results presented were obtained in the research project "Determination of the process limits for the re-manufacturing of damaged gears by using precision post-forming at elevated temperatures" which is financed by the German Research Foundation (DFG) under the project number 319564414. The authors would like to thank the German Research Foundation.

References

1. Haupt, M., Vadenbo, C., Zeltner, C., Hellweg, S.: Influence of input-scrap quality on the environmental impact of secondary steel production. J. Ind. Ecol. **21**(2), 391–401 (2016)
2. Bureau of Intenational Recycling: Steel scrap – a raw material of steelmaking. World Steel Recycling in Figures 2013–2017, p. 15 (2018)
3. Weibring, M., Gondecki, L., Tenberge, P.: Simulation of fatigue failure on tooth flanks in consideration of pitting initiation and growth. Tribol. Int. **131**, 299–307 (2019)
4. Behrens, B.-A., Odening, D.: Process and tool design for precision forging of geared components. In: The 12th International ESAFORM Conference on Material Forming, 125–128 (2009)
5. Feng, K., Borghesani, P., Smith, W.A., Randall, R.B., Chin, Z.Y., Ren, J., Peng, Z.: Vibration-based updating of wear prediction for spur gears. Wear **426–427**, 1410–1415 (2019)
6. Morales-Espejel, G.-E., Gabelli, A.: A model for gear life with surface and subsurface survival: tribological effects. Wear **404–405**, 133–142 (2018)
7. Santus, C., Beghini, M., Bartilotta, I., Facchini, M.: Surface and subsurface rolling contact fatigue characteristic depths and proposal of stress indexes. Int. J. Fatigue **45**, 71–81 (2012)
8. Morales-Espejel, G.-E., Rycerz, P., Kadiric, A.: Prediction of micropitting damage in gear teeth contacts considering the concurrent effects of surface fatigue and mild wear. Wear **398–399**, 99–115 (2018)
9. Bach, F.-W., Behrens, B.-A., Dähndel, H.: Integration of heat treatment in precision forging of gear wheels. Arab. J. Sci. Eng. **30**, 103–112 (2005)
10. Kuwert, P., Brunotte, K., Behrens, B.-A.: Process Development for the Remanufacturing of Geared Components. Production at the Leading Edge of Technology, 53–61. Springer, Berlin (2019)

11. Radev, T.: Einfluss des Schmierstoffes auf die Grübchentragfähigkeit einsatzgehärteter Zahnräder. Dissertation, Technical University Munich (2013)
12. Chaturvedi, H., Chaturvedi, S., Sharma, S.: Optimisation of heat treatment process for 16MnCr5. Int. J. Eng. Sci. Technol. **4**(3), 998–1004 (2012)
13. Cryderman, R., Whitely, B., Speer, J.: Microstructural evolution in microalloyed steels with high-speed thermomechanical bar and rod rolling. In: Proceedings of International Federation of Heat Treating and Surface Engineering (2016)
14. Argoud, V., Morel, F., Pessard, E., Bellett, D., Thibault, S., Gourdin, S.: Fatigue behavior of gear teeth made of case hardened steel: from competing mechanisms to lifetime variability. Procedia Struct. Integrity **19**, 719–728 (2019)
15. Demet, S.M., Ersoyoglu, A.S.: Experimental study on fatigue fracture damage of symmetric spur gear tooth. Procedia Struct. Integr. **13**, 2036–2039 (2018)

A New Approach for the Evaluation of Component and Joint Loads Based on Load Path Analysis

C. Steinfelder(✉) and A. Brosius

Chair of Forming and Machining Processes, Technische Universität Dresden, 01062 Dresden, Germany
christian.steinfelder@tu-dresden.de

Abstract. The load specific design of structural components and their associated mechanical joints is challenging due to the large number of interdependencies between different joining parameters. The definition of a transferable knowledge base is also difficult and in most cases the design phase has to be accompanied by numerous experimental tests. Therefore, in most cases, the components and/or the clinched joints are manufactured with a significant safety factor, generally leading to oversizing and corresponding reduced economic and environmental efficiency. This paper describes the basics of a new approach for the evaluation of component and joint loads based on load path analysis by using a simple specimen geometry. The new method is independent of the orientation of the component and the force input in relation to the coordinate system used for evaluation. It is presented using simple plate with a square hole and compared to the classical approach. The primary goal of this methodical approach is an even load distribution over the joining point and component, thereby developing a basis for future design approaches aiming at the reduction of oversizing in joint structures.

Keywords: Clinching · Load path analysis · Finite element analysis

1 Introduction

The design of structural components containing mechanical joints still often leads to oversizing for safety reasons. Since the design variables and process parameters have numerous interdependencies, a comprehensive design approach for both the component and the distributed joining points does not yet exist. Instead, experimental testing and extensive numerical studies are used to define the final geometries and process parameters. Then, in order to account for errors and possible material scatter, additional safety margins are added that reduce both the economic viability as well as resource efficiency of the jointed parts.

This paper describes the basic methodology of the novel method for the analysis of a load path which is independent of the alignment of the analyzed geometry by

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 134–141, 2021.
https://doi.org/10.1007/978-3-662-62138-7_14

using a simple plate with a square hole. It is the basis for a new transferable methodical approach for the design of component and joint.

The goal of the new method presented here is to comprehensively analyze the load distribution over a structural component while considering the placement and geometry of the clinched joints. By analyzing the load distribution, it is possible to identify the areas that are particularly heavily loaded and also those where the material's load capacity is not sufficiently utilized. This knowledge can be used to adapt the component and/or the joint in order to achieve an even load distribution, reducing oversizing and also the risk of critical failure. The evaluation whether the joint quality and the full component functionality can still be guaranteed after adapting the component structure or the joints' position requires a fundamental understanding of the complex interactions and sensitivities.

The new approach for evaluating and modifying the component and joint load is based on force flow models for finite element analysis (FEA). In a further step, load cases of the assemblies, components and joints and thus the local transfer of the force flow into the mechanical joints are analyzed. The aim of this procedure is to reduce or homogenize the load on the different load paths under consideration, so that the existing joints are subjected to a load that can be specifically adjusted. The overall goal is to develop a holistic calculation method for the design of the assembly, i.e. the system component - joint, taking into consideration the various interdependencies.

2 State of the Art

Marhadi and Venkataraman presented a good overview of the current methods of loadflow analysis, including an assessment of performance of the different methods [1] Qualitative and quantitative visualizations of load paths in components have an important function in the design process of structures. They allow the assessment of the design and structural integrity and can qualitatively show how the loads are redistributed in case of local damage [2]. They can also be used to create alternative load paths for changes in lifetime [1].

The easiest form of force load flow analysis is based on the direction vectors of the largest principal stress in the integration points from the FEA [2]. This approach provides good results for regions with low shear stress and without multi-axial stress states, because the principal stress vectors align along an expected load path when the proportion of the smallest principal stress is low. This approach is suitable for identifying critical regions and is used, among other things, for designing the fiber orientation in fiber reinforced plastic composites [3]. Steel structures represent another field of application for force flow analyses [4].

In order to overcome the limitations regarding the visualization of shear stress, Kelly develops an alternative calculation approach for load paths [2]. For this purpose, the determined stresses are projected so that the x-component of the force vector is constant in the global x-direction. The angle of the projection is chosen so that the projections of the stresses are eliminated at the element under consideration to avoid an increase in force in the x-direction. Waldmann later developed a closed solution

for this iterative equation for determining the projection angle [5]. The disadvantage of this approach is the dependence of the solution from the alignment of the global coordinate system to the effective load and it is therefore not transferable to complex structures without adjustments. Further publications on basic methods of load path analysis are not known to the authors.

3 New Approach for Load Path Analysis

As shown in the state of the art, the available methods for load path analysis are dependent on the selected coordinate system or the existing stress state. Therefore, they all have deficits regarding a generally valid applicability since the results are not invariant from the chosen coordinate system. For this reason, a new method of analysis is being developed to assess the load and capacity utilization of the component and the joint. A simple test case is used to demonstrate that for this load path analysis method, the results of the stress tensors are invariant from the chosen coordinate system.

3.1 Test Case for Invariant Stress Tensors

The test case consists of a two-dimensional plate with a hole that is subjected to a unidirectional load by displacement of the upper nodes by 1 mm. The resulting stress tensors are evaluated for a load aligned with the both the global and local y-axis (Fig. 1a)) and then compared to a 45° rotated coordinate system with the load aligned to only the local y-axis, see Fig. 1b).

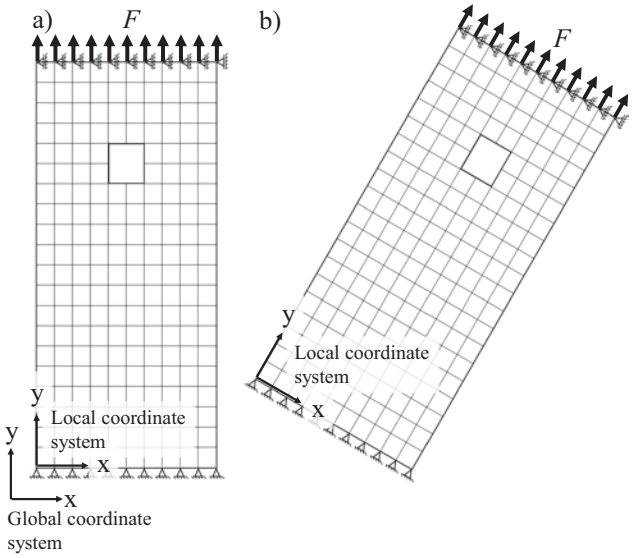


Fig. 1. Meshed plate with hole in 2D with global and local coordinate systems

If the proposed analysis method is correct, the resulting stress tensors for both cases have to be identical, meaning that the results do not depend on the chosen orientation of the analyzed component. A new theoretical approach was tested using a simple finite element analysis (FEA). The simulation was done using LS-DYNA R11.1 with an explicit solver [6]. The 2D-model uses the element formulation ELFORM12 with plane stress in the x-y plane of the model. A fully elastic material behavior using MAT001 with a Young's modulus of 210,000 MPa, a density of 7.85 kg/dm^3 and a Poisson's ratio of 0.3 was assumed [7]. The dimensions of the simulated plate were 20 mm × 40 mm meshed with quad elements of 1 mm × 1 mm. The hole had the dimensions 4 mm × 4 mm.

3.2 Calculation Approach for Load Path Analysis

The hypothesis is based on the assumption that the individual component areas are loaded by a single resulting force that corresponds in its effect to the existing stress components at the respective location. By evaluating an FEA, a stress tensor can be specified for each integration point of a finite element. By calculating the eigenvalues and the eigenvectors of this stress tensor an effective stress vector $\boldsymbol{\sigma}_{\text{eff}}$ can be specified. The eigenvalues λ_i are principal stresses and the eigenvectors $e_{i,j}$ indicate the principal stress direction. Related to an infinitesimal area element dA, this vector can be regarded as an effective force $\boldsymbol{F}_{\text{eff}}$, see Eq. (1).

$$\boldsymbol{F}_{\text{eff}} = \boldsymbol{\sigma}_{\text{eff}} \cdot dA = \left\{ \begin{array}{l} \lambda_{\text{I}} \cdot e_{\text{I,x}} + \lambda_{\text{II}} \cdot e_{\text{II,x}} + \lambda_{\text{III}} \cdot e_{\text{III,x}} \\ \lambda_{\text{I}} \cdot e_{\text{I,y}} + \lambda_{\text{II}} \cdot e_{\text{II,y}} + \lambda_{\text{III}} \cdot e_{\text{III,y}} \\ \lambda_{\text{I}} \cdot e_{\text{I,z}} + \lambda_{\text{II}} \cdot e_{\text{II,z}} + \lambda_{\text{III}} \cdot e_{\text{III,z}} \end{array} \right\} dA \tag{1}$$

The resulting vector fields are evaluated using trajectories. As a criterion for the calculation of the trajectory, the direction of the minimum, maximum or constant amounts of the vector $\boldsymbol{F}_{\text{eff}}$ can be used. The advantage of this proposed method is the independence of the result from the selected coordinate direction, because the local invariant principal stresses and their eigenvectors are included in the calculation.

3.3 Evaluation of the Stress Tensors and Data Processing

For the evaluation of the stress tensors and the data processing *Python 3.7* was used in combination with the libraries *LS-Reader*, *NumPy* [8], *SciPy* [9] and *Matplotlib* [10]. The stress information was read from the d3plot file and saved as a stress tensor $\boldsymbol{\sigma}$, Eq. (2).

$$\boldsymbol{\sigma} = \begin{bmatrix} \sigma_{\text{xx}} & \tau_{\text{xy}} \\ \tau_{\text{xy}} & \sigma_{\text{xy}} \end{bmatrix} \tag{2}$$

In the next step the eigenvalues λ_i and eigenvectors e_i of the stress tensor were calculated, Eqs. (3)–(5)

$$\text{Eigenvals}(\sigma) = \begin{bmatrix} \lambda_{\text{I}} & \lambda_{\text{II}} \end{bmatrix} \tag{3}$$

$$\lambda_{\mathrm{III}} \stackrel{!}{=} 0 \tag{4}$$

A *factor* was introduced in Eq. (5) for the possibility to correct the direction of the vectors. The Mohr's circle was used to decide the directional correction of the vector. By default, the *factor* has the value 1 and can under conditions assume the value −1.

$$\text{eigenvector}(\sigma) * \mathit{factor} = \begin{bmatrix} e_{\mathrm{I}} & e_{\mathrm{II}} \end{bmatrix} \tag{5}$$

To rotate the stress tensor, the angle $\boldsymbol{\varphi}$ between the first principal stress (first eigenvalue) and the x-axis of the coordinate system was determined according to Eq. (6).

$$\boldsymbol{\varphi} = \cos^{-1}\left[\frac{e_{\mathrm{I}} \cdot \begin{bmatrix} 1 \\ 0 \end{bmatrix}}{\left[|e_{\mathrm{I}}| \cdot \left|\begin{matrix} 1 \\ 0 \end{matrix}\right|\right]}\right] \tag{6}$$

With the angle $\boldsymbol{\varphi}$ from Eq. (6) the rotation of the stress tensor was done according to Eq. (7).

$$\boldsymbol{\sigma}_{\mathbf{rot}} = \begin{bmatrix} \sigma_{\mathrm{rot,xx}} \\ \sigma_{\mathrm{rot,yy}} \\ \sigma_{\mathrm{rot,xy}} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}\left(\sigma_{\mathrm{xx}} + \sigma_{\mathrm{yy}}\right) + \frac{1}{2}\left(\left(\sigma_{\mathrm{xx}} - \sigma_{\mathrm{yy}}\right)\cos 2\varphi\right) + \sigma_{\mathrm{xy}} \sin 2\varphi \\ \frac{1}{2}\left(\sigma_{\mathrm{xx}} + \sigma_{\mathrm{yy}}\right) - \frac{1}{2}\left(\left(\sigma_{\mathrm{xx}} - \sigma_{\mathrm{yy}}\right)\cos 2\varphi\right) - \sigma_{\mathrm{xy}} \sin 2\varphi \\ -\frac{1}{2}\left(\sigma_{\mathrm{xx}-\sigma_{\mathrm{yy}}}\right)\sin 2\varphi + \sigma_{\mathrm{xy}} \cos 2\varphi \end{bmatrix} \tag{7}$$

The shear stress component $\sigma_{\mathrm{rot,xy}}$ of the stress tensor $\boldsymbol{\sigma}_{\mathrm{rot}}$ is changed to zero. The maximum and minimum values of the other values of the stress tensor $\boldsymbol{\sigma}_{\mathrm{rot}}$ must correspond to the maximum or minimum values of the eigenvalues λ_{i} (principal stresses) of the stress tensor $\boldsymbol{\sigma}$. If this is fulfilled and the deviation of Δ_{min} and Δ_{max} is in the range of a permitted tolerance, the alignment of the vector is correct. If the values for Δ are not within the permitted tolerance, a direction correction is required. The deviation is due to discretization errors in the numerical calculation. In theory, Δ_{min} and Δ_{max} are zero. For this purpose, the *factor* in Eq. (5) is set to −1 and the check is repeated according to Eqs. (6)–(8).

$$\begin{aligned} \Delta\mathrm{max} &= \max\left(\sigma_{\mathrm{rot,xx}}, \sigma_{\mathrm{rot,yy}}\right) - \max(\lambda_{\mathrm{i}}) \\ \Delta\mathrm{min} &= \min\left(\sigma_{\mathrm{rot,xx}}, \sigma_{\mathrm{rot,yy}}\right) - \min(\lambda_{\mathrm{i}}) \end{aligned} \tag{8}$$

4 Results

The evaluation and comparison of the first results was performed using a simple plate with a square hole, which was deformed elastically. To compare the results, the x-component $\boldsymbol{\sigma}_{\mathrm{x}}$ and the y-component $\boldsymbol{\sigma}_{\mathrm{y}}$ of the stress tensor were compared to the effective force $\boldsymbol{F}_{\mathrm{eff}}$ determined by the new evaluation method. The comparison was done with specimen in two different orientations in respect to the global coordinate system. In Fig. 2 the results of different variants are shown. When comparing the stress component σ_{x} form the different aligned specimens (see Fig. 2a) and b)), a dependence of the results on the orientation of the sample in the global coordinate

system can be seen. The same can also be seen when comparing the stress component σ_y of different aligned plates (see Fig. 2c) and d)). For the newly developed method, in comparison, it could be shown that this is invariant to the orientation of the specimen (see Fig. 2e) and f)).

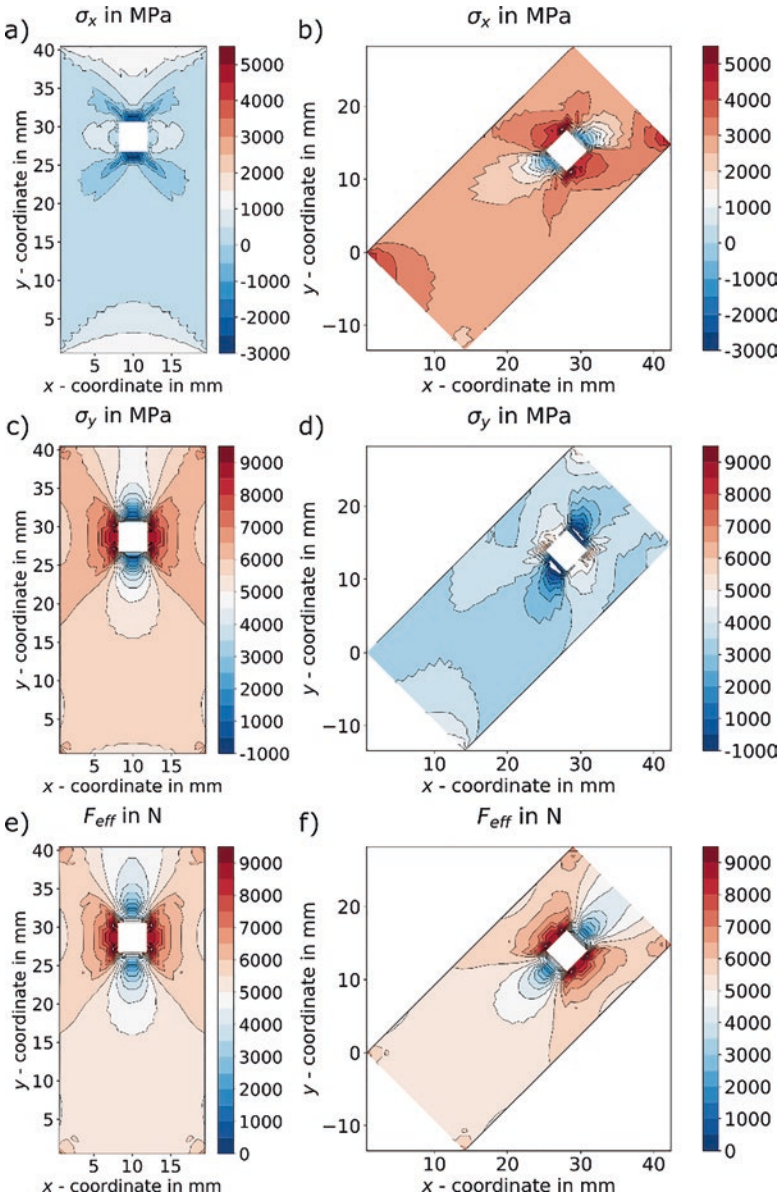


Fig. 2. Comparison of σ_x and σ_y of the stress tensor and the effective force $\boldsymbol{F}_{\text{eff}}$ determined by the new evaluation method for different orientations of the specimen

When comparing $\boldsymbol{\sigma}_x$ and $\boldsymbol{\sigma}_y$ in different orientations, clear differences in the stress fields were identified. In comparison, the new method was able to determine almost identical force fields regardless of the sample alignment. Thus, an independence of the sample orientation in the global coordinate system could be confirmed for the newly developed method. The independence of the results from the alignment of the specimen in the global coordinate system can be achieved, because the calculation of the effective force $\boldsymbol{F}_{\text{eff}}$ includes all local invariant principal stresses and their eigenvectors. These results will be used in the next steps to generate trajectories from the vector fields and can thus be used for a general analysis of the load paths which stresses the part and the joint.

5 Conclusion and Outlook

The basic assumption of the new method for the analysis of the force flow could be confirmed by the conducted studies by using a simple specimen with a square hole in different orientations. It was shown that the newly developed approach for the analysis of load flows is invariant to the orientation in the global coordinate system.

In further steps the method is extended to complex geometries, mechanical jointed components up to 3D assemblies. For the evaluation of the vector fields and the analysis of the load paths, algorithms for the building of trajectories are implemented according to different criteria. The presented approach is the fundament of a new design method that allows the positioning and design of joints based on load cases and given component structure and provides an evaluation criterion for the safety of the joint. Therefore, the loading of different regions in the part can be identified, similar to the concept of equivalent stress. However, this new method gives also the direction of loading.

Acknowledgement. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – TRR 285 – Project-ID 418701707, subproject B01. The computations were performed on a PC-Cluster at the Center for Information Services and High Performance Computing (ZIH) at TU Dresden.

References

1. Marhadi, K., Venkataraman, S.: Comparison of quantitative and qualitative information provided by different structural load path definitions. Int. J. Simul. Multi. Design Optim. **3**(3), 384–400 (2009)
2. Kelly, D.W., Elsley, M.: A procedure for determining load paths in elastic continua. Eng. Comput. **12**(5), 415–424 (1995)
3. Li, R., Kelly, D., Crosky, A., Schoen, H., Smollich, L.: Improving the efficiency of fiber steered composite joints using load path trajectories. J. Compos. Mater. **40**(18), 1645–1658 (2006)
4. Moldenhauer, H.: Die Visualisierung des Kraftflusses in Stahlbaukonstruktionen. Stahlbau **81**(1), 32–40 (2012)

5. Waldman, W., Heller, M., Kaye, R., Rose, F.: Advances in two-dimensional structural load-flow visualisation. Eng. Comput. **19**(3), 305–326 (2002)
6. Livermore LSTC: LS-DYNA Keyword user's manual - Volume I. 03/30/18 (r:9845) (2018)
7. Livermore LSTC: LS-DYNA Keyword user's manual - Volume II. 04/03/18 (r:9846) (2018)
8. van der Walt, S., Colbert, S.C., Varoquaux, G.: The NumPy array: a structure for efficient numerical computation. Comput. Sci. Eng. **13**(2), 22–30 (2011)
9. Virtanen, P., Gommers, R., Oliphant, T. E. et al.: SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat. Methods **17**, 261–272 (2020)
10. Hunter, J.D.: Matplotlib: a 2D graphics environment. Comput. Sci. Eng. **9**(3), 90–95 (2007)

Microstructure and Mechanical Properties of Thermomechanically Forged Tempering Steel 42CrMo4

J. Diefenbach(✉), K. Brunotte, and B.-A. Behrens

Institut für Umformtechnik und Umformmaschinen, Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany
diefenbach@ifum.uni-hannover.de

Abstract. Thermomechanically treated (TMT) materials are characterized by a fine-grained microstructure, which leads to extraordinary mechanical properties. In this study, the alloyed tempering steel 42CrMo4 (AISI 4137) is used to set up a two-step TMT upsetting process with intermediate cooling. A water-air based cooling system was used to adjust different phase configurations between the two steps by varying the target temperature and cooling rate. Standardized test specimens for Charpy impact tests and tensile tests as well as for metallographic analyses are cut out of the formed parts. Tensile tests showed that the yield strength can be enhanced up to 1188 MPa while the elongation at break is about 12% without any additional heat treatment by forming the material after rapid cooling. This fulfills the demands of the standard in quenched and tempered state. This process route allows for local-load tailored part design and manufacturing by adjusting the forming conditions conveniently.

Keywords: Thermomechanical treatment · Forging · Mechanical properties

1 Introduction

Forged components are able to fulfil high demands on load capacity and safety aspects, especially in the automotive sector. On the one hand, these parts are subjected to high dynamic loads and on the other hand they must meet the demands of lightweight engineering [1]. An additional challenge in production technology is to design resource efficient and sustainable process chains at low costs for these energy-intensive goods. Forging is a central production process in the manufacturing of such parts. Beside this process, especially the material and the following heat treatment are playing an important role in defining parts properties [2]. An excellent mixture of mechanical properties can be achieved by thermomechanical treatment (TMT). The combined influences of material, forming and heat treatment on the microstructure allow for properties, which are superior to those reached in separately performed production steps. TMT parts are characterized by a fine-grained microstructure, which

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 142–150, 2021.
https://doi.org/10.1007/978-3-662-62138-7_15

gives them an excellent relation between strength and toughness [3]. Additionally, by integrating heat treatment in the forming operation, significant savings of time, energy and costs are possible [4]. The process chain can be shortened significantly by modifying the process route towards an integrated TMT step. Thermomechanically treated materials are known from sheet metal processing [5] and extrusion [6] in particular. This approach is now being transferred to forging process chains.

2 Process-Integrated Thermomechanical Treatment

Thermomechanical treatment processes combine the mechanical influence of forming on material microstructure with the thermal influence of heat treatment. They have gained particular importance in the production of fine-grained structural steels. In these steel grades, the high strengths are not achieved by large proportions of alloying elements, but mainly by a highly fine-grained structure and precipitation of micro-alloy elements [7]. Due to complex process control, this technology is only adopted in the area of individual parts produced by forming technology, although its importance in the research has already been recognized. For example, Tekkaya et al. discussed the numerous influencing factors, measuring methods and simulation possibilities [2]. For the integration in forging process chains, mainly small-scale demonstration processes have been designed, which focus on the investigation of microstructures and the testing of mechanical properties in reduced test procedures [8]. Additionally, the basic influence of process-integrated TMT regarding the transformation behavior is a current field of research [9]. The material microstructure that is being formed has also a significant impact. In addition to the temperature at the time of forming, the type of microstructure depends above all on the cooling rate with which the forming temperature is approached from the austenitizing temperature [10].

3 Objectives

The aim of this work is to adjust microstructure and component properties by using a process-integrated thermomechanical treatment and to provide reliable values of the corresponding mechanical properties. The approach was to use TMT to modify the mechanical properties of forged parts made from tempering steel 42CrMo4 that are close to those of the quenched and tempered material. Of central importance for the usability of such properties in the industry is to assign them to the process conditions under which they are created. In addition, mechanical properties must be determined for later applications in such a way that they are available in reliable and applicable quantities. It is therefore necessary to carry out standardized test procedures such as tensile and notched bar impact tests (Charpy tests) with macroscopic test specimens.

4 Process Design and Test Setting

With respect to its high relevance in the industrial production of high performance components, the tempering steel 42CrMo4 (AISI 4137) was chosen for the examination. The chemical composition of the tested material is given in Table 1. All determined contents of alloying elements are in accordance with the corresponding standard (DIN EN ISO 683-2).

Table 1. Chemical composition of tested 42CrMo4 material determined by spark spectrometry (selected elements)

	C%	Si%	Mn%	S%	Cr%	Mo%	Ni%
42CrMo4	0.428	0.241	0.757	0.014	1.080	0.228	0.074
	−0,01 +0,006	−0,006 +0,004	−0,003 +0,011	−0,001 +0,001	−0,01 +0,0	−0,002 +0,004	−0,0 +0,001

Forming operations representing a thermomechanical treatment step need to be carried out in a temperature range that is significantly lower than the typical hot forging temperature. To achieve a specific and fine-grained microstructure with the desired properties, it is necessary to avoid instant recrystallization. The applied energy during forming results in an increase of dislocation density. This energy should be used to form new grain boundaries. At this reduced forming temperature (usually less than 800°C), the required forming force is already elevated and the material's total formability is reduced. For this reason, forming is divided into two steps. The major forming takes place at the hot forging temperature (1250 °C), followed by a post-forming step at reduced temperature which is used to adjust the microstructure and mechanical properties. As shown in Fig. 1 schematically, four thermal routes were designed using different phase states of the materials. In two of the investigated routes, forming is effected at a comparatively high temperature of 610 °C (hh and lh in Fig. 1) and in the other two at a reduced temperature of 460 °C (hl and ll in Fig. 1). The high temperature was chosen because it is closely above the transformation temperature. In previous examinations it was determined that at this temperature no transformation from austenite to a low temperature modification takes place, even when a low cooling rate is applied. At the lower temperature, this transformation takes place but the formability of the material is still enhanced due to the elevated temperature. Additionally, the cooling rate plays an important role when defining the phase state which dominates the material. Changes in the heat transfer behavior and exothermic transformations lead to strong temperature dependent variations of the cooling rate. Therefore, the cooling rate is given approximately only for the relevant temperature range before the transformation starts(900 °C–650 °C). Using a low cooling rate (2.5 $Ks{-}^{1}$) to reach the lower thermal level, the material partly transforms from austenite (A) to the ferritic-perlitic (F+P) modification. Post-forming is applied after the transformation (ll route). The other possibility is to suppress the transformation to F+P by rapid cooling (5 Ks^{-1}) to this temperature level. The material transforms to bainite (B) before and during forming, which leads to a microstructure with superior properties (hl route).

Initially, the basic material is heated to the hot forging temperature and then formed in the major forming operation. Afterwards the post-forming temperature is adjusted. Subsequently, the post-forming step is carried out and the parts are cooled down uncontrolled at ambient atmosphere. Furthermore, reference parts are forged in a single step to the final height at hot temperature.

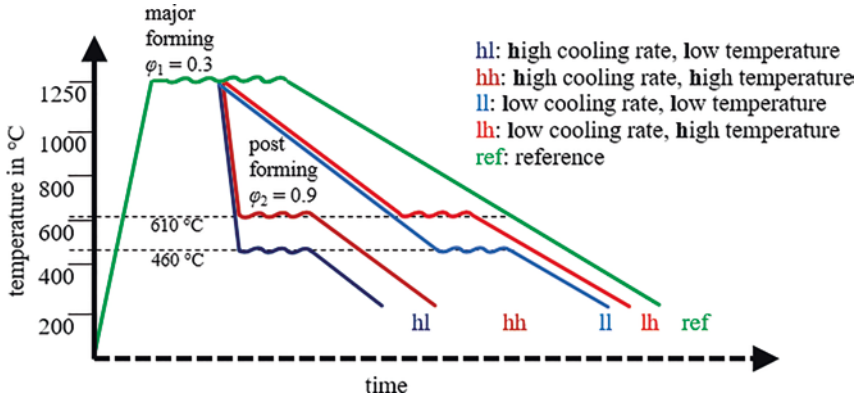


Fig. 1. Schematic TMT process routes and reference route

These thermal routes require a cooling device, which allows for adjusting the temperature by applying an individual cooling rate. The developed cooling station (Fig. 2) consists of six concentrically arranged spray nozzles atomizing water with compressed air. The flow rate and the pressure of water and air can be modified to adjust different cooling rates. The part's temperature can be measured by thermocouples and an integrated pyrometer. This setup was used in the final forging tests and for prior examinations of the transformation during cooling which led to the specific post-forming temperatures.

During the prior examinations, the core temperature was measured by thermocouples whilst the surface temperature was simultaneously measured by a pyrometer. The core temperature was used to determine the process routes with the given temperatures. Furthermore, the correlation between the core and surface temperature was determined in order to supervise the cooling process during the forging tests.

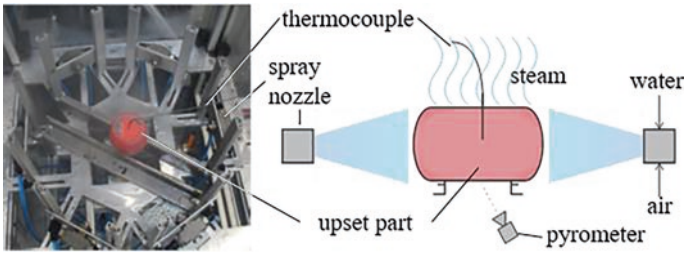


Fig. 2. Cooling station in operation (left) and schematically (right)

The geometry of the demonstrator part in this study was chosen considering two central criteria. The first criterion is the possibility to shape the parts in a two-step process. A forged and cooled part has to be processible in a second stage. If possible, it would be beneficial to perform post forming in the same tool as the major forming operation. The second criterion is the suitability for the following testing procedure. To supply industrial processes with reliable values for mechanical properties, the most suitable way is to use standardized tests, such as Charpy impact tests and tensile tests with macroscopic test specimens. The forged parts need to provide an area with approximately homogeneous deformation which is large enough to extract the specimens. As illustrated in Fig. 3 a cylindrical upsetting geometry with an initial diameter of 50 mm and a height of 80 mm was chosen. The schematic distribution of plastic strain in the forged part is also shown in the figure as well as the location of the test specimens that are cut out of the part's centre by wire cutting. In addition, the area between the round tensile specimen and the square Charpy specimen is also cut out and prepared for metallographic examination. One test piece per test procedure was manufactured from each component. This makes it possible to use the three described examination methods on a single component and thus to establish a clear assignment of microstructure, mechanical properties and process conditions.

A hydraulic press (Schirmer & Plate) was used to perform the upsetting with a ram speed of 30 mms^{-1}. The part temperature was only recorded during cooling in the cooling station.

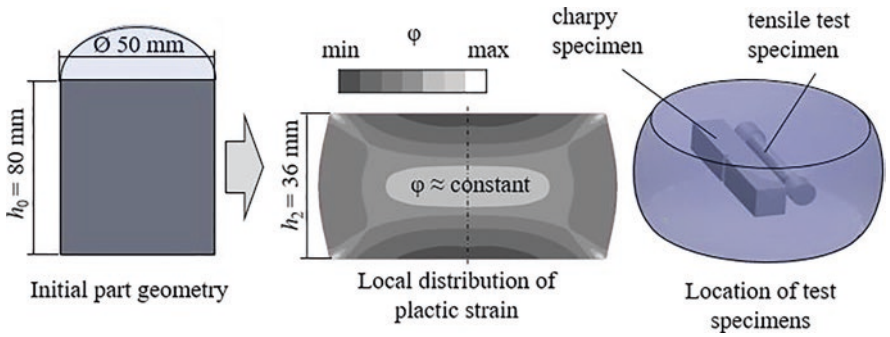


Fig. 3. Initial part geometry and forged part with schematic distribution of effective plastic strain and location of the test specimens

According to the standard DIN 50125, round (form B) tensile test specimens with a diameter (d_0) of 5 mm in the relevant area were manufactured. Specimens for Charpy impact tests with the dimensions 10 × 10 × 55 mm and v-notch were made according to standard DIN EN ISO 148.

In this paper, the focus is on the thermal component of the different TMT routes. During the major forming step, an effective plastic strain is applied to the relevant central part area (see Fig. 3 (central)) of $\varphi_1 = 0.3$ (height 62 mm). Subsequently, in the TMT post-forming operation, the part is upset with an effective plastic strain of $\varphi_2 = 0.9$ (final height 36 mm). Parts from the reference route were upset to the final height of 36 mm in one step at hot forging temperature.

5 Results

The different thermal process routes show a clear specific influence on the microstructure of the investigated material as presented in Fig. 4. First, it can be observed that forming, even on the reference route, leads to a considerable change in and refinement of the microstructure. During the manufacturing of the initial material by hot rolling, a very slow cooling from high temperature leads to a microstructure close to equilibrium with large ferrite (white) and perlite (dark/brown) areas. This is the initial state in which the raw billets are present (Fig. 4 top left). In perlite, cementite and ferrite are present side by side. These parallel structures can have different degrees of fineness so that the individual lamellae in the micro section can be partially seen under the light microscope. In both deformed microstructures, the white areas, i.e. the ferrite grains and the bainitic structures consisting of ferrite and cementite in non-lamellar orientation, are smaller. A lamellar structure as in perlite grains is mostly not visible here with the light microscope. Ferrite and precipitated cementite (dark) are present irregularly next to each other. This bainite microstructure represents an intermediate stage between the ferritic-perlitic modification and the martensitic one, which would result from very rapid cooling. The greatest effect on the microstructure can be seen in the process route with rapid cooling to low temperature (hl). In this modification, an arrangement of needle-like phases can be seen, which mark the structure as lower bainite. This microstructure is known for its excellent strength and toughness properties in a fine-grained form. The least effect can be seen when post-forming was performed at high temperature (lh and hh). Regardless of the cooling, a microstructure is seen which is very similar to that of the reference route and only slightly finer structured.

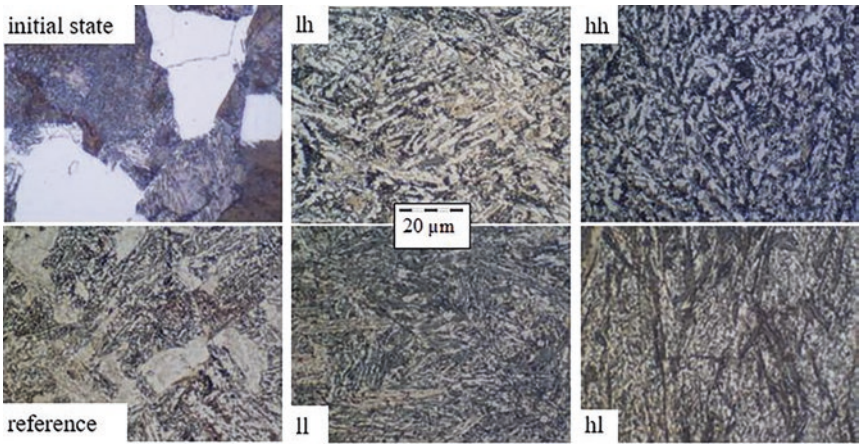


Fig. 4. Micrographs from different TMT routes, reference and initial state, magnification 1000:1, etched with nitric acid (5%)

Parallel to the light-microscopic analysis, the hardness of the specimens extracted from the core of the forged parts was measured in the metallographic examination according to the Vickers method (HV1). In Fig. 5 it can be seen that forming at elevated temperature always leads to an increase in hardness compared to the initial state. As expected, this increase is greatest for the hl route. The increase compared to the reference is clearly at 27%. Due to the close resemblance to the reference process at hot forming temperature, the hardness of the lh route shows no significant impact (−3%). The others are in between (ll: +12%, hh: +7%).

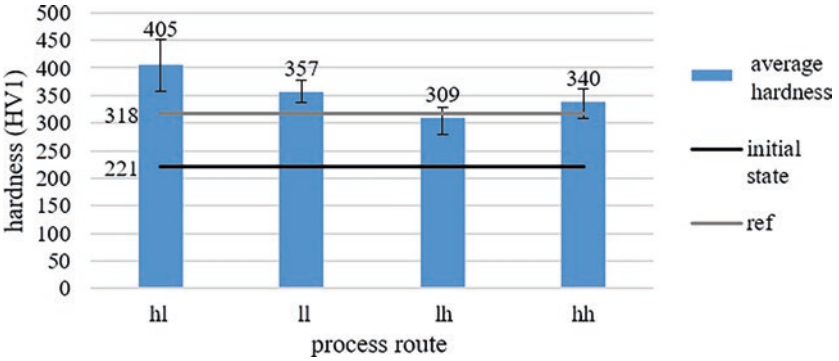


Fig. 5. Hardness HV1 from different TMT routes, reference and initial material, average of six indents

The mechanical properties were determined in tensile tests and Charpy impact tests. Figure 6 shows a summary of the mechanical properties tensile strength (R_m), elongation at break (A) and absorbed impact energy (Kv). It becomes clear that the hl route leads to the highest tensile strength similar to the evaluation of hardness. With 1188 MPa this is above the 900–1100 MPa range required by the standard (DIN EN ISO 683) for the relevant cross-section of the components (40 mm to 100 mm) in quenched and tempered state. This requirement is met by components from all thermal routes investigated. No significant influence of the thermal routes on the elongation at break is recorded. All examined components show an elongation at break of approx. 12% and thus just about meet the requirements of the standard. The development of the impact energy is particularly remarkable. This increases significantly during processing along the hl-route. Compared to the reference process, it is higher by a factor seven at 69 J. The ll-route also shows a significant increase in impact energy (59.5 J, more than factor six higher compared to parts from the reference route). The minimum impact energy of 35 J specified in the standard is only achieved by components from these two TMT routes.

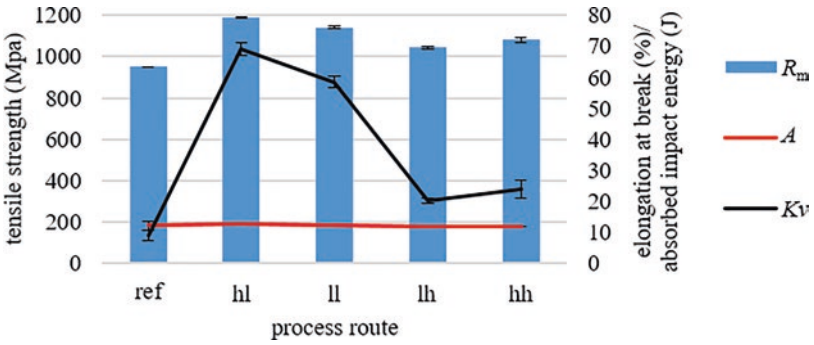


Fig. 6. Mechanical properties (tensile strength (R_m), elongation at break (A) and absorbed impact energy (Kv)) from different TMT routes and reference, average of two to three specimen for each testing method and process route

6 Summary and Conclusion

The investigations of various thermomechanical process routes during the processing of the tempering steel 42CrMo4 have shown that hot forming with the associated cooling to ambient temperature has a significant influence on the microstructure and mechanical properties of components. By varying both, the temperature during forming and the cooling rate starting from the hot forming temperature, bainitic microstructures with different characteristics can be produced. Standardized test methods with macroscopic test specimens were used to determine the mechanical properties. The strength properties of these thermomechanically forged materials are predominantly very good (up to 1188 MPa). In addition, good ductility properties (absorbed impact energy up to 69 J) can also be achieved, if rapid cooling to a low temperature level before post-forming is used selectively. This makes it possible to meet the requirements of the relevant standard for the quenched and tempered material without additional time and resource-intensive heat treatment. Finally, it can be stated that the hl route (with post forming at low temperature after rapid cooling) leads to the best mechanical properties.

7 Outlook

Besides the variation of thermal process routes, different mechanical routes with variations of the effective plastic strain applied in the two forming steps will be investigated. Furthermore, test series with different cooling conditions after post forming as well as with different heating rates before the main forming step are planned.

Acknowledgement. The results presented in this paper were obtained within the research project "Improvement and identification of mechanical properties in forged parts by immediate post-forming after hot forging on different temperature levels by influencing microstructure of tempering steel" (359921546). The authors would like to thank the Deutsche Forschungsgemeinschaft (German Research Foundation, DFG) for the financial and organisational support of this project.

References

1. Carruth, M.A., Allwood, J.M., Moynihan, M.C.: The technical potential for reducing metal requirements through lightweight product design. Resour. Conserv. Recycl. **57**, 48–60 (2011)
2. Tekkaya, A.E., Allwood, J.M., Bariani, P.F., Bruschi, S., et al.: Metal forming beyond shaping. Predicting and setting product properties. CIRP Ann. **64**(2), 629–653 (2015)
3. Calcagnotto, M., Ponge, D., Raabe, D.: Effect of grain refinement to 1 μm on strength and toughness of dual-phase steels. Mater. Sci. Eng., A **527**(29–30), 7832–7840 (2010)
4. Fischer, M., Dickert, H.-H., Bleck, W., Behrens, B.-A., et al.: EcoForge: energy-efficient process chain of a production of high forging parts. HTM **69**(4), 209–219 (2014)
5. Lis, A.K.: Mechanical properties and microstructure of ULCB steels affected by thermomechanical rolling, quenching and tempering. J. Mater. Process. Technol. **106**(1–3), 212–218 (2000)
6. Foydl, A., Segatori, A., Tekkaya, A.E., Brosius, A., et al.: Grain size evolution simulation in aluminium alloys AA 6082 and AA 7020 during hot forward extrusion process. Mater. Sci. Technol. **29**(1), 100–110 (2013)
7. Hou, H., Chen, Q., Liu, Q., Dong, H.: Grain refinement of a Nb–Ti microalloyed steel through heavy deformation controlled cooling. J. Mater. Process. Technol. **137**, 173–176 (2003)
8. Russo Spena, P., Firrao, D.: Thermomechanical warm forging of Ti–V, Ti–Nb, and Ti–B microalloyed medium carbon steels. Mater. Sci. Eng., A **560**, 208–215 (2013)
9. Behrens, B.-A., Chugreev, A., Kock, C.: Experimental-numerical approach to efficient TTT-generation for simulation of phase transformations in thermomechanical forming processes. IOP Conf. Ser. Mater. Sci. Eng. **461**, 012040 (2018) doi:https://doi.org/10.1088/1757-899X/461/1/012040
10. Xu, X.-X., Bai, B.-Z., Liu, D.-Y., Yuan, Y.: Effect of thermomechanical treatment temperature on structure and properties of CFB/M ultra-high strength steel. J. Iron. Steel Res. Int. **17**, 66–72 (2010)

Dynamic Performance of Polymer-Steel-Hybrids Manufactured by Means of Process Integration

M. Demes(✉), T. Ossowski, P. Kabala, S. Bienia, and K. Dröder

Institute of Machine Tools and Production Technology (IWF), Technische Universität Braunschweig, Langer Kamp 19b, 38106 Brunswick, Germany
m.demes@tu-braunschweig.de

Abstract. Hybrid components are commonly produced in two or more separate manufacturing steps. In order to apply fiber-reinforced thermoplastics (FRTP) adhesively to a hot stamped blank, it has to be reheated. Through the integration of production processes of processing hot stamped metal blank and FRTP, it is possible to form and join the two materials simultaneously without reheating the metal component. By setting up an automated experiment to interrupt the cooling step during the hot stamping process at an elevated temperature level, the residual temperature of the metal blank in the quenching phase can be used to enable an adhesive bond of the FRTP to the formed steel. Within the scope of this investigation, a process window suitable for the interruption of the hot stamping process and the subsequent application of the FRTP is determined. Finally, hybrid u-shaped profiles are examined under dynamic loads, demonstrating an increase in the mechanical performance.

Keywords: Manufacturing process integration · Hot stamping · Dynamic testing · Hybrid components

1 Introduction

Polymer-Steel-Hybrids for the reduction of component mass are nowadays processed in separate process steps. Hot stamping of boron-manganese steels is an established process in the automotive industry for the manufacturing of light vehicle body components [1, 2]. By using boron-manganese steel blanks, which are heated above the austenitizing temperature before forming, an easy formability is achieved due to reduced flow stresses of the heated material. The high material strength is achieved by subsequent cooling and hardening in the press tool [3, 4]. As a result, a high-performance, hot stamped car body in white component with reduced mass through reduced blank thickness can be manufactured [5, 6]. A further reduction in blank thickness through the use of higher strength hot stamped steels would result in a further reduction in the mass of the component. Another approach to realize lightweight design with hot stamped steels is the combined use of these with

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 151–160, 2021.
https://doi.org/10.1007/978-3-662-62138-7_16

fiber-reinforced thermoplastic (FRTP) [7]. FRTP with its unidirectional characteristics are particularly suitable for implementing lightweight design in the vehicles body in white [8, 9]. Due to high semi-finished product costs, unestablished manufacturing processes and incompatible joining processes, endless fiber-reinforced plastics have so far only been used in the high-priced vehicle segment [10].

By manufacturing hybrid components consisting of FRTP and metal blank, the targeted application of FRTP to the metal blank can create a hybrid component that enables the economic use of FRTP in large-scale automotive production [11, 12]. An established technique to manufacture hybrid components is the thermoforming of FRTP onto a metal blank [8, 13, 14]. For this purpose, the blank is heated before the application of the FRTP in order to provide a sufficient bonding. The application of FRTP on hot stamped blanks already enables the manufacturing of hybrid body in white components that offer high performance with reduced component weight [13]. However, blank processing and application of the FRTP are executed separately in the manufacturing of these particular hybrid components. This requires a reheating of the blank and thus a higher demand of process time and energy. Figure 1 shows the temperature profile of the materials during manufacturing of a hybrid component based on hot stamped steel and FRTP. After heating up the steel blank to the austenitizing temperature of 950 °C, it is transported to the tool, formed and quenched.

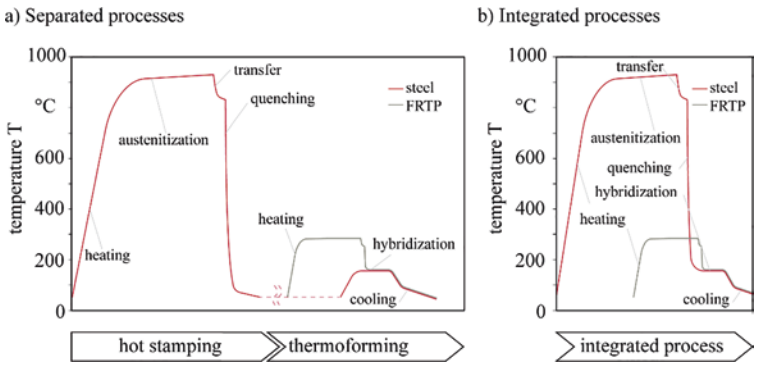


Fig. 1. Temperature profiles of separated processes and integrated process

After completion of the hot stamping process, the blank is fed into the thermoforming process and therefore reheated (Fig. 1a)). In order to integrate the hot stamping and thermoforming process, the quenching process of the blank is interrupted at an elevated temperature level of approx. 170 °C. This temperature range achieves the highest bond strength between FRTP and hot stamped blank and can be used for the integration of the separated process steps of hot stamping and thermoforming to manufacture a polymer steel hybrid by adhesive bonding (Fig. 1b)) [15]. Due to non-uniform contact conditions of the blank in the forming and quenching phase, a non-uniform temperature distribution in blank occurs after conventional hot stamping.

The local inhomogeneous temperature distribution in the blank does not result in the best bonding strength between FRTP and hot stamped steel. Through local tempering of the tool at the end of the quenching process allows achieving a more homogeneous temperature distribution on the blanks surface after hot stamping can be achieved, which consequently results in higher bonding strength of the hybrid component along the entire contact area.

In this paper, the dynamic structural behavior of application-oriented hybrid u-shaped profiles manufactured by a process-integrated hot stamping and thermoforming process, is examined. The dynamic investigation of the quality of hybrid u-shaped profiles allows conclusions to be drawn about the manufacturing process. The aim is to show how a temperature change in the metal blank before the application of FRTP affects the dynamic properties of hybrid demonstrators in an integrated process.

2 Demonstrator, Materials and Manufacturing of Specimen

This section begins with the derivation of a generic hybrid vehicle structure in the form of a demonstrator based on press-hardened steel and FRTP. The process-integrated production of hybrid demonstrators is described. By locally adjusting the die temperature, it is possible to bring the blank temperature to a homogeneous temperature distribution after hot stamping. It is expected that a more homogeneous sheet temperature after hot stamping leads to a better local connection between the materials. It is also expected that this will have an influence on the mechanical performance of hybrid structures.

2.1 Demonstrator Geometry

To investigate the properties of a process integrated manufactured generic body in white structure, a hybrid u-shaped profile as demonstrator is derived. Figure 2a) shows the cross-section with typical material thicknesses used e.g. in a B-pillar of a vehicle body in white design. The profile has a total length of 600 mm.

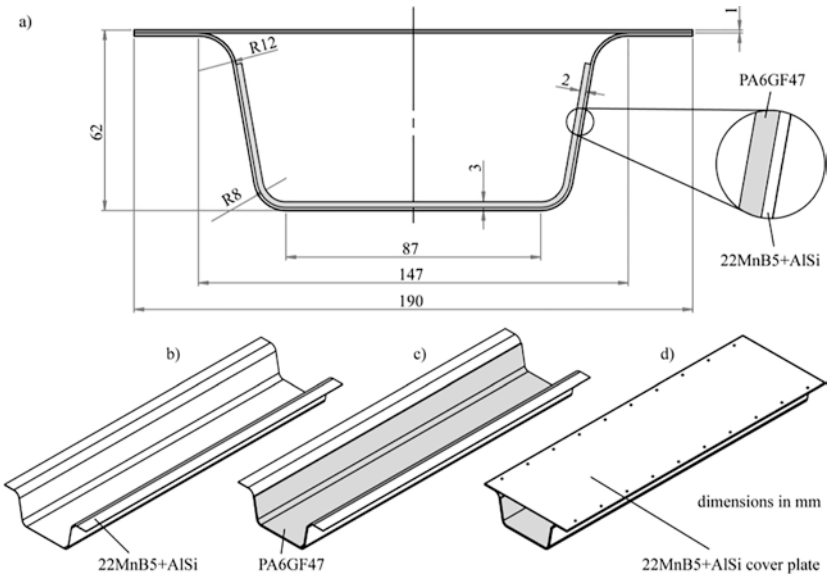


Fig. 2. Dimensions of the demonstrator geometry. a) Cross-section of hybrid u-shaped profile, b) hot stamped u-shaped profile, c) FRTP applied u-shaped profile and d) u-shaped profile closed with cover plate.

Besides the cross-section, the hot stamped u-shaped profile is shown after hot stamping (Fig. 2b)), after application of FRTP through thermoforming (Fig. 2c)) and after closing by a cover plate (Fig. 2d)).

2.2 Materials

In this study, a hybrid structure consisting of hot stamped steel and FRTP is investigated (Table 1). The FRTP used is a 2.0 mm thick glass-fiber-reinforced Polyamide 6 with a fiber volume fraction of 47% (PA6GF47). The hot stamped steel used is a 22MnB5 with an aluminum-silicon coating and a common used blank thickness of 1.0 mm. In delivery condition, it has a tensile strength of 532 MPa. After austenitizing to 950°C and completion of the forming and simultaneous quenching process, a tensile strength of 1518 MPa is achieved.

Table 1. Material properties of joining partners

	PA6GF47	22MnB5+AlSi (as delivered)	22MnB5+AlSi (hot stamped)
Fiber	Glass	–	–
Density ($g \cdot cm-3$)	1.8	7.9	7.9
Polymer	PA 6	–	–
Melt temperature (°C)	220	–	–
Fiber volume fraction (vol. %)	47	–	–
Thickness (mm)	2.0	1.0	1.0
Tensile Strength (MPa)	404 (long./trans.)	532	1518

2.3 Integrated Manufacturing of Specimen

The demonstrator profiles are manufactured using an integrated process consisting of hot stamping and immediate subsequent thermoforming, which is carried out on an automated experimental setup. The hot stamping process begins with the automated insertion of the steel blank into a chamber furnace where it is heated to a temperature of 950 °C. After a duration of 5 min and complete austenitization, it is automatically placed into a modified press tool within a reproducible transfer time of 7 s. Subsequently, the blank is formed and quenched to an elevated temperature level in a press by a force of 500 t. (Fig. 2c)). Previous investigations show, that the blank can be quenched to an elevated temperature level of up to 200 °C and yet a microstructure transformation from a ferritic-pearlitic microstructure to a martensitic microstructure takes place [17, 18]. By quenching to an elevated temperature level, the tensile strength and martensite volume percentage of the metal blank is slightly reduced. This loss of performance can be more than compensated by the application of FRTP after hot stamping on coupon specimens [17]. Based on previously carried out experimental work, the highest bonding strength between FRTP and hot stamped steel is achieved at a metal blank temperature of 170 °C after hot stamping. Therefore, a target temperature of 170 °C is defined for the metal blank after hot stamping. The hot stamping process is investigated using a tool modified with heating cartridges in order to quench the metal blank to an elevated temperature level. After hot stamping and quenching of the metal blank with an isothermal tool temperature, the temperature distribution on the metal blank surface can be inhomogeneous due to different contact conditions (e.g. pressure and contact time) between the blank and the tool. By the selective control of individual heating cartridges, a homogenization of the metal blank temperature after hot stamping is achieved. The required setting of the heating cartridges was determined in advance by simulating the process, taking into account the effective heat transfer mechanisms and material properties of the tool. Figure 3 shows the results of this inhomogeneous tool temperature control. To simplify the system, the punch is tempered homogeneously to 100 °C and the die is heated homogeneously to 170 °C (Fig. 3a)). An inhomogeneous temperature distribution of the die is set for the specific adjustment of the blank temperature after hot stamping (Fig. 3b).

Hot stamping of a u-shaped profile with the homogeneously tempered tool results in a temperature difference of $\Delta T = 45°C$. Due to the inhomogeneous temperature control of the die, the difference between the maximum temperature and target temperature of 170 °C in the blank can be reduced to $\Delta T = 18$ °C. This allows achieving an optimum and homogeneous bonding strength along the blank. Within the scope of the investigations, the hot stamping and thermoforming process is carried out in two stages. In the first stage, a modified hot stamping process is investigated in which the metal blank is quenched. In the second stage, knowledge of the blank temperature after hot stamping from the first stage in thermoforming is used.

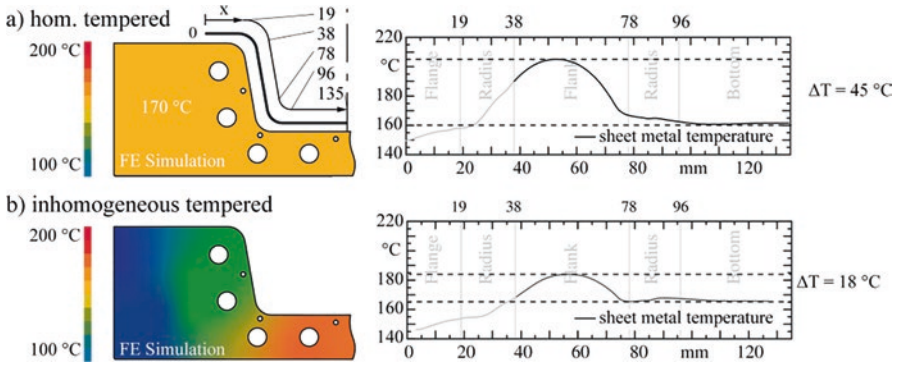


Fig. 3. a) Homogeneous tempered tool and b) inhomogeneous tempered tool for hot stamping

The temperature profile of the blank after hot stamping is mapped by the targeted control of heating cartridges in the thermoforming tool. The adhesive bond between the FRTP and the metal blank is thus enabled by the increased temperature of the blank after hot stamping. The adhesive bonding is supported by a powder based adhesion promoter of Vestamelt Hylink, which is applied on the blank with a thickness of 100 µm. The FRTP is heated up above its melting temperature to 300 °C in a furnace and then fed into the thermoforming process by an automated transfer using needle grippers. With a pressing force of 1000 kN and a dwell time of 10 s, the FRTP is consolidated and bonded to the formed hot stamped blank (Fig. 2c)). To illustrate an assembly situation in a generic vehicle structure, a cover sheet of hot stamped 22MnB5 (1.0 mm) by means of spot welding (Fig. 2d)) closes the u-shaped profiles.

3 Dynamic Performance of Hybrid U-Shaped Profiles

The dynamic investigation of the process-integrated manufactured demonstrators is intended to demonstrate the influence of the adjustment of the blank temperature before the application of FRTP. For this purpose, the experimental setup for the dynamic investigation of the demonstrators is described. In the evaluation, the

mechanical performance of the demonstrators will be assessed with regard to the process control during production.

3.1 Experimental Setup

Figure 4 shows the experimental setup for the dynamic testing of the hybrid u-shaped profiles manufactured in an integrated process. The profiles are positioned centered on two supports under the impactor. The total impactor mass is 438 kg and is pre-positioned at a height of 250 mm above the hybrid u-shaped profiles, resulting in a potential energy input of 1 kJ on impact.

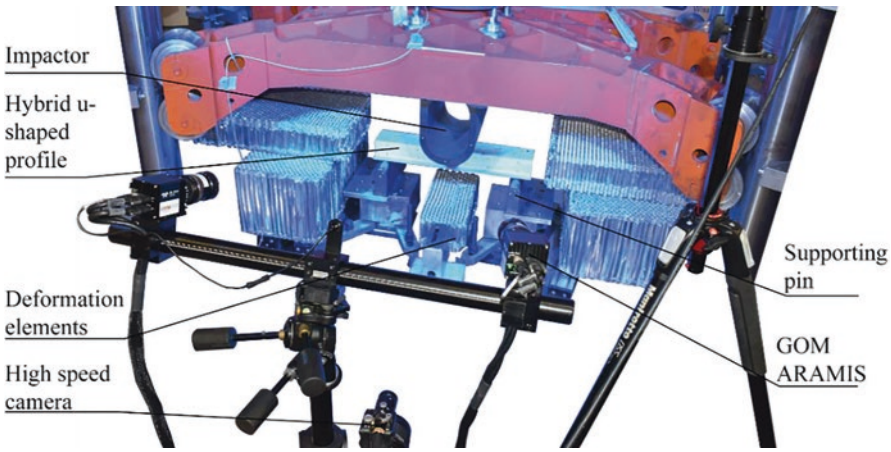


Fig. 4. Setup for the dynamic testing of u-shaped profiles manufactured in an integrated process

The position of the test specimen and the speed of the impactor was recorded by a high-speed camera and a GOM ARAMIS system. Deformation elements absorb energy in case of premature collapse of the u-shaped profiles. After releasing the impactor, it accelerates up to a speed of 1.9 m/s before the first contact with the u-shaped profiles. During the tests, the force-time diagrams and the energy-intrusion diagrams are recorded. In order to investigate the influence of tool tempering, hybrid u-shaped profiles manufactured with different process setups are evaluated. Table 2 shows an overview of the tested specimens and their composition. As a reference, a non-hot stamped u-shaped profile (NHS-1), a non-hybridized u-shaped profile without FRTP (NHP-2), a semi-finished product quenched to a higher inhomogeneous blank temperature level of 170°C (HUP-3) and homogeneous blank temperature levels of 170 °C (HUP-4) and 180 °C (HUP-5) are investigated. All investigated u-shaped profiles contain a hot stamped steel cover plate with a plate thickness of 1.0 mm.

Table 2. Overview of tested u-shaped profiles

	Hot stamped	Blank temperature	Blank temperature distribution	FRTP	Max. force	Max. intrusion
NHS-1	–	–	–	–	–	–
NHP-2	x	50 °C	Inhom.	–	–	–
HUP-3	x	170 °C	Inhom.	x	28.1 kN	46.0 mm
HUP-4	x	170 °C	Hom.	x	34.3 kN	40.5 mm
HUP-5	x	180 °C	Hom.	x	31.1 kN	40.5 mm

3.2 Results

Figure 5 shows the resulting force-time and energy-intrusion diagram for the examined hybrid profiles. For the reference u-shaped profiles of a non-hot stamped steel (NHS-1), and a hot stamped u-shaped profile without FRTP (NHP-2), no improvements can be observed, since the u-shaped profiles fail completely and the energy is absorbed by the deformation elements. The remaining hybrid u-shaped profiles result in a maximum force after 15 ms at different levels. The homogeneously tempered metal blank with a temperature of 170 °C after hot stamping (HUP-4) achieves a maximum force of 34.1 kN. For specimen HUP-3 an inhomogeneous temperature distribution in the blank is set by a homogeneously tempered 170 °C die. The maximum force obtained is 28.1 kN. If the homogeneous temperature distribution on the metal blank surface is increased to 180 °C after hot stamping (HUP-5), the maximum force reached is 31.1 kN. The minimum detected intrusion in the three examined hybrid profiles is 40.5 mm and results in both homogeneous temperature distributions in both the metal blank of 170 °C (HUP-4) and 180 °C (HUP-5) after hot stamping. An inhomogeneous metal blank surface temperature after hot stamping (HUP-3) results in an intrusion of 46.0 mm. The targeted quenching of the blank to a homogeneous sheet temperature of 170 °C ensures that the highest bond strength between the metal blank and TFRP are present in all areas of the sheet semi-finished product. The results show that the highest maximum force of 34.3 kN is observed for a hybrid u-shaped profile based on this homogeneous metal blank temperature after hot stamping (HUP-4). Quenching in a homogeneously tempered tool of 170°C results in an inhomogeneous blank temperature after hot stamping. Due to the temperature distribution in the metal blank, there are local differences in strength when the TFRP is bonded to the steel sheet, which explains a lower performance of 28.1 kN of the u-shaped profile (HUP-3). Hence, the experimental results emphasize, that a local tempering of the tool is necessary to increase the structural performance under dynamic loading.

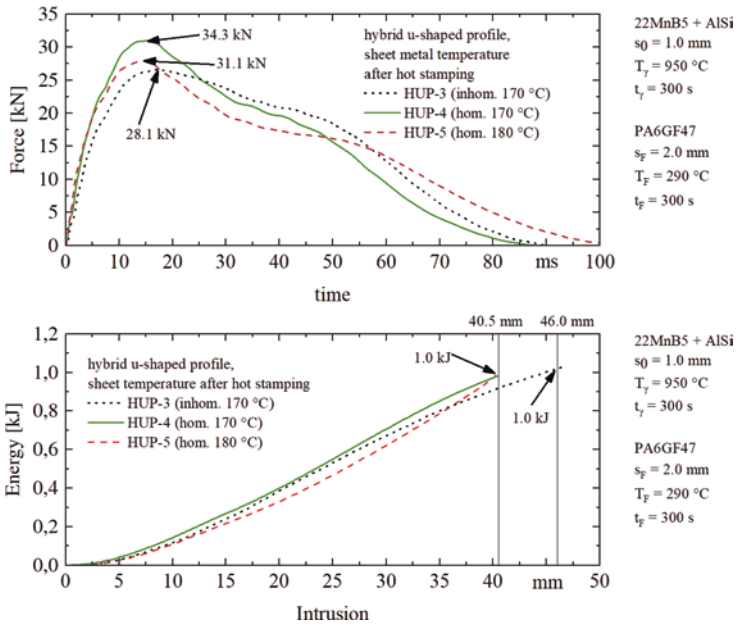


Fig. 5. Resulting force-time and energy-intrusion diagram for hybrid u-shaped profiles manufactured based on different blank temperatures after hot stamping

4 Conclusions

In an experimental setup for the dynamic testing of the hybrid u-shaped profiles of process-integrated manufactured hybrid profiles, it was proven that the performance in means of a maximum force of 34.3 kN could be achieved by setting a homogeneous blank temperature of 170 °C after hot stamping. Investigations of reference steel profiles without applied FRTP have shown that the energy cannot be absorbed and the test specimens fail prematurely. The reference investigation of a homogeneously tempered hybrid profile at 180 °C, which reaches a lower maximum force of 31.1 kN, shows that a precise temperature adjustment is necessary for the process integrated manufacturing of hybrid profiles. In this investigation, it was shown that the process integration in combination with a homogeneous blank temperature after hot stamping leads to a better mechanical performance. However, detailed investigations on the bonding behaviour in local areas needs to be carried out in detail to identify the factors that influence the process. Further, the characterization of the steel blank with regard to its influence on the mechanical properties in terms of hardness and strength is planned.

References

1. Karbasian, H., Tekkaya, A.E.: A review on hot stamping. J. Mater. Process. Technol. **15**, 2103–2118 (2010)
2. Taylor, T., Clough, A.: Critical review of automotive hot-stamped sheet steel from an industrial perspective. Mater. Sci. Technol. **10**, 1–53 (2018)
3. Radhankata, R., Singh, S.B.: Automotive Steels - Design, Processing and Applications. Woodhead Publishing, Metallurgy (2016)
4. Åkerström, P., Oldenburg, M.: Austenite decomposition during press hardening of a boron steel—Computer simulation and test. J. Mater. Process. Technol. **174**(1–3), 399–406 (2006)
5. Liu, H., Liu, W., Bao, J., Xing, Z., Song, B., Lei, C.: Numerical and experimental investigation into hot forming of ultra high strength steel sheet. J. Mater. Eng. Perform. **20**(1), 1–10 (2011)
6. Hu, P., Ying, L., He, B.: Hot Stamping Advanced Manufacturing Technology of Lightweight Car Body. Springer, Singapore (2017)
7. LANXESS Homepage. https://lanxess.com/de-DE/Presse/Presseinformationen/2019/09/. Accessed 15 May 2020
8. Klemt, C.: Verfahrensentwicklung zur Einbringung endlosfaserverstärkter Thermoplaste in metallische Strukturen mittels Patchen, TU Dresden, Dissertation (2016)
9. Plattform FOREL: PROLEI. Prozesskette für das Fügen endlosfaserverstärkter Kunststoffe mit Metallen in Leichtbaustrukturen
10. Dau, J., Lauter, C., Damerow, U., Homberg, W., Tröster, T.: Multimaterial systems for tailored automotive structural components. In: Conference Proceedings, 18th International Conference on Composite Materials (2011)
11. Bader, B., Berlin, W., Demes, M.: Steigerung der Leichtbaugüte durch interdisziplinäre Zusammenarbeit. Lightweight Des. **12**(3), 52–57 (2019)
12. Dreessen, P. C.: Verfahrensentwicklung zur spritzgussintegrierten Herstellung eigenschaftsoptimierter Kunststoff-Metall-Hybridstrukturen, Technische Universität Braunschweig u. Vulkan-Verlag GmbH, Dissertation, Braunschweig (2018)
13. FOREL Homepage. https://plattform-forel.de/wp-content/uploads/2015/02/Satellit_3D-Hybrid.pdf. Accessed 15 May 2020
14. Roth, S., Reg, Y., Götz, P., Masseria, F., Hühn, D., Bublitz, D.: Qualitätsgesicherte Prozesskettenverknüpfung zur Herstellung höchstbelastbarer intrinsischer Metall-FKV-Verbunde in 3D-Hybrid-Bauweise - Q-Pro, Dresden (2018)
15. Demes, M., Beuscher, J., Kühn, M., Dröder, K.: Influencing intermetallic layers of hot stamped steel for adhesively bonded plastic metal hybrids. Key Eng. Mater. **801**, 258–263 (2019)
16. Kühn, M.: Prozessabhängige Eigenschaften strukturierter Warmumformstähle in hybriden Materialverbunden. Schriftenreihe des Instituts für Werkzeugmaschinen und Fertigungstechnik der TU Braunschweig. Vulkan-Verlag, Essen (2019)
17. Demes, M., Janke, N., Beuscher, J., Kühn, M., Dröder, K.: Process integration of hot stamping steel and thermoforming fibre-reinforced thermoplastics. Procedia CIRP **85**, 141–146 (2019)
18. Gebken, T., Kühn, M., Demes, M., Müller, A., Dröder, K.: Integral manufacturing of plastic-metal hybrids consisting of endless fiber reinforcement using injection molding. MaterialsToday Proc. **8**, 804–811 (2019)

Investigation of the Scaling of Friction Coefficients from the Nano to the Micro Level for Base Materials and Coatings

N. Heimes[1(✉)], F. Pape[2], D. Konopka[2], S. Schöler[3], K. Möhwald[3], G. Poll[2], and B.-A. Behrens[1]

[1] Institut für Umformtechnik und Umformmaschinen (IFUM), Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany
heimes@ifum.uni-hannover.de
[2] Institut für Maschinenkonstruktion und Tribologie (IMKT), Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany
[3] Institut für Werkstoffkunde (IW), Leibniz Universität Hannover, An der Universität 2, 30823 Garbsen, Germany

Abstract. The application of solid lubricants offers high potential in terms of resource-efficiency in contacts under relative motion. The disadvantage of this lubrication method is the high risk of interface failure after the coating has been worn. In order to predict the service life of coated surfaces, it is necessary to determine the microscopic wear behaviour by analysing the coating hardness and coefficient of friction (CoF). Due to layer thicknesses under 3 µm, these parameters can only be determined using nanoindents and nano-scratch-tests. These tests require a transfer of material properties and tribological parameters from the nano to the micro level. The transfer method to the micro level and validity of the values determined at the nano level is presented while also describing challenges due to different influences of the respective size level. Additionally, oscillating sliding tests are carried out to determine the friction properties during sliding, which correspond to reference tests performed on the micro level.

Keywords: Nanoindentation · CoF · Wear behaviour · Micro tribology · Coatings

1 Introduction

Solid lubricants extend the life of rolling bearings, especially in applications that do not allow liquid-lubricated systems, such as the food industry or under vacuum conditions [1]. For this reason, research is being carried out in the priority program 2074 (project 2) on molybdenum-based solid lubricant systems, which are intended to extend the service life by means of self-regenerating lubricating layers. Thus, the solid lubricant molybdenum trioxide (MoO_3), which is removed in tribological contact, can

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 161–170, 2021.
https://doi.org/10.1007/978-3-662-62138-7_17

be formed again from a molybdenum deposit by oxidative processes. The molybdenum layers are applied to hardened rolling bearing steel (AISI 52100) by physical vapour deposition. A titanium layer is applied as an adhesion promoter to ensure a good bonding of the molybdenum layer to the bearing body. The pure molybdenum layer is completed by a thin oxide top layer. The aim is to produce a molybdenum trioxide layer on the surface, as such an interface offers very good lubricating properties [2]. The Archard wear model is frequently used approach for describing abrasive wear behaviour and is therefore used to determine the service life of the coating system. Archard describes the amount of wear W as a function of the sliding distance s, the hardness H and the normal force F_N with which the bodies are pressed together. The wear coefficient k depends on the process, friction and material pairing [3]. An extension of the Archard wear model is the definition according to Sarkar (Eq. 1). In this extension, the prevailing tribological conditions are introduced with a friction term by using the CoF μ [4].

$$W = k * F_N * s * \sqrt{\left(1 + 3\mu^2\right)}/H \tag{1}$$

The model requires material-dependent parameters, such as the material hardness H and the CoF μ, to be meaningful. For mono-materials, these values can be easily measured microscopically. Coatings in the range of less than 3 µm can only be characterised on the nano level. Furthermore, the testing of local areas on a component can be referred to as micro level testing, whereas the testing of the whole component is referred to as macro-testing. Bhushan et al. investigated the level dependence of the micro- and nanotribological properties [5]. They investigated different materials, coatings and lubricants for the application of electromechanical systems. The result of the study was that the CoF is influenced by two independent factors namely the scan size (size of the test area) and the used tip radius when transferring the CoFs from the nano to the micro level. As the scan size increases, the roughness values also increase, which in turn leads to an increase in the CoF, since the contact angle of the tip changes with larger roughness peaks. Due to the very small tip radius, very high Hertzian pressures occur at the nano level compared to the micro level [5].

Essentially, in this study, the nano-hardness, the surface roughness and the CoF under furrows are investigated at the nano level for the molybdenum coating and the mono-material. In addition, sliding tests on nano level with different test loads and tip sizes are performed to investigate the sliding value for different Hertzian pressures on all specimens. For a transfer to the micro level, sliding tests are performed on polished and ground specimens to determine the influence of the grinding direction on the evolution of the CoF. The scaling of the CoF is limited to mono-material to verify the scaling method. Performing the sliding test for the thin molybdenum layer would always be influenced by the substrate and thus be an additional interference factor for checking the scaling method. Finally, the evolution of the sliding values and the resulting wear marks are analysed and compared on both levels.

A possible scaling of the characteristic values from the nano to the micro level would be transferable to further applications. This would allow wear models such as the one according to Archard to be determined completely at the nano level and thus predict, for example, the wear behaviour of coated forming dies. Especially with

regard to the high contact pressures in coated forging dies, modelling at the nano level could be promising. With reliable wear models, the failure time of coated rolling bearings as well as of coated forging dies can be precisely determined, thus increasing safety in operation and reducing failure costs.

2 Experimental Setup

2.1 Specimen Preparation

Hardened rolling bearing steel with different surface conditions was used as substrate specimen. On the one hand, the specimen was ground plane-parallel and subsequently polished to a mirror finish and on the other hand ground to a technical surface roughness. Specimens with molybdenum layers were first ground and plasma-etched before the sputtering process. The layer structure consists of the hardened substrate, an approx. 200 nm thick adhesion promoter layer of titanium and an approx. 2 µm molybdenum layer. The dimensions of the specimen as well as their surface preparation and naming are listed in Table 1.

Table 1. Experimental setup of specimen

Material	Specimen name	Size (mm)	Manufacturing
AISI 51200	G_p	∅ 40 × 4	Polished to mirror finish
AISI 51200	G_g	∅ 15 × 5	Grounded
Molybdenum	M_o	∅ 15 × 5	Grounded + sputtered

2.2 Test Setup for Nano Level

The investigations at the nano level were performed on the Hysitron TriboIndenter® TI 950, where a force range from 2 µN to 10000 µN can be applied for different tests. A triangular Berkovich diamond tip with a tip radius of 77 nm was used to determine the nano-hardness and the modulus of elasticity. For the indents, a trapezoidal force-displacement function was used to neutralise dynamic effects [6, 7]. To determine the hardness over an area size of 36 × 36 µm, hardness maps were created using XPM for accelerated property mapping. Up to six indents per second can be carried out and accelerate the workflow significantly.

Scratch tests are used to characterise the elastic and plastic material behaviour. Additionally, the CoF can be determined by measuring the resulting lateral force. A conical diamond tip with a tip radius of 0.29 µm was used to avoid the preferred direction of the tip during scratching. The normal force is continuously increased over a length of 8 µm during scratching and reaches the set maximum force at the end of the scratching. To analyse the surface roughness, the specimens were tactilely measured before each test. This is achieved with a minimal force of 2 µN at a high-resolution scanning rate. The resulting resolution of the nano-surface roughness depends on the radius of the tip analogous to measurements on the micro level. On the nano level, five fields with a size of 40 x 40 µm per specimen were measured tactilely with the Berkovich tip.

The sliding test was derived from the test program of the scratch test. According to a method applied by Pape et al. an oscillating movement of the tip was used to achieve wear of the coating [8]. The programmed course of the normal force and the lateral displacement over the test time is shown in Fig. 1 for a sliding force of 200 µN. Analogous to the scratch test, the sliding distance of 8 µm was scanned by means of a prescan with a force of 2 µN, green segment. Within the first two sliding segments, the final sliding force is applied step by step. In the first segment (coloured in blue), the force is increased from 0 µN to half of the final sliding force. On the way back, the sliding force is further increased until the set sliding force is reached (coloured in red). The stepwise increase reduces the load on the tip on the way back. Then, the tip oscillates with a sliding speed of 0.8 µm/s, 600 cycles over the sliding distance. Finally, the wear track is scanned in the post scan with a normal force of 2 µN. For the sliding test, conical tips with a tip radius of 0.29 µm and a tip radius of 1.64 µm were used to test different Hertzian pressures. To induce Hertzian pressures as low as possible, small forces are used for the sliding test.

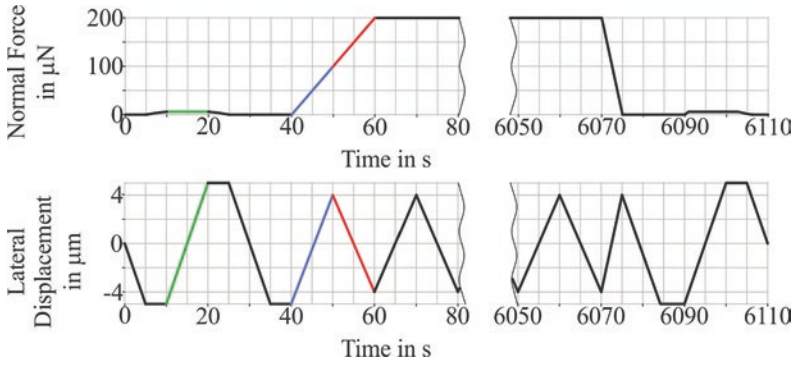


Fig. 1. Programmed course of normal force and lateral displacement over the time of the sliding test

2.3 Test Setup for Micro Level

A milli-tribometer is used for investigations on the micro level. For the investigations focused in this study, a sliding test (ball-on-disc) is used. A steel ball is placed on the specimen with a defined normal force. An oscillating movement with constant speed is performed by a linear stage on which the specimen is mounted. The oscillation takes place over 600 cycles (according to the nano tests) with a sliding distance per cycle of 2.5 mm, whereby the influences of different sliding velocities are tested. For the variation of the Hertzian pressure, balls with a radius of 3 mm and 1.5 mm are used, whereby the normal force is set to the maximum of 3 N to achieve a maximum Hertzian pressure. During sliding, the normal and lateral forces are measured continuously. After the test, the surfaces are examined regarding their wear behaviour.

The measurement of surface roughness at the micro level was carried out on the manual roughness tester HOMMEL-ETAMIC TURBO WAVE V7.53. For each

specimen six scans with a length of 4.8 mm and a tip radius of 500 µm were performed to determine the roughness.

3 Results and Discussion

3.1 Surface Roughness on Nano and Micro Level

The surface roughness was measured tactilely on nano and micro level. The results are shown in Table 2. For the ground specimens G_g and M_o, the values are close together. The reduced roughness for M_o has been expected, as the roughness is reduced when the layer is applied. The differences between nano and micro measurement at sample G_p can be caused by the different scan sizes.

Table 2. Roughness values of the specimens measured at the nano and micro level

Specimen name	Ra at nano level (µm)	Ra at micro level (µm)
G_p	0.009	0.081
G_g	0.149	0.137
M_o	0.095	0.113

3.2 Nano Hardness

First, the nano hardness was determined for the specimens M_o and G_p. For this purpose, 100 indents were measured at each of three measuring positions. One field is shown in Fig. 2a as an example of a hardness map for G_p, where the hard and softer phases are clearly visible. Hardness values of all measuring points are shown in Fig. 2b. The M_o layer was determined with a test load of 500 µN and the bearing steel with a test load of 1000 µN. The lower test load for M_o was chosen to ensure that only the hardness of the layer could be determined without the subtract affecting the hardness values. A M_o-hardness of 2.8 ± 0.3 GPa correlates very well with the hardness determined in further studies presented where molybdenum layers with different thicknesses were tested [9].

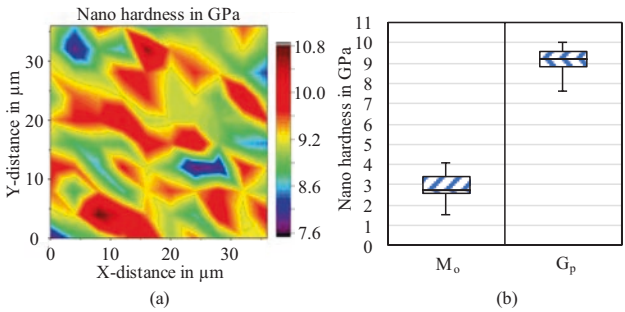


Fig. 2. Nano hardness map for G_p (a), nano hardness box plot for M_o and G_p (b)

3.3 Nano Scratch Tests

The elastic and plastic material behaviour on the nano level is determined by means of scratch tests. Figure 3 shows the results of the scratch test for M_o and G_p. The scratch force was chosen similar to the hardness measurements. Compared to bearing steel, molybdenum behaves very plastically, since the curve of the scratch and post-scan is very narrow. This behaviour was to be expected, since the E-modulus of the G_p with 225 ± 20 GPa is higher than the E-modulus of the M_o layer with 170 ± 15 GPa. For scratch tests, the CoF under furrows can be calculated additionally by the measured normal and lateral forces. The molybdenum shows its friction-reducing effect with a CoF of $\mu = 0.2$, while a CoF of $\mu = 0.55$ was determined for the bearing steel.

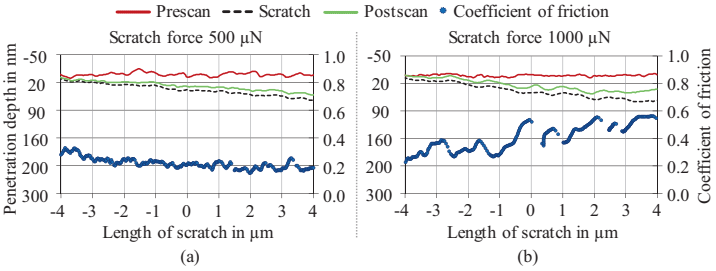


Fig. 3. Results of the nano scratch tests for M_o (a) and G_p (b)

3.4 Sliding Test on the Nano Level

For the sliding tests, the normal forces were varied between 25 µN to 500 µN. The sliding tracks for the varied normal forces are shown in Fig. 4a, which were performed with a tip radius of 0.29 µm on the polished bearing steel. Above a force of 75 µN, the sliding traces become clearly visible because the material pile-up increases with increasing force, as the tip penetrates deeper into the material.

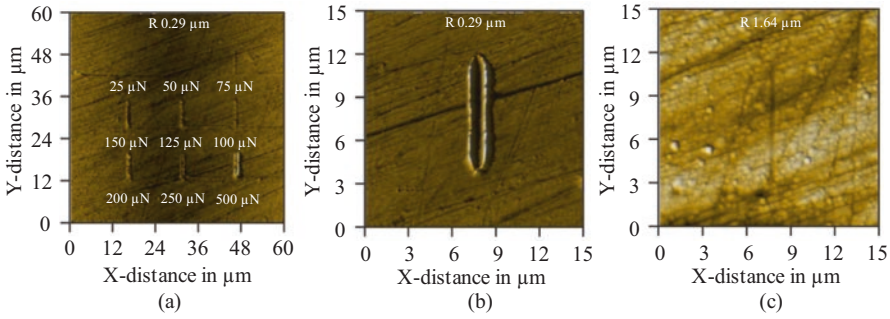


Fig. 4. Overview for sliding tracks for different sliding forces on the nano level (a) sliding track with 500 µN and a tip radius R of 0.29 µm (b) sliding track with 500 µN and a tip radius R of 1.64 µm (c) all tracks located at G_p

At forces of 25 µN and 50 µN, the sliding traces were very weak and no significant penetration of the tip could be detected over the sliding test. In addition, the specified normal force could not always be maintained at low forces, so that the force signal deviated from the nominal value. In Fig. 4b and c the sliding tracks for the different tip radii at a test force of 500 µN are shown magnified. With a smaller tip radius, the tip penetrates deeper and thus displaces more material (Fig. 4b). The sliding track with the larger radius does not show any major material displacement (Fig. 4c). Larger tip radius results in lower Hertzian pressure and thus the load is transmitted over a larger contact area into the material resulting in less penetration. It is also visible that the larger radius of the tip reduces the resolution, as the tip can no longer penetrate into the fine hills and valleys of the polished specimen.

3.5 Sliding Tests on the Micro Level

An overview of the sliding track for different angles to the grinding direction on specimen G_g is shown in Fig. 5. The uneven longitudinal edges of the sliding tracks for the angles of 45° and 90° compared to 0° are clearly visible. This indicates uneven contact of the ball due to the sliding grooves.

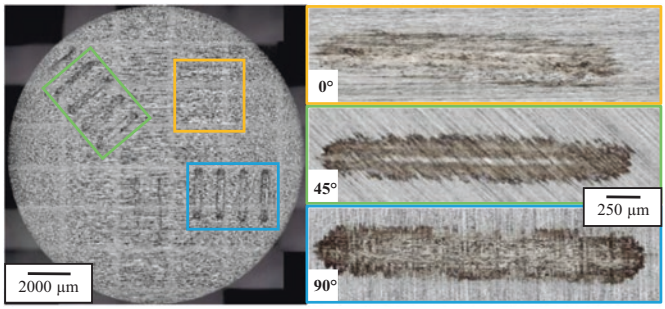


Fig. 5. Overview of the sliding tracks on micro level for different angles to the grinding direction on G_g

3.6 Comparison of the Sliding Tests

Sliding test results at the nano (blue box) and micro (red box) levels are shown in Fig. 6. The progressions on specimen G_p are shown in dotted lines. The CoF at the nano level decreases rapidly after the application of the normal force of 75 µN to a constant level of $\mu = 0.15$ independent of the used tip radius. In the further course, the CoF with a tip radius of 1.64 µm is slightly above the measurement with a tip radius of 0.29 µm. At the micro level, the sliding tests were performed with a spherical radius of 3 mm and a maximum force of 3 N. The resulting CoF suddenly increases to $\mu = 0.7$ and settles over the segments at $\mu = 0.8$. On specimen G_g, sliding tests were carried out on the micro level with different orientations to the grinding direction, which are shown as a solid line in Fig. 6. The influence of the sliding direction could not be determined on the nano level, because the sliding distance is too small there to test several sliding grooves in one test. The tests were performed in parallel 0°, diagonal 45° and transverse 90° to the grinding direction. The course 0° to the grinding

direction behaves like the sliding test on the polished specimen G_p. While the results of the sliding test at 45° and 90° to the grinding direction initially form a small plateau for approx. 35 segments with a CoF of $\mu = 0.19$. In this range, the CoFs are equal on the nano- and micro-levels. During the further course of the test, the CoF increases to $\mu = 0.8$ as with all tests on the micro level.

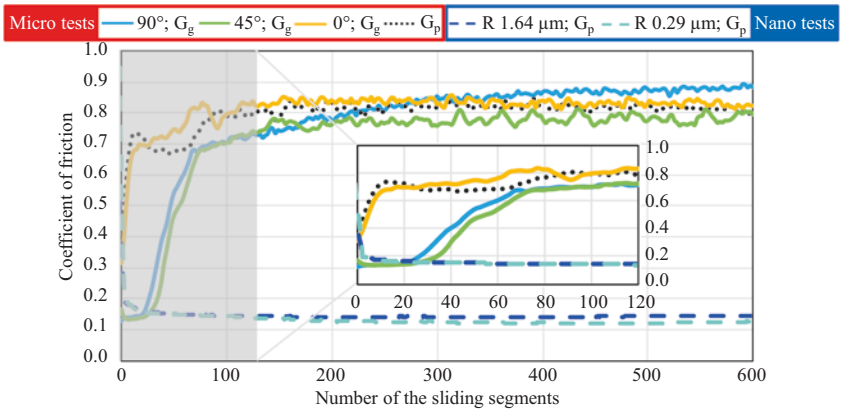


Fig. 6. Development of the CoF over the sliding segments for the nano level (0.8 µm/s, 75 µN) and micro level (6 mm/s, 3 N, R 3 mm)

The large difference in the CoF between the nano and micro experiments could be caused by the strongly differing Hertzian pressure. The development of the Hertzian pressure as a function of the used tip radius and normal force is shown in Fig. 7. The forces of the micro level are plotted on the vertical primary axis and those of the nano level on the secondary axis. In order to approximate the Hertzian pressures, the forces on the micro level have to be chosen as high as possible with a small tip radius. However, on the nano level, the forces must be as small as possible and the radius as large as possible to bring the pressures closer together. For the used tip radii and normal forces in the sliding test the following differences in the Hertzian pressure result, for R 3 mm/R 1.64 µm (mirko/nano) is Δ 4.92 GPa and for R 3 mm/R 0.29 µm is Δ 17.61 GPa.

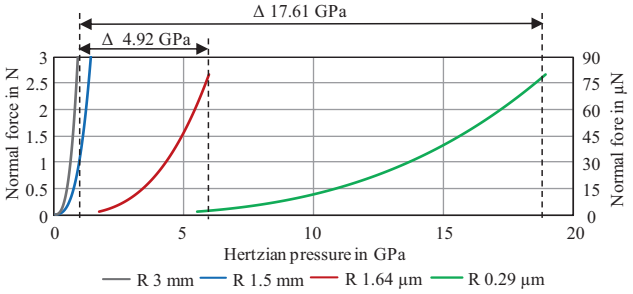


Fig. 7. Development of the Hertzian pressure as a function of the used tip radius and normal force for nano and micro level

However, the sliding test on the specimen G_g for the tests 45° and 90° to the grinding direction within the first 35 segments show a good agreement with the friction value on the nano level. One reason for this could be the small contact area within the first 35 segments, as the roughness peaks are initially smoothed in there. For better understanding, Fig. 8a shows a schematic representation of that kind of smoothing. In phase I, the ball is only in contact with a few roughness peaks, which increases the Hertzian pressure. As a result, the difference in Hertzian pressure between the nano and micro levels is smaller and the CoFs are equal. During phase II, the roughness is smoothened more and more and the contact area increases, which again increases the pressure difference and the CoFs differ from each other. In phase III, the pressure difference is even greater, because the ball is now completely in tribological contact. In tests 0° G_g and G_p, the CoF immediately increases steeply, since the contact surface is larger at the beginning. At 0° G_g, the larger contact surface is due to the parallel grinding grooves. The steep increase at G_p is due to the low surface roughness, as this was polished to a mirror finish. To further reduce the difference in Hertzian pressure, tests were carried out on the micro level with a sphere radius of 1.5 mm. Figure 8b shows the averaged CoFs for all tests on the specimen G_p on the micro level. Different sliding speeds were tested, but no significant influence could be determined in agreement with the investigation by Koinkar et al. [10]. The tests with a sphere radius of 1.5 mm show a large scattering because the ball holder could not always clamp the ball, so that it could partially roll during the test. However, it can be seen that an increase in Hertzian pressure by using a smaller ball radius reduces the CoF.

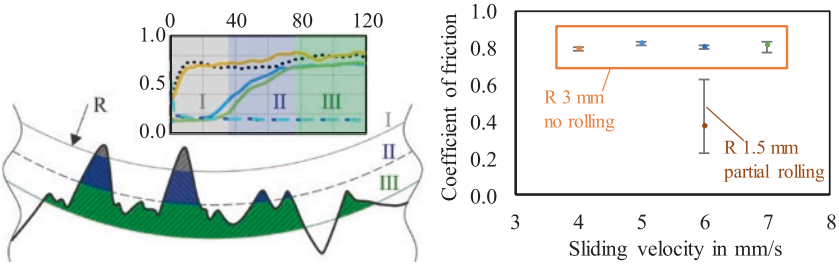


Fig. 8. Smoothing of the roughness peaks during sliding on the micro level (a), CoF for different velocities and ball radii in sliding test on the micro level (b)

4 Conclusions and Outlook

Conclusions of this study can be summarised as follows:

- The nano-hardness and the modulus of elasticity were determined for the substrate and the molybdenum layer. In addition, the friction coefficients under furrows were recorded on the nano level.
- The friction coefficients recorded in the sliding test at the nano and micro level showed large differences, so that a simple scaling with the selected setup is not possible.

- It could be shown that the Hertzian pressure is essential for the transferability of the friction coefficients. By bringing the Hertzian pressure closer together, the coefficients of friction can also be determined from the nano and micro levels.

In the future, further tests with a smaller sphere radius should be carried out on the micro level, in order to bring the Hertzian pressure closer to the pressure on the nano level. For this purpose, a new holder for smaller spheres must be used to prevent the sphere from rolling during the sliding tests. After successful implementation, the sliding tests for the molybdenum layers should be performed at the nano- and micro-levels to verify the validity of the method for layer systems. The overall goal is to develop a wear model on the nano level and subsequently transfer it to the micro level.

5 Acknowledgments

The financial support of this study by the German Research Foundation (DFG funding code 407673224) within the scope of the priority program “Fluidless Lubrication systems with high mechanical Load” (SPP 2074) is gratefully acknowledged.

References

1. Birkhofer, H., Kümmerle, T.: Feststoffgeschmierte Wälzlager - Einsatz, Grundlagen und Auslegung. Springer, Berlin (2012)
2. Dhanasekaran, S., Gnanamoorthy, R.: Microstructure, strength and tribological behavior of Fe–C–Cu–Ni sintered steels prepared with MoS2 addition. J. Mater. Sci. **42**, 4659–4666 (2007)
3. Archard, J.F.: Contact and Rubbing of Flat Surfaces. J. Appl. Phys. **24**(8), 981–988 (1953)
4. Sarkar, A.D.: Friction and Wear. Academic Press, London (1980)
5. Bhushan, B., Tambe, N.S.: Scale dependence of micro/nano-friction and adhesion of MEMS/NEMS materials coatings and lubricants. Nanotechnology **15**, 1561–1570 (2004)
6. Fischer-Cripps, A.-C.: Nanoindentation, 3rd edn. Springer, New York (2011)
7. Behrens B.-A., Bouguecha A., Vucetic M., Peshekhodov I., Matthias T., Kolbasnikov N., Sokolov S., Ganin S.: Experimental investigations on the state of the friction- welded joint zone in steel hybrid components after process-relevant thermo-mechanical loadings. AIP Conf. Proc. **1769**, 130013 (2016)
8. Pape, F., Gatzen, H.H.: Resistance of Multilayer Diamond-like Carbon (DLC) Coatings for MEMS Applications against Oscillating Wear. World Tribology Congress 2009, Japan (2009)
9. Heimes, N., Pape, F., Konopka, D., Schöler, S., Möhwald, K., Poll, G., Behrens, B.-A.: Characterisation of self-regenerative dry lubricated layers on mo-basis by nano mechanical testing, production at the leading edge of technology. In: Proceedings of the 9th Congress of the German Academic Association for Production Technology (WGP), Hamburg, pp. 139–148 (2019)
10. Koinkar, V.N., Bhushan, B.: Effect of scan size and surface roughness on microscale friction measurements. J. Appl. Phys. **81**, 2472–2479 (1997)

Investigation of Parameters Influencing the Producibility of Anodes for Sodium-Ion Battery Cells

J. Hofmann[1(✉)], A.-K. Wurba[1], B. Bold[1], S. Maliha[1], P. Schollmeyer[1], J. Fleischer[1], J. Klemens[2], P. Scharfer[2], and W. Schabel[2]

[1] Karlsruhe Institute of Technology, Wbk Institute of Production Science, 76131 Karlsruhe, Germany
janna.hofmann@kit.edu
[2] Karlsruhe Institute of Technology, Institute of Thermal Process Engineering (TVT), TFT Thin Film Technology, 76131 Karlsruhe, Germany

Abstract. Lithium-ion battery cells will dominate the market in the next 10 years. However, the use of certain materials as cobalt is a critical issue today and is constantly being reduced. Sodium-ion batteries are an alternative, which has already been researched on a laboratory scale. Increasing of the individual production steps are serious bottlenecks for bringing basic cell concepts into application. Within this paper a systematic investigation of parameters influencing the producibility for sodium-ion battery cells will be taken into account. For this purpose, the characteristic process variables and challenges along the production chain are presented along the process chain of lithium-ion battery cells. The influence of various process-machine interactions on the properties of the electrode is illustrated using the anode of sodium-ion batteries as an example. First conclusions whether the production technology can be adapted to the cell chemistry of the future at an early stage will be made.

Keywords: Battery cell production · Post-Lithium batteries · Sodium-Ion batteries

1 Motivation and Introduction

With the increasing demand for lithium-ion batteries (LIB), the earth's lithium resources are decreasing and it is already foreseeable that the currently known reserves of lithium and cobalt will be used up within the next 10–15 years. (see Fig. 1a). As shown in Fig. 1b other elements as sodium or magnesium have a much higher earth crust abundance and thus are more suited to cover future demands [1].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 171–181, 2021.
https://doi.org/10.1007/978-3-662-62138-7_18

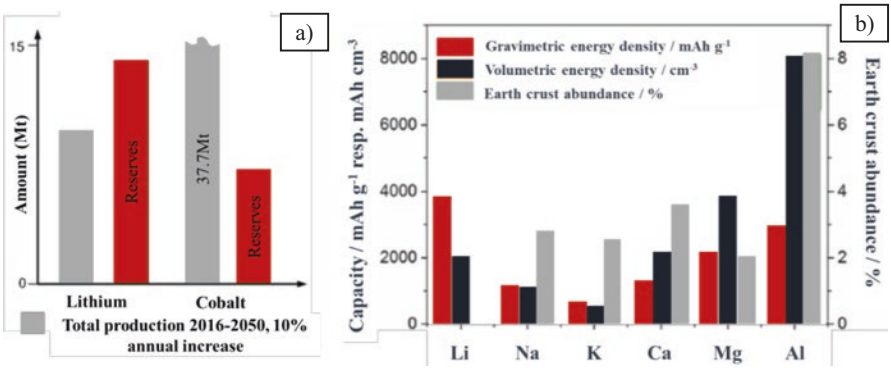


Fig. 1. (a) Forecast of lithium and cobalt reserves [1], (b) earth crust abundance of sodium [2] (own illustrations)

The mission of the DFG-funded Excellence Cluster Energy Storage beyond Lithium (POLiS) is to lay the foundation for new electrochemical energy storage systems which overcome the limitations of current battery technology. The investigation and the synthesis of new materials and their processability are to be predicted systematically in order to be able to exert an influence from the production technology already during the development process of the materials. This article gives a first step towards the prediction of the producibility of sodium-ion battery (SIB) anodes.

The operating principle of SIB, like that of the LIB, is based on the reversible inter- and deintercalation of ions [3]. The components of both battery systems are identical and differ essentially only in the materials used and the ionic charge carrier. Nevertheless, similar compounds can be used on the cathode side in the material system. However, the use of sodium-ions leads to significant differences between the two systems [4]. On the one hand, sodium-ions, in contrast to lithium-ions, are considerably larger (+25% to +55%, depending on the coordination number) [5] which has an effect on the phase stability, the transportation characteristics and the interphase formation [6]. Furthermore, the electrochemical standard electrode potential for sodium ($E0 = -2.71$ V vs. NHE) is higher than that for lithium ($E0 = -3.05$ V vs. NHE) [7]. As a result, the sodium-ion battery is promised a lower energy density [8] but due to the lower potential water-based electrolytes are used [9]. While graphite is applied as a material system for the anode of commercial LIB systems [10], the material is unsuitable for SIB, since the Na+ ions can only be stored under certain circumstances [11]. Hard carbon is the state-of-the-art anode material for SIB [12]. In contrast to sodium, lithium reacts with aluminum and forms alloys. Therefore, the less expensive aluminum can be used instead of copper as current collector for the anode in SIB [13].

2 Electrode Manufacturing and Challenges Along the Process Chain

2.1 Slurry Production

In the process chain for the production of battery electrodes, the process steps of slurry making, coating and drying are of essential importance for the formation of a specifically adjusted quality and the microstructure of the electrode.

To process the electrodes, the particulate components such as active material and conductive additives are first processed together with the solvent and with addition of the binder to form a homogeneous electrode slurry. Due to the mechanical stress of the mixing process, consisting of several successive steps, the existing agglomerates are disagglomerated, crushed and evenly distributed [14]. Different energy inputs and shear effects of different mixer types lead to a change in the distribution of the components in the dispersion medium [15–17]. Furthermore, the type and amount of the individual components in the slurry determine the flow properties and homogeneity of the slurry [18]. This leads to large differences in the rheological paste properties and the morphology as well as the properties of the electrode [19–22].

2.2 Coating and Drying of Battery Electrodes

In the coating step, the electrode slurry is applied to a metallic current collector foil by using an application tool. The rheological properties with regard to the processability of the slurry when using different coating techniques are crucial here. The state-of-the-art is the slot coating process, in which the slurry is applied through a stationary slot nozzle onto the moving foil. The homogeneity of the film application during the coating process must be ensured and coating defects must be avoided [23, 24]. Defects and inhomogeneities in the layer thickness distribution can influence the subsequent mechanical and electrical properties of the electrode. This requires high demands on the manufacturing accuracy of the slot nozzles, control of the flow behavior and stability of the slurry as well as the fluid supply of the coating tool [14].

The subsequent drying step aims to remove the solvent from the coated slurry and to ensure the formation of a porous electrode structure. For this purpose, dryers consisting of several drying zones with different conditions are often used in industry in a continuously operated roll-to-roll process.

The drying process is a complex multi-stage process and has a considerable influence on the formation of the microstructure and the distribution of the individual components [25]. In order to withstand the mechanical stress during cycling, the particles must have sufficient adhesion within the electrode layer and with the collector foil. Otherwise there may be a loss of capacity and cell failure. Depending on the drying rate and the slurry properties, a binder gradient can occur in the electrode, which influences the stability and adhesion on the current collector [26–28].

2.3 Calendering

One of the main objectives of the calendering process is to increase the volumetric energy density by compacting the coating and thus reducing the porosity of the dry electrode to 30 to 35%. The resulting modification of the pore structure has a significant impact on the electrochemical and mechanical properties. The compaction rate is configurable by the applied line load and the gap size between the calender rolls.

The various active material properties demand for individually adjusted process parameters and achieve different resulting electrode properties [29]. Furthermore critical issues of previous process steps as well as difficulties directly correlated to the process parameters often result in defects that influence the future cell performance. Geometrical defects are plastic deformations, usually characterized by periodic waves and are classified according to their appearance and location. Geometric defects result in a decrease of adhesive force. It is assumed that all geometric defects can be reduced by a decrease of line load. This can be enabled by roll tempering. Structural defects like local thickness and density deviations caused by uneven surface loadings lead to inconsistent compaction rates and come along to inhomogeneous porosities. Furthermore sealed surface pores and coating detachment pose challenges to the cell performance. Mechanical defects are mostly detected during electrode handling. The migration and encountering of particles under high pressures may induce the formation of cracks within the coating and the particles themselves. Foil tear and embrittlement are further challenging the production process [14, 30, 31].

3 Manufacturing of Sodium Ion Anodes

The processing of the SIB anode is performed according to known process steps for the LIB and the influence of different process-machine interactions on the properties of the material is investigated. The investigations are carried out with hard carbon and are intended to provide information for predicting the processability of new materials.

3.1 Mixing, Coating and Drying of SIB Anode Electrode

To produce the slurry, a multi-stage mixing process is transferred from the literature for a graphite anode slurry for LIB [21]. The active material is processed into a slurry with a conductive additive and a water-based binder system. The composition of the components for the slurry is also transferred from the literature and hard carbon as main component is used instead of graphite [21, 27]. The hard carbon used is spherical and has an average particle size of 10 µm (D_{50}).

The hard carbon (BHC-240, Shandong Gelon LIB Co., LTD) and the carbon black (C65, Imerys) are prepared in a dry mixing step. For slurry preparation, CMC (Sunrose, Nippon Paper Industries) is dissolved with water and added to the dry mixture in several dilution steps in a kneading device. The solids content in the slurry is adjusted to 43 wt% by adding water. Before the last mixing step, SBR (Zeon Europe GmbH, Japan) was added (Table 1).

Table 1. Slurry composition for the anode of the SIB

Material	Mass fraction solids [wt%]
Hard carbon	93
Carbon black	1.4
CMC	1.87
SBR	3.73

To determine the viscosity a rotational rheometer with a plate-plate measuring system with a diameter of the measuring head of 25 mm was used (Anton Paar, Germany) (Fig. 2). The hard carbon anode slurry shows a very high zero shear viscosity and a strongly shear-thinning viscosity behavior. With increasing shear rate the viscosity decreases accordingly. The viscosity profile shows a good agreement with the literature for the same multi-stage production process of a graphite slurry [21]. Their viscosity profile also shows a strong shear-thinning behaviour with 129 ± 10 Pa s at low shear rate for all slurries and from 0,16 Pa s to 0,26 Pa s very high shear rate.

The production of the electrode is carried out on a roll-to-roll pilot plant (Coatema, Germany) at the TFT - Thin Film Technology Group at the KIT.

The slurry is applied to the collector foil via a dosing unit and a slot die coater. The coating width is 150 mm and the wet film thickness of the slurry on the collector foil is 200 µm. Used current collectors are a copper foil with 10 µm and an aluminum foil with 20 µm. The coating process was carried out free of defects at a constant speed of the collector foil of 0.3 m min^{-1}. After coating, the wet film is dried convectively in two identical dryers connected in series at different temperatures (120 °C and 100 °C). The porosity of the coating on the aluminum foil is 0.21 and 0.29 on the copper foil.

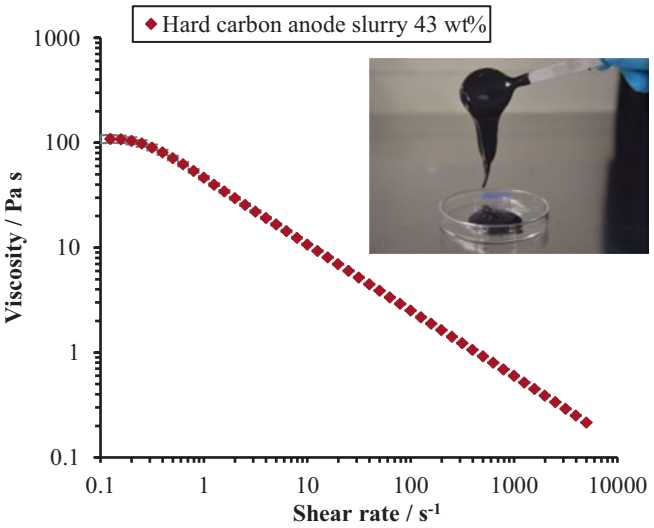


Fig. 2. Average viscosity in a shear rate range from 0.01 to 5000 s^{-1} of the hard carbon anode slurry with 43 wt%.

3.2 Calendering

The investigations show the differences that occur when calendering to different degrees of compaction calculated by thickness and using different current collectors. Different degrees of compaction are aimed at, since there is no application with specified requirements for the electrodes and the aim is to build models to be able to react flexibly to new materials for calendering on an industrial scale. The influence of the machine is reduced in order to better observe the material behaviour. For this purpose, the calendering speed is kept constant at 1 m/min, the roll temperature at 90 °C and the web tension at 50 N for the aluminum and 30 N for the copper foil. The processing to different compaction rates is achieved by the adjustment of the hydraulic cylinders and thus a reduction of the roll gap but also by an increase of the line load. The calender (Saueressig) at the wbk Institute for Production Science at the KIT enables a compaction with a line load of 2000 N/mm.

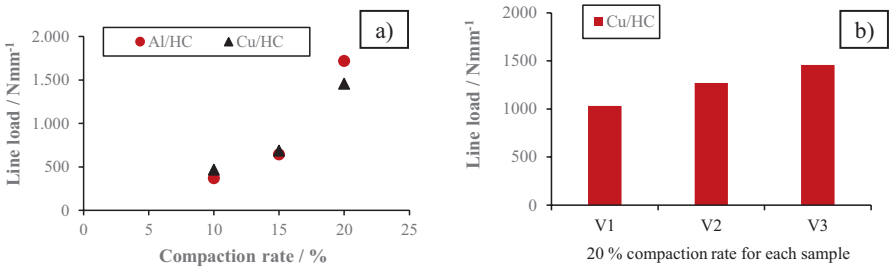


Fig. 3. Comparing line load for calendering hard carbon with different compaction rates and different current collectors

Figure 3a illustrates that the line load increases with a higher degree of compaction as a stronger compaction is necessary. The particles are held together by the binder, so that compaction involves not only moving the particles towards each other but also overcoming the shear forces on the binder. In the course of the line load it can be seen that saturation occurs. This means that the line load continues to increase, but the material reduction is still negligible. If a maximum compaction, in this case of 20%, is thus reached, the particles are all close together. Figure 3b shows this effect. The increase of the line load can no longer cause the particles to move towards each other but leads to the particles breaking. This can be seen in the SEM images of Fig. 4, which compare the both states and in Fig. 5d.

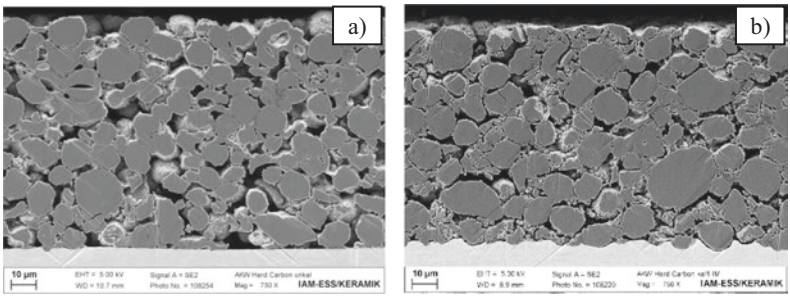


Fig. 4. (a) uncalendered (b) calendered electrode

This was also confirmed by the adhesive force tests shown in Fig. 5, since an increase in the degree of compression until 15% and thus pressing the particles into the collector causes an increase in the adhesive force. For a compaction rate of 20% the adhesive force is decreasing. Thus an increase in the line load leads to a displacement of the weak point from the film into the coating, since broken particles are located there.

The increase of the adhesive force until the compaction rate to 15% is related to the increased penetration of the particles into the foil, which is shown by the photographs Figs. 4b, 5c and d of the partially removed coating. The higher the degree of calendering, the more particles are trapped on the film base, as shown in Fig. 4b. This also reinforces the assumption that the weak point moves into the coating, because of cracking particles.

The SEM images in Fig. 5c and d show that the particles are pressing themselves into the foil, which can be seen just by looking at the surface as well as in the cross-section. This means that the film becomes tapered and distorted due to the compression of the particles.

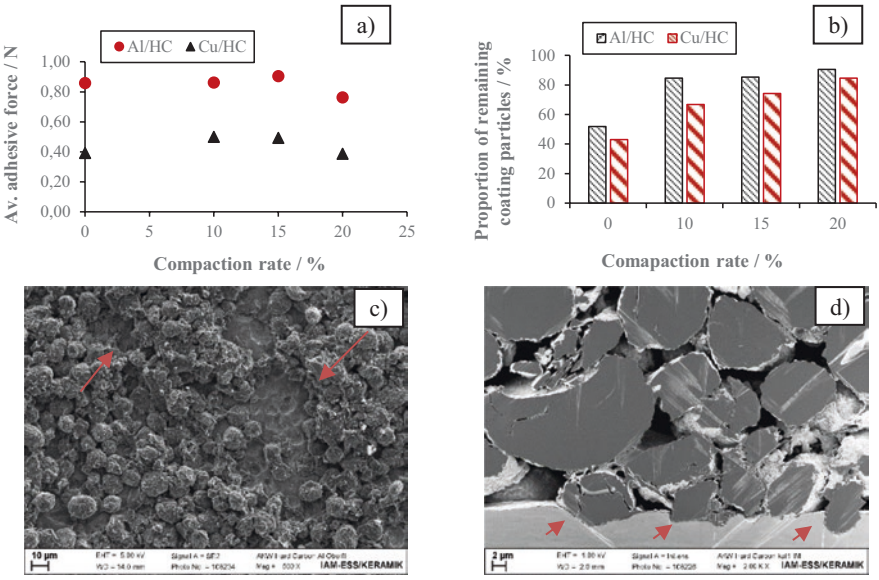


Fig. 5. Behavior of the adhesive force at different compaction rates influenced by particle penetration

The distortion can also be detected on the surface with a stereo camera system (gom ARAMIS), which can compare the changes of an applied point pattern in uncalandered and calendered state. The results for the highest compaction rate are presented in Fig. 6. This shows that the coating as well as the collector is exposed to mechanical stress. The load leads to an elongation of the film below the coating and thus to a difference in the states from the coated to the uncoated area. In addition, the investigations have revealed differences in the use of the collector that is also detected for example in the distortion in Fig. 4a. The adhesive force is lower for copper, and fewer particles remain attached to the collector. The distortion of the electrode is greater with the use of copper foil. This finding needs further investigation, which will be done in future research. Thus, comprehensive results are available which allow for a tendency in the behavior of the electrode due to different degrees of compression and the collector used.

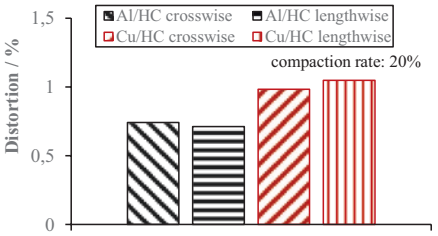


Fig. 6. Distortion for the compaction rate of 20%

4 Conclusion and Outlook

The manufacturing process for a slurry could be transferred from the literature for a graphite anode of LIB to the anode of SIB with hard carbon. For this purpose, a multi-stage mixing process was carried out and the slurry obtained was rheologically examined. A similar viscosity behaviour and range of zero viscosity was shown. This suggests a similar processing as for the anode for LIB. It is also necessary to investigate the influence of different particle properties such as diameter and shape, different compositions and the manufacturing process on the properties of the slurry such as component distribution and stability. In particular, the relationship between the properties of the slurry on the forming microstructure and the mechanical and electrical properties of the electrode must be worked out. This is to create the connection between the process and the optimum properties of the electrode for the desired application.

The anode slurry for the SIB could be coated and dried without defects in a roll-to-roll process in a pilot plant. The behaviour of the slurry during the coating and drying process suggests similar characteristics and process engineering issues as with the graphite anode of the LIB. Once the optimization of the microstructure from the component composition to the process engineering of the electrode for a single layer has been investigated, the optimization of the characteristics of the electrode should be developed with the aid of a multilayer structure. This poses particular challenges for the coating and drying process. In addition, the edge geometry of the coating must be optimized with regard to the calendering of the electrode.

With regard to calendering, the investigations on hard carbon show the expected results for increasing the degree of calendering for a material. As the compaction rate increases, the line load and thus the load on the material also increases. There is an optimum for achieving the required porosity and the necessary process parameters. For example, compaction to a 20% degree of calendering shows that a too high load on the material results in a loss of adhesive strength. Thus, the material shows a typical behaviour with regard to the change in the degree of compaction. Despite the use of the material as anode material, it exhibits properties of an NMC cathode compared to the previous materials, as described in [29], especially with regard to the necessary line load. The only characteristic these materials have in common is the spherical morphology. Therefore, these investigations show that the morphology of the particles plays a decisive role in the modelling of calendering with respect to new materials. Thus, further research activities include the description of the mechanical properties of the particles, such as diameter and hardness. Of course, the influences of the binder have to be considered as well. The knowledge concerning the particles could be used to react to new materials in a simple way with regard to the adjustment of the process parameters, where only these few properties of the particles are known.

5 Acknowledgements

This work contributes to the research performed at CELEST (Center for Electrochemical Energy Storage Ulm Karlsruhe) and KIT Battery Technology Center. It was funded by the German Research Foundation (DFG) under Project ID 390874152 (POLiS Cluster of Excellence).

References

1. Vaalma, C. et al.: A cost and resource analysis of sodium-ion batteries. Nat. Rev. Mater. **3**(4) (2018). https://doi.org/10.1038/natrevmats.2018.13
2. Elia, G.A., et al.: An overview and future perspectives of aluminum batteries. Adv. Mater. Weinh. **28**(35), 7564–7579 (2016)
3. Kulova, T.L., Skundin, A.M.: From lithium-ion to sodium-ion battery. Russ. Chem. Bull. **66**(8), 1329–1335 (2017)
4. Hwang, J.-Y., Myung, S.-T., Sun, Y.-K.: Sodium-ion batteries: present and future. Chem. Soc. Rev. **46**(12), 3529–3614 (2017)
5. Huheey, J., Keiter, E., Keiter, R.: Anorganische Chemie. DE GRUYTER, Berlin (2014)
6. Adelhelm, P., et al.: From lithium to sodium: cell chemistry of room temperature sodium-air and sodium-sulfur batteries. Beilstein J. Nanotechnol. **6**, 1016–1055 (2015)
7. Felixberger, J.K.: Chemie für Einsteiger. Springer, Heidelberg (2017)
8. Pan, H., Hu, Y.-S., Chen, L.: Room-temperature stationary sodium-ion batteries for large-scale electric energy storage. Energy Environ. Sci. **6**(8), 2338 (2013)
9. Palomares, V., et al.: Na-ion batteries, recent advances and present challenges to become low cost energy storage systems. Energy Environ. Sci. **5**(3), 5884 (2012)
10. Eftekhari, A., Kim, D.-W.: Sodium-ion batteries: New opportunities beyond energy storage by lithium. J. Power Sources **395**, 336–348 (2018)
11. Nayak, P.K., et al.: From lithium-ion to sodium-ion batteries: advantages, challenges, and surprises. Angew. Chem. Int. Ed. Engl. **57**(1), 102–120 (2018)
12. Zhang, W., et al.: Sodium-ion battery anodes: status and future trends. EnergyChem **1**(2), 100012 (2019)
13. El Kharbachi, A., et al.: Exploits, advances and challenges benefiting beyond Li-ion battery technologies. J. Alloy. Compd. **817**, 153261 (2020)
14. Kaiser, J., et al.: Prozess- und Produktentwicklung von Elektroden für Li-Ionen-Zellen. Chem. Ing. Tech. **86**(5), 695–706 (2014)
15. Bauer, W., et al.: Influence of dry mixing and distribution of conductive additives in cathodes for lithium ion batteries. J. Power Sources **288**, 359–367 (2015)
16. Bockholt, H., Haselrieder, W., Kwade, A.: Intensive powder mixing for dry dispersing of carbon black and its relevance for lithium-ion battery cathodes. Powder Technol. **297**, 266–274 (2016)
17. Lee, G.-W., et al.: Effect of slurry preparation process on electro-chemical performances of $LiCoO_2$ composite electrode. J. Power Sources **195**(18), 6049–6054 (2010)
18. Bitsch, B., et al.: Einflüsse der mechanischen Verfahrenstechnik auf die Herstellung von Elektroden für Lithium-Ionen-Batterien. Chem. Ing. Tech. **87**(4), 466–474 (2015)
19. Su, F.-Y., et al.: Micro-structure evolution and control of lithium-ion battery electrode laminate. J. Energy Storage **14**, 82–93 (2017)
20. Bockholt, H., et al.: The interaction of consecutive process steps in the manufacturing of lithium-ion battery electrodes with regard to structural and electrochemical properties. J. Power Sources **325**, 140–151 (2016)
21. Diehm, R., et al.: In-situ investigations of simultaneous two-layer slot die coating of component graded anodes for improved high energy li-ion batteries. Energy Technol. **8**(5), (2020)
22. Wenzel, V., Nirschl, H., Nötzel, D.: Challenges in lithium-ion-battery slurry preparation and potential of modifying electrode structures by different mixing processes. Energy Technol. **3**(7), 692–698 (2015)

23. Schmitt, M., et al.: Slot-die processing of lithium-ion battery electrodes—coating window characterization. Chem. Eng. Process. **68**, 32–37 (2013)
24. Ding, X., Liu, J., Harris, T.: A review of the operating limits in slot die coating processes. AIChE J. **62**(7), 2508–2524 (2016)
25. Jaiser, S., et al.: Development of a three-stage drying pro-file based on characteristic drying stages for lithium-ion battery anodes. Drying Technol. **35**(10), 1266–1275 (2017)
26. Jaiser, S., et al.: Investigation of film solidification and binder migration during drying of Li-Ion battery anodes. J. Power Sources **318**, 210–219 (2016)
27. Kumberg, J., et al.: Drying of lithium-ion battery anodes for use in high-energy cells: influence of electrode thickness on drying time, adhesion, and crack formation. Energy Technol. **7**(11), 1900722 (2019)
28. Müller, M., et al.: Investigation of binder distribution in graphite anodes for lithium-ion batteries. J. Power Sources **340**, 1–5 (2017)
29. Meyer, C., et al.: Characterization of the calendering process for compaction of electrodes for lithium-ion batteries. J. Mater. Process. Technol. **249**, 172–178 (2017)
30. Bold, B., Fleischer, J.: Kalandrieren von Elektroden für Li-Ionen-Batterien. ZWF **113**(9), 571–575 (2018)
31. Günther, T., et al.: Classification of calendering-induced electrode defects and their influence on subsequent processes of lithium-ion battery production. Energy Technol. **8**(2), 1900026 (2020)

Numerical Investigation of an Extruded Shaft for High Temperature Applications Manufactured by Tailored Forming

C. Büdenbender[1(✉)], I. Ross[1], H. Wester[1], A. Zaitsev[2], and B. A. Behrens[1]

[1] Institut für Umformtechnik und Umformmaschinen, Leibniz Universität Hannover, Hannover, Germany
buedenbender@ifum.uni-hannover.de

[2] Institute of Metallurgy, Mechanical Engineering and Transport, Peter the Great St.-Petersburg Polytechnic University, St. Petersburg, Russia

Abstract. Components used in high temperature applications require special material properties to be able to withstand these external conditions. However, oftentimes only individual component areas are exposed to such requirements. Multi-material solutions facilitate the use of the right material at the right place, thus saving resources and costs. One approach, considered in the CRC "Tailored Forming", is the use of pre-joined hybrid semi-finished products to manufacture high-performance multi material components. Within the scope of this study, the forming process for the production of a hybrid shaft made of the nickel-based alloys AISI alloy 625 and AISI 304 is numerically investigated. A material characterisation was carried out to analyse the different thermomechanical properties of the materials and to define a suitable process window in which the flow properties are adjusted. Furthermore, the influence of various die angles on the joining zone was investigated. A tool load analysis was finally carried out.

Keywords: Material characterisation · High temperature application · Tailored forming

1 Introduction

Bulk metal forming components are often used in areas exposed to high temperatures and loads. They are applied in the automotive and aviation as well as electrical and chemical industry. The nickel-based alloy AISI alloy 625 has been widely used in high temperature applications such as aerospace, petrochemical, marine and nuclear industries because of its excellent oxidation resistance and superior mechanical properties [1]. This alloy has been developed for service temperatures above 600 °C, and it possesses high strength and excellent fabrication characteristics [2]. The AISI alloy 625 contains relatively high levels of chromium, molybdenum, carbon and

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 182–192, 2021.
https://doi.org/10.1007/978-3-662-62138-7_19

niobium [3]. The material AISI 304 is an austenitic chromium-nickel. It also offers good mechanical properties as well as corrosion resistance and can be used at temperatures up to 300 °C. In order to save resources, the collaborative research centre (CRC) 1153 at the Leibniz Universität Hannover develops a new process chain with the objective of manufacturing novel hybrid components based on the use of pre-joined hybrid semi-finished products [4]. This innovative concept, called tailored forming, is based on the use of pre-joined hybrid semi-finished products, which allows a thermo-mechanical treatment of the joining zone during the forming process to improve their properties. For the joining of the hybrid semi-finished products, different processes like friction welding or build-up welding are investigated [6]. The influence of forming is investigated by means of various processes like die forging, cross-wedge rolling or impact extrusion are designed to manufacture high performance multi-material components [5].

The performance of the hybrid components was proven in [7] by means of tensile specimens taken from the joining zone of a formed serial hybrid semi–finished product made of a steel-steel combination. Since, the failure of the components did not occur in the joining zone, the strength of the joining zone appears to be at least as high as that of the used base materials.

In the present work, an extruded shaft made by the tailored forming process chain using a hybrid semi-finished product consisting of AISI alloy 625 and AISI 304 in serial configuration is presented. The intention is to use the AISI alloy 605 only in those areas of the component that are exposed to high thermo-mechanical loads and to use the more cost-effective AISI 304 in the other areas. In this way, a component with locally adapted properties can be provided and resources and costs can be saved. Furthermore, the forming process of hybrid semi-finished products is a complex process close to the limit of feasibility. By applying the complex process chain of tailored forming, it is possible to produce hybrid components in large quantities, which considerably reduces time and costs. At the same time, the transferability of the CRC process chain to high-alloy materials can be investigated. First, a material characterisation was carried out. In the past decade these two materials have been studied extensively using hot torsion or hot compression experiments [8]. The results show that a dynamic recrystallisation (DRX), which is one of the main softening mechanisms at high temperature, takes place after reaching a critical strain φ_p [9]. When the softening process is governed by dynamic recrystallisation, the flow stress passes through a single peak or shows a cyclic behaviour and subsequently drops to a steady state regime [10, 11]. All the presented works [9–11] used a temperature range between 800 °C and 1200 °C, but only strain rates below 1 $s-^{1}$ or high plastic strains up to 2 were considered. Yet, in the present study the flow behaviour at a strain rate up to 100 s^{-1} will be investigated by means of a hot compression test. The results will be used to compute the coefficients of the Hensel-Spittel-10 flow curve approach. Subsequently, an impact extrusion forming process with three different angles of the forming tool is considered in order to achieve the highest possible surface enlargement of the joining zone. This is necessary to get a sufficient thermo-mechanical treatment of the joining zone to enhance the performance of the component [7]. Finally, a stress analysis of the forming tools is presented.

2 Experimental Setup of the Material Characterisation

Uniaxial compression tests were carried out to investigate the flow behaviour of the two steel grades AISI alloy 625 and 304 to identify the temperature range at which the flow properties are similar. Cylindrical specimens with a diameter of 10 mm and a height of 15 mm were used. The upsetting tests were carried out with the forming simulator Gleeble 3800-GTC (cf. Fig. 1). Investigated temperatures ranged between 800 °C and 1200 °C with steps of 100 °C. The specimens were heated conductively in an evacuated test chamber. Since the flow behaviour depends on the temperature and the strain rate, four different strain rates were also investigated. The examined strain rates $\dot{\varphi}$ were 1.5 s^{-1}, 10 s^{-1}, 50 s^{-1} and 100 s^{-1}. The specimens were upset to a plastic strain of $\varphi = 0.7$. Up to this point, the influence of friction on the measurement result is negligibly small [12]. Due to the high investigated strain rates, an isothermal compression test cannot be guaranteed. The heating due to plastic dissipation of the specimen is considered in the calculation of the flow curves. The resulting test data will subsequently be used to calculate the coefficients of the Hensel-Spittel-10 flow curve approach, which will be used in the numerical model. The analytical flow curve approach is shown in Eq. (1). The resulting flow stress is represented by σf. The coefficients A, m_1, m_2, m_3, m_4, m_5, m_6 and m_7 were computed with the GRG nonlinear optimisation algorithm. Analytical flow curve approaches like the Hensel-Spittel-10 approach describe the flow behaviour in a defined area of strain, temperature and strain rate. Due to the use of the flow curve approach, it is unnecessary to investigate the entire experimental matrix. The executed tests are marked in the experimental matrix in Fig. 1. The tests are sufficient to compute the coefficients of the flow curve approach presented and the simulation of the uniaxial compression will additionally validate the coefficients. The resulting flow curves, the coefficients of the flow curve approach and the validation of the coefficients are discussed in Sect. 4.

$$\sigma_f = A \cdot e^{m_1 \cdot T} \cdot T^{m_8} \cdot \varphi^{m_2} \cdot e^{(m_4 \varphi + m_6 \varphi)} \cdot (1+\varphi)^{(m_5 \cdot T)} \cdot \varphi^{(m_3 + m_7 \cdot T)} \tag{1}$$

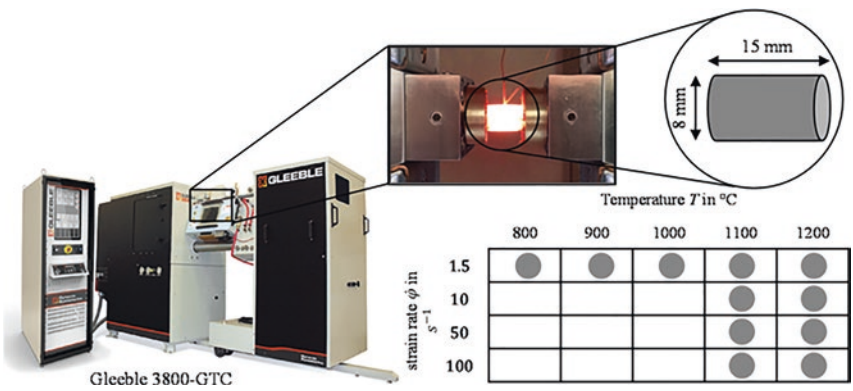


Fig. 1. Gleeble 3800-GTC with a view into the test chamber and the presentation of the specimen used and the experimental test matrix

3 Modelling of the Numerical Forming Process

In Fig. 6(a), the numerical model of the forging process, which was set up with simufact.forming v16, is shown. The impact extrusion tool system consists of a punch, a die and the serially arranged semi-finished products with the AISI alloy 625 on top made by friction welding. An axisymmetric numerical model was created. The diameter of the semi-finish product is 40 mm and will be reduced by the full forward impact extrusion process to 30 mm. The height of the semi-finished product is 100 mm, where the proportion of AISI alloy 625 is 60 mm. The proportion of the AISI 304 is 40 mm. An elastic-plastic material model was used to represent the material behaviour. Quadratic elements of the type quads (10) were chosen to discretise the semi-finished products. The initial temperature of the punch and die was 250 °C. The punch and the die were modelled as deformable bodies to calculate heat conduction and mechanical stress within the tools. The joining zone of the semi-finished product was modelled by an adhesive contact. To map the friction between semi-finished product and the punch and the die, the combined friction model was used. The combined friction model consists of the tresca-formulation with $m = 0.3$ and the coulomb-formulation with $\mu = 0.1$. Three dies with different angles α (25°, 30° and 35°) were investigated to analyse the effect on the joining zone. The die consisted of the material AISI L6. For the stress analysis, the material data was taken from the simufact.forming database.

4 Results

Results of the material characterisation

In Fig. 2, the resulting flow curves at a strain rate of $\dot{\varphi} = 1.5$ and various temperatures are presented; Fig. 2(a) considers AISI alloy 625 and Fig. 2(b) AISI 304. The general characteristics of the flow curves of the alloy 625 are similar in hardening at low plastic strains and subsequently softening at various temperature conditions.

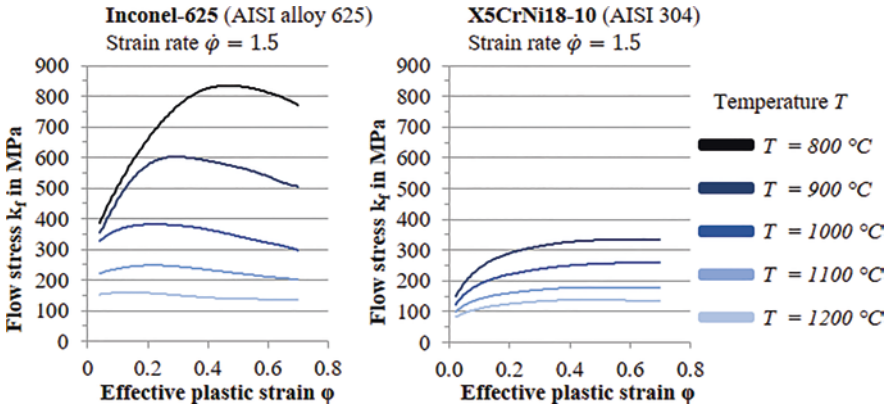


Fig. 2. (a) Resulting stress strain curves of alloy 625 and (b) AISI 304 at a strain rate of 1.5 s^{-1} and test temperatures of 800 °C, 900 °C, 1000 °C, 1100 °C and 1200 °C

However, the hardening behaviour decreases with increasing temperature. In particular, at the lower investigated temperatures, the material AISI 304 shows a significantly lower yield stress level. At a temperature of 800 °C, the AISI alloy 625 reaches the peak flow stress after a critical plastic strain of $\varphi_p = 0.4$. Due to the DRX with a single peak behaviour, a softening process is subsequently observed. With increasing temperature, the critical plastic strain φ_p decreases. Accordingly, the test temperature strongly influences both the level of flow stress and the critical plastic strain φ_p at which the DRX takes place [1]. These results were also observed by other authors [13–15]. In contrast to the AISI alloy 625, the AISI 304 steel shows significantly lower softening behaviour. At the investigation settings of $T = 900$ °C and $\dot{\varphi} = 1.5\ s^{-1}$, the effect of DRX cannot be observed clearly. Generally, with a decrease of the temperature and increases of the strain rate, the peak of the flow stress curve is less obvious, which indicates that the critical plastic strain φ_p is not reached yet [17]. As shown by Dehghan-Manshadi et al., at test conditions of 900 °C and $\dot{\varphi} = 1\ s^{-1}$, the critical plastic strain φ_p is reached below $\varphi = 1$ [16]. Meanwhile, Kim and Yoo prove that a critical strain φ_p of 1 is detected at a test temperature of 1000 °C at a strain rate of $\dot{\varphi} = 5\ s^{-1}$ [17]. This indicates that the critical plastic strain increases with increasing temperature and strain rate. In [18], at test conditions of 1150 °C and strain rates in the range of $0.1\ s^{-1}$ and $0.001\ s^{-1}$, a clear critical plastic strain of $\varphi_p = 0.3$ was detected. So the effect of DRX becomes clearly visible with a plastic strain below $\varphi = 1$. Summarising these studies' results, an obvious DRX cannot be detected for the described test conditions between 900 °C and 1200 °C at $\dot{\varphi} = 1.5$ and above for the AISI alloy 304. Furthermore, Fig. 2 shows that the level of flow stress of AISI alloy 625 at 1200 °C is nearly equal to the flow stress level of AISI 304 at 1100 °C. This does not apply at higher strain rates, though, as shown in Fig. 3. The resulting flow curves at a temperature of 1100 °C and the strain rates $10\ s^{-1}$, $50\ s^{-1}$ and $100\ s^{-1}$ are depicted in Fig. 3. For AISI alloy 625 (cf. Fig. 3(a)), the flow stress increases with increasing strain rate. An explicit DRX can be obtained for the strain rates $50\ s^{-1}$ and $100\ s^{-1}$.

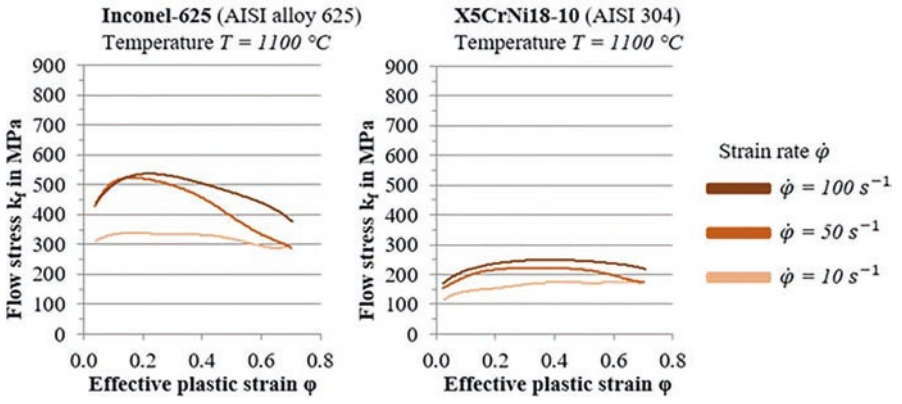


Fig. 3. (a) Resulting stress strain curves of alloy 625 and (b) AISI 304 at strain rates of $10\ s^{-1}$, $50\ s^{-1}$ and $100\ s^{-1}$ and a temperature of 1100°C

For the tested strain rate of 10 s^{-1}, the effect of DRX decreases significantly. This can also be seen in [15], where the same test conditions were used. Again, no significant DRX can be obtained for AISI 304 (cf. Fig. 3(b)). However, the level of flow stress increases with an increase of the strain rate. Furthermore, for all tested strain rates, the level of flow stress is lower compared to AISI alloy 625. In Fig. 4(a), flow curves at 1200 °C and various strain rates are shown. The level of flow stress decreases due to the higher test temperature of 1200 °C for the alloy 304.

Again, the level of flow stress increases when the strain rate rises. The results of the AISI alloy 625 cannot be used at a test temperature of 1200 °C. The specimens were broken after the upsetting test. It is assumed that the heat dissipation due to plastic work caused the temperature of the material to rise to a temperature close to the melting point, which is about 1350 °C, which was observed for all tested strain rates above 10 s^{-1}.

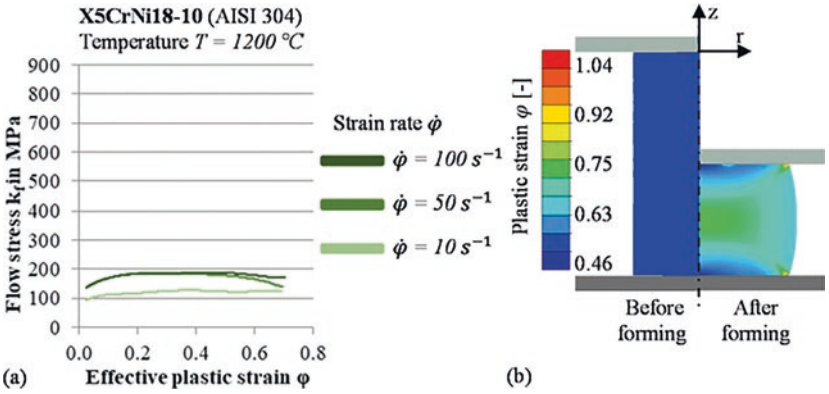


Fig. 4. (a) Resulting stress strain curves of alloy 304 at strain rates of 10 s^{-1}, 50 s^{-1} and 100 s^{-1} and a temperature of 1200 °C; (b) numerical model for the validation of the flow curve approach

The presented data of the flow stress curves was used to compute the coefficients of the Hensel-Spittel-10 flow curve approach. Due to the strong temperature dependency of the alloy 625, the coefficients were computed for two temperature areas (AISI 625 I from 900 °C to 980 °C and AISI 625 II from 980 °C to 1200 °C). The resulting coefficients are presented in Table 1.

Table 1. Computed coefficients for the Hensel-Spittel-10 flow curve approach

AISI	A	m_1	m_2	m_3	m_4	m_5	m_6	m_7	m_8
304	29.1	−0.00339	0.4275	0.0127	0.005064	−0.002	0.27	0.00012	0.87
625 I	12.5	−0.0046	0.619	0.095	0.00046	−0.007	3.13	9.91e−5	1.39
625 II	69.9	−0.0061	0.706	−4.308	0.00577	−0.007	3.07	0.0041	1.4

The flow curve approach was validated by the simulation of the uniaxial compression test. The numerical model is presented in Fig. 4(b). Figure 5 shows the comparison of the force-displacement curves. Figure 5(a) depicts the results of the AISI alloy 625 for the strain rate 1.5 s^{-1} and test temperatures of 900 °C and 1100 °C. In Fig. 5(b), the results of the AISI 304 are presented. The numerically calculated force-displacement curves are in good agreement with the experimentally measured curves. Thus, the Hensel-Spittel-10 flow curve approach is validated.

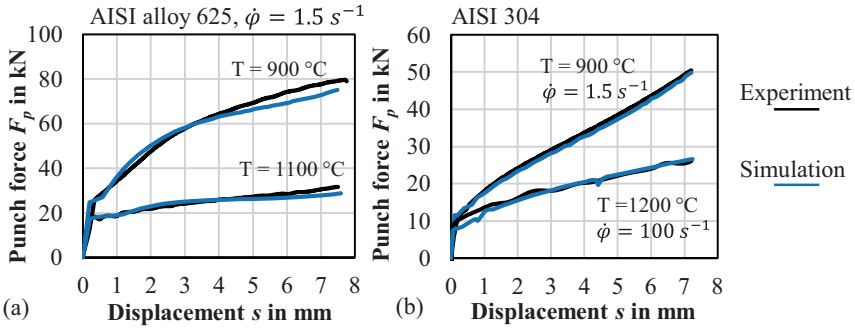


Fig. 5. Comparison of the force-displacement curves of the upsetting test for varying test conditions, (a) for the AISI alloy 625 and (b) for AISI 304

Results of the material flow investigation

The results of the material characterisation indicate nearly similar flow behaviour for a temperature of 1200 °C for the AISI alloy 625 and 1100 °C for the AISI 304 at a strain rate of 1.5 s^{-1}. Thus, these temperatures were chosen for the numerical process design. The implementation of inhomogeneous heating concepts using induction has already been demonstrated for other material combinations in the CRC in experimental tests. Within the parts of the serially arranged semi-finished product, an initially homogenous temperature distribution was assumed for a first numerical process approach. Figure 6(b) shows the results of the forming process after the alloy 625 has passed the forming zone of the die. With increasing angles of the forming die the resulting plastic strain increases as well. It was observed that with rising die angle the alloy 625 displaces the alloy 304 more and more. Due to that, a larger surface enlargement in the area of the joining zone was achieved. Through the thermo-mechanical treatment of the joining zone, higher bond strengths of the formed component can be achieved [19].

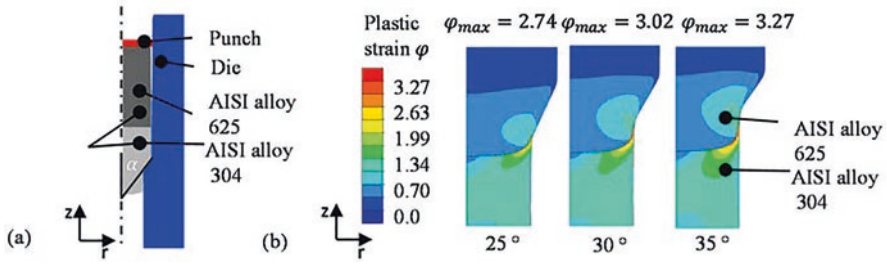


Fig. 6. (a) Numerical model of the forging process with the die angle α; (b) distribution of plastic strain for the three die angles

In Fig. 7, the numerical force-displacement curves of the impact extrusion process with a hybrid semi-finished product are shown. In part A, the steady-state area of the impact extrusion process is achieved, which is typical for this process, if a mono semi-finished product is used. In area B, the force increases because the alloy 625 flows in radial direction due to the punch displacement. The frictional force between tool and work piece increases the force requirement. In the last section of the diagram, named C, the force increases again with the next part of the semi-finished product situated in the forming zone. Since the flow stress of the alloy 625 is higher than the flow stress of alloy 304, the force requirement increases again. Another conclusion which can be drawn based on the force-displacement curve is that the die angle has a small influence on the force requirement. The maximum required force is nearly on the same level for all three different used dies.

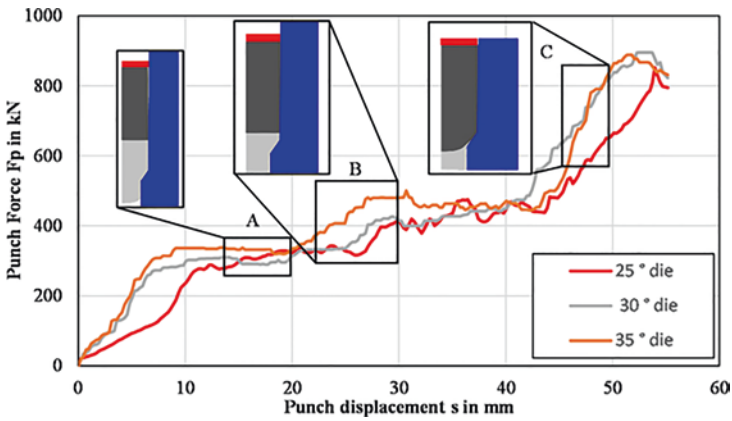


Fig. 7. Force-displacement of the impact extrusion process with a hybrid semi-finished product

Tool analysis

In Fig. 8, the von Mises stress distribution of the forming die is presented. The results show the state with the highest recognized punch force (cf. Fig. 7) which was applied on the forming die. Since, the maximum required force for the impact extrusion process for the forming process under consideration with a hybrid semi-finished product is nearly on the same level for all three used dies, there is also no significantly difference to be observed during tool loading. The maximum of the observed von Mises stress is about 1100 MPa. This load can be taken easily by the chosen tool material and indicates that no plastic deformation of the die occur with the selected die geometries. A reinforcement of the tool is not required for presented setting of the impact extrusion process.

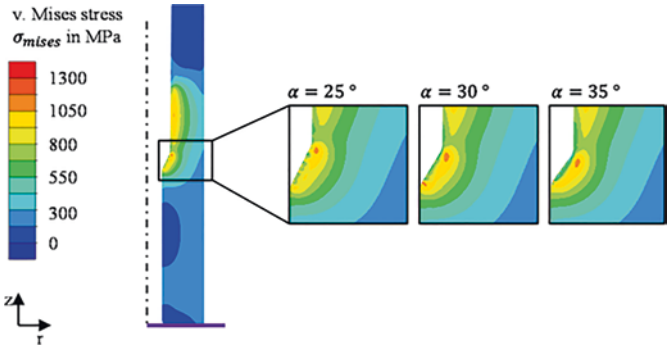


Fig. 8. Calculated von Mises stress distribution of the forming die at the stage of the highest punch load for three different die angles

5 Summary and Outlook

The development of hybrid bulk metal components could lead to a new level of high performance products. The novel process chain tailored forming uses hybrid semi-finished products for the forming process. To enable the simultaneous forming of different foreign materials, knowledge of the temperature-dependent flow properties is of great importance. Therefore, the strain, temperature and strain rate dependent flow behaviour of the alloys 625 and 304 was investigated by means of upsetting tests. The results were used to compute the coefficients of the Hensel-Spittel-10 flow curve approach. Based on a material characterisation, a suitable temperature concept was chosen. Subsequently, a numerical investigation of the impact extrusion process was conducted to identify a die angle that achieves a big surface enlargement. After that, a tool analysis was carried out. All tested dies show almost the same forming force, therefore no significant difference can be observed when the dies are loaded by means of the resulting force. The observed loading of the die can be taken easily be the chosen tool material.

However, the die with an angle of 35° seems to be the optimal angle of the tested configurations for the impact extrusion process, because in this case the highest surface enlargement of the joining zone was achieved. The presented numerical simulations will be used to design the experimental test setup. Subsequently, the achieved strength of the joining zone will be evaluated by experimental tensile test. Furthermore, numerical models will be validated by experimental tests in further studies.

Acknowledgements. This paper's results were obtained at the Collaborative Research Centre 1153 "Process chain to produce hybrid high-performance components by Tailored Forming" in subproject C1. The authors would like to thank the German Research Foundation (DFG/252662854) for the financial and organisational support. In addition, the authors would like to thank subproject B3 for the provision of design of the impact extrusion tools.

References

1. Shankar, V., Valsan, M., Rao, K.B.S., et al.: Effects of temperature and strain rate on tensile properties and activation energy for dynamic strain aging in alloy 625. Metall. Mater. Trans. A **35**, 3129–3139 (2004)
2. Dokme, F., Kulekci, M., Esme, U.: Microstructural and mechanical characterization of dissimilar metal welding of Inconel 625 and AISI 316L. Metals **8**, 797 (2018)
3. Mittra, J., Dubey, J.S., Banerjee, S.: Acoustic emission technique used for detecting early stages of precipitation during aging of Inconel 625. Scripta Mater. **49**, 1209–1214 (2003)
4. Behrens, B.A., Bouguecha, A., Bonk, C., et al.: FE-based design of a forging tool system for a hybrid bevel gear. Key Eng. Mater. **742**, 544–551 (2017)
5. Behrens, B.A., Chugreev, A., Matthias, T., et al.: Manufacturing and evaluation of multi-material axial-bearing washers by tailored forming. Metals **9**, 232 (2019)
6. Pape, F., Coors, T., Barroi, A., et al.: Tribological study on tailored-formed axial bearing washers. Tribol. Online **13**, 320–326 (2018)
7. Behrens, B.A.; Chugreev, A.; Matthias, T.: Characterisation of the joining zone of serially arranged hybrid semi-finished components. In: AIP Conference Proceedings, Palermo, Italy (2018)
8. Cai, D., Xiong, L., Liu, W., et al.: Characterization of hot deformation behaviour of a Ni-base superalloy using processing map. Mater. Des. **30**, 921–925 (2009)
9. Rollett, A.D., Rohrer, G.S.: Recrystallization and Related Annealing Phenomena. Elsevier, Amsterdam (2017)
10. Trigg, G.L. (ed.): Encyclopedia of Applied Physics. VCH, New York (1991)
11. Sakai, T., Jonas, J.J.: Dynamic recrystallisation: mechanical and microstructural considerations. Acta Metall. **32**, 189–209 (1984)
12. Pöhlandt, K.: Werkstoffprüfung für die Umformtechnik – Grundlagen, Anwendungen, WFT Werkstoff-Forschung und -Technik. Springer, Berlin (1986)
13. Li, D., Guo, Q., Guo, S., et al.: The microstructure evolution and nucleation mechanisms of dynamic recrystallization in hot-deformed Inconel 625 superalloy. Mater. Des. **32**, 696–705 (2011)
14. Oliveira, M.M., Couto, A.A., Almeida, G.F.C., et al.: Mechanical behavior of Inconel 625 at elevated temperatures. Metals **9**, 301 (2018)
15. Guo, Q.M., Li, D.F., Guo, S.L.: Microstructural models of dynamic recrystallization in hot-deformed Inconel 625 superalloy. Mater. Manuf. Process. **27**, 990–995 (2012)

16. Dehghan-Manshadi, A., Barnett, M.R., Hodgson, P.D.: Recrystallization in AISI 304 austenitic stainless steel during and after hot deformation. Mater. Sci. Eng. A **485**(1–2), 664–672 (2008)
17. Kim, S.-I., Yoo, Y.-C.: Dynamic recrystallisation behaviour of AISI 304 stainless steel. Mater. Sci. Eng. A **311**(1–2), 108–113 (2001)
18. El Wahabi, M., Cabrera, J.M., Prado, J.M.: Hot working of two AISI 304 steels: a comparative study. Mater. Sci. Eng. A **343**(1–2), 116–125 (2003)
19. Behrens, B.A., Chugreev, A., Matthias, T., et al.: Investigation of the composite strength of hybrid steel-steel semi-finished products manufactured by laser beam welding and friction welding. IOP Conf. Ser. Mater. Sci. Eng. **461**, 12049 (2018)

Introduction of Composite Hot Extrusion with Tubular Reinforcements for Subsequent Cold Forging

P. Kotzyba[1(✉)], K. C. Grötzinger[2], O. Hering[1], M. Liewald[2], and A. E. Tekkaya[1]

[1] Institute of Forming Technology and Lightweight Components, TU Dortmund University, Baroper Str. 303, 44227 Dortmund, Germany
Patrick.Kotzyba@iul.tu-dortmund.de
[2] Institute for Metal Forming Technology, University of Stuttgart, Holzgartenstr. 17, 70174 Stuttgart, Germany

Abstract. A new process is presented in which semi-finished products (AA6060) with tubular internal reinforcements of higher strength (AA7075) are produced by composite hot extrusion. Flange-shaped parts are manufactured from the extrudate by subsequent lateral extrusion and upsetting. The aim is to develop a robust process route for flange-shaped components by composite hot extrusion and subsequent cold forging. For this purpose, numerical investigations were carried out. The process can be conducted without errors if certain limits are observed, which can be shown numerically. The probability of solid state bonding can be increased by choosing the proper tube geometry, so that subsequent lateral extrusion can be carried out. The knowledge generated is to be used to expand the spectrum of components in the future to include higher strength tubular reinforcements.

Keywords: Composite hot extrusion · Cold forging · Solid bonding

1 Introduction

When aluminum is combined with a reinforcing element, two advantages can be created. Firstly, replacing a heavy material such as steel with the lightweight material aluminum reduces the component weight. On the other hand, it is possible to generate a locally load-adapted profile. These advantages are the reason hybrid components are in great demand and are part of current research.

Composite hot extrusion is described as a process in which a profile is produced from two or more materials. The basic idea of hot extrusion in general is to press a billet preheated to forming temperature in a recipient through a forming tool. Characteristic of this process is the high degree of deformation of the workpiece, as well as the large variety of possible profile geometries, which allow solid and hollow profiles [1].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 193–201, 2021.
https://doi.org/10.1007/978-3-662-62138-7_20

The process of composite hot extrusion is divided into two variants. Either the composite partner is added during the process or the billet is already available in reinforced form [2]. As an example, it is possible to insert isolated electrical conductors [3] or shape memory alloys [4] into aluminum profiles. Reinforced billets can be separated in particle reinforced billets [5] or bi-metallic billets [6].

Composite extrusion of tubular reinforcing elements is a novel process and object of this paper. Previous investigations showed a failure of the steel reinforcing element as shown in Fig. 1, therefore this material should be replaced by the softer aluminum AA7075.

Fig. 1. Failure at the co-extrusion of aluminum with a steel tube

The overall aim is to achieve a stable process route of hot extrusion with subsequent lateral extrusion as shown in Fig. 2. The key point of feasibility is the achievement of a material bond that does not detach during the subsequent extrusion process.

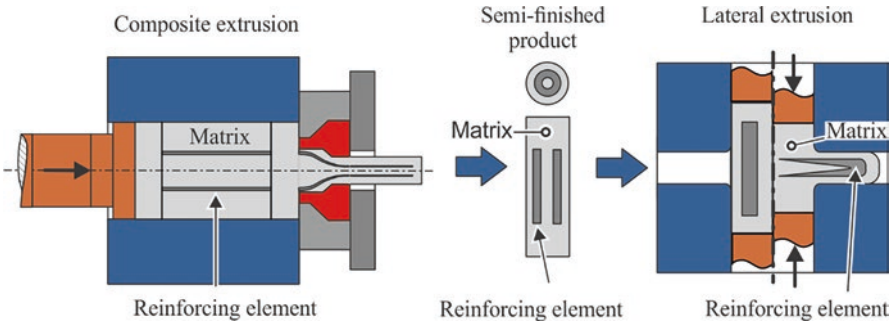


Fig. 2. Process route of co-extrusion with tubular reinforcements and subsequent lateral extrusion

2 Composite Hot Extrusion with Tubular Reinforcements

A prepared extrusion billet consisting of an aluminum 6060 matrix and a tubular reinforcing element made of 7075 is extruded to form a composite profile. As shown in Fig. 3, the tube is inserted in the center of a cylindrical billet, and the composite is then extruded together. Technically, an AA6060 core is inserted into the AA7075 tube, both together are then inserted into the AA6060 sheath. First experimental

investigations show that an error-free process is possible. This can be seen in the microscopic image in Fig. 3. To obtain a closed billet, short AA6060 rods are welded to the ends. The diameter of the aluminum billet and the diameter and wall thickness of the reinforcing tube are to be differentiated.

After extrusion, the steel tube showed local necking and fracture. This failure is caused by the local differences in velocity, therefore the aluminum core is leading. This causes a high hydrostatic stress, the aluminum is torn apart and the steel tube collapses. The difference in strength between matrix and reinforcing element proved to be too high and is replaced in this paper by a lower ratio. The matrix still consists of AA6060, the reinforcement is replaced by AA7075.

Experimental investigations have shown that the newly selected composite enables the process to be carried out without failures. Figure 3 also shows a light microscope image showing the embedding of the reinforcing element in the matrix.

To ensure that the subsequent step, lateral extrusion, is also error-free, it is necessary to weld the composite partners together.

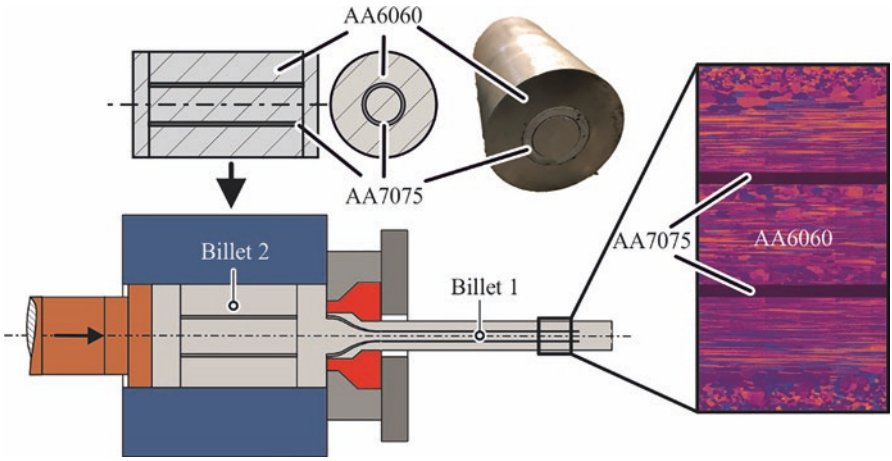


Fig. 3. Process route of composite hot extrusion with tubular reinforcements

According to the theories of Bay [7], a material bond is created when aluminum is joined by two main factors: pressure and strain. When looking at the pressure distribution (above) and strain distribution (below) in Fig. 4, it is noticeable that they are in opposite directions.

The highest pressures are generated on the left-hand side of the picture near the ram. There, however, the strains are very low, which only increases in the direction of forming in front of the die. This is where the pressure drops. This condition leads to a narrow process window in which solid state bonding is theoretically possible, but it can be seen that both pressure and strain tend to increase towards the outside. This is the first finding that should be taken into account when geometrically designing the tube inside the billet. The higher the diameter of the reinforcing tube, the more likely it is that solid state bonding will occur.

The knowledge about pressure and strain must be considered together so that a more reliable statement about the weld quality can be made.

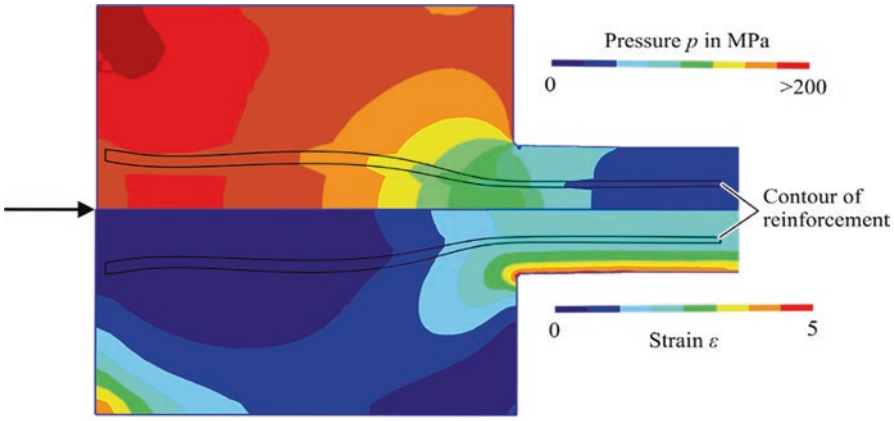


Fig. 4. Pressure and Strain distribution in hot extrusion

3 Analysis of Weld Quality

To predict bond strength between metals Bay [7] developed a physical model and validated it with roll bonding. For metal to bond to metal, the surfaces must come into direct contact. Since aluminum forms an oxide layer on contact with oxygen, these layers must be broken up to achieve direct contact between the metal surfaces. An expansion of the surfaces is required to break up these layers. If the contact pressure is sufficiently high, a bond is achieved, as the metal flows through the cracks in the broken layers.

Cooper and Allwood [8] conducted this model and predicted the bond strength by

$$\frac{\sigma_b}{\sigma_f} = \left(\frac{0.8}{\sigma_f}\sqrt{\sigma_n^2 + 3\tau^2}\right) v\left(0.8\frac{\sigma_n - p_{ex}}{\sigma_f}\right), \tag{1}$$

the quotient $0 < \frac{\sigma_b}{\sigma_f} < 1$ represents the weld quality, is the absolute bond strength and σ_f is the flow stress. σ_n is the contact normal stress, τ is the shear stress and p_{ex} is the micro extrusion pressure, which has to be exceeded by σ_n in order to push aluminum through the cracks in the oxide layers. v stands for the strain which is at least necessary to generate those cracks. Although the investigated parameter ranges do not correspond to the elevated temperatures of hot extrusion, Kolpak et al. [9] found out that the model is capable of predicting the influence of the hot extrusion parameters on the resulting weld quality.

To estimate the weld quality of different tube diameters a finite element simulation in DEFORM 2D with 8,000 elements is used. Due to strong distortions in the area of the forming zone mesh windows were used, which is reduces the mesh in two steps from 2 mm to 0.5 mm element edge length.

The shear friction model was used with m = 1 at the front page of the die and with m = 0.7 at the press channel. The values of the heat transfer conditions can be seen in Table 1. The Cooper and Allwood model was implemented in DEFORM and used to evaluate the weld quality between the matrix material and the reinforcement tube. It must be recognised that the results are qualitative, it can always only be understood as probability of welding.

Table 1. Values of the heat transfer conditions during simulation

Measure	Value AA6060	Value AA7075	Unit
Thermal conductivity	180.175	180.181	W/m * K
Thermal expansion	2.2e−05	2.2e−05	$K-^{1}$
Heat capacity	2.43398	2.43369	$J * K^{-1}$

The evaluation of the diameters is done by point tracking, the result is evaluated starting from the starting point of the selected points along the path. The maximum value for the weld quality is then selected.

Depending on the application of the component, the reinforcing element should be located in different areas, therefore the influence of the position of the reinforcing element on the process success and the weld strength was investigated in a first numerical test series. At Fig. 5 those paths are shown, starting from different initial tube diameters. The container diameter is $d_0 = 66$ mm, the profile has a diameter of $d_1 = 20$ mm, which results in an extrusion ratio $R = 10{,}89$. The ram velocity is set at $v = 1$ mm/s, the billet has a temperature of $T = 430$ °C. The inner block radius r_i is varied in 4 mm steps, starting at 4 mm and ending at 20 mm. The wall thickness of the tube is $s = 2$ mm. Part a on the left side of the picture shows an example of the path of a reinforcing element with an inner radius of 4 mm.

On the right side of the picture the contour plot of the weld quality factor is shown, the black lines represent the paths of the varied internal surfaces. The area of the enlarged image represents the forming zone, where the highest weld quality factor is achieved. This can be explained by the previous image. Sufficiently high pressure and high strains meet there. The green circles mark the measuring points for the respective weld quality factor and are plotted in Fig. 6a) over the initial inner radius related to Fig. 5. It can be seen that the probability of welding increases linearly with increasing radius. The blue curve represents the probability of welding on the outside of the pipe, also plotted over the initial radius. It should be noted, however, that the comparison over the same radius comes from different simulations, due to the wall thickness this is shifted by 2 mm. It is noticeable that the probability of welding is higher outside the tube than inside. The reason for this is the strain distribution in the container, external surfaces of the tubes experience a higher strain than the inner surfaces. At a initial radius $r_i = 20$ mm the weld quality factor for the inner surface is $\sigma_b/\sigma_f = 0.96$, for the outer surface $\sigma_b/\sigma_f = 0.99$. At these values it can be assumed that solid state bonding takes place. However, if the initial radius is increased to more than 20 mm, process errors will occur as shown in Fig. 6b). Further evaluations show that the wall thickness has no influence on the process window.

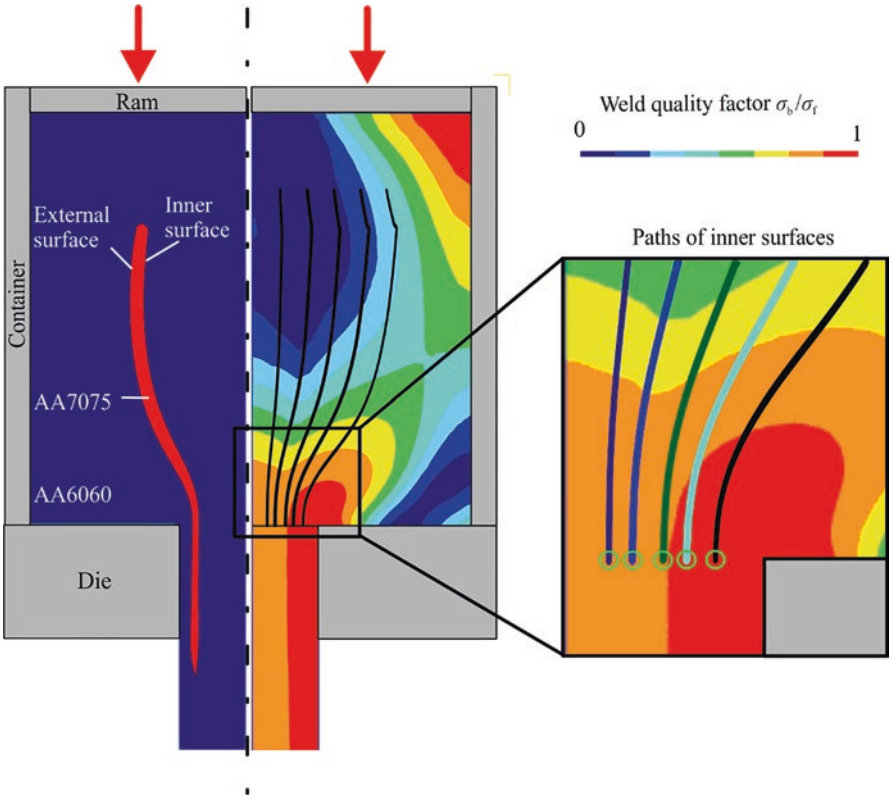


Fig. 5. Paths of reinforcements with different radii with contour plot of the weld quality factor

Regardless of the radius of the tube, the probability of welding can be increased by using a different die diameter. An increase in the extrusion ratio during extrusion increases pressure and stress, thus increasing the probability of welding.

However, there are geometrical and technological limits to maximising the two factors that would significantly increase the probability of welding. The extrusion ratio is particularly limited by the maximum force of an extrusion press, since the extrusion force increases with the extrusion ratio. The limit of the tube diameter is more complex. The higher the tube diameter, the more the tube is stretched and can constrict. In simulations with tube radii larger than 20 mm, errors such as separation of the reinforcing element from the aluminum matrix or necking occur.

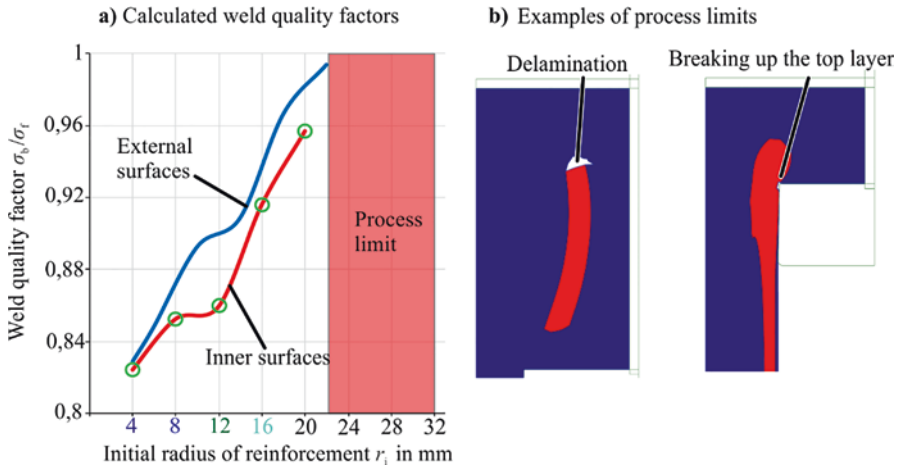


Fig. 6. a) Weld quality factor of external and inner surfaces plotted over different initial radii related to Fig. 4 (extrusion ratio $R = 10.89$); b) examples of process limits

4 Subsequent Cold Forging of the Composite Extrudate

For further investigation of the metallic bond strength of the hybrid extrudate, a subsequent cold forging process is considered. The manufacture of a flange-shaped component by means of cold forging leads to critical stress-strain states and thus, delamination in the interface between the different materials might be initiated. The selection of suitable cold forging process parameters might appear difficult, that is why a thorough simulative consideration is reasonable. The aim of the investigation is to define the prevailing contact pressure distribution in the interface between the different workpiece materials and subsequently to identify process windows for the cold forging process.

For that reason, a simulation procedure, based on the hybrid extrusion results, was developed. To achieve reliable simulation results, compression tests were conducted to obtain flow curve data for both aluminum alloys of the composite, namely AA6060 (matrix) and AA7075 (reinforcement). The specimens were machined from monolithic extrudates maintaining the same extrusion ratios as in the composite extrusion process. In a first simulation step, a boolean operation was applied to the composite extrudate to cut out the initial workpiece geometry for the subsequent cold forging process (see Fig. 7 a). While maintaining in simulation the prevailing stress-strain distribution of the billet, the flow stress $\bar{\sigma}\left(\bar{\varepsilon}, \dot{\bar{\varepsilon}}, T\right)$ was defined as strain, strain rate and temperature dependent. Two different cold forging process variants were investigated: Lateral extrusion as well as conventional upsetting. The different process characteristics led to entirely different stress-strain states in the interface between the two components. The interface contact pressure in the last calculation step of an exemplary lateral extrusion process is shown in Fig. 7b). As depicted, the contact pressure decreases significantly when increasing flange radius (x-direction).

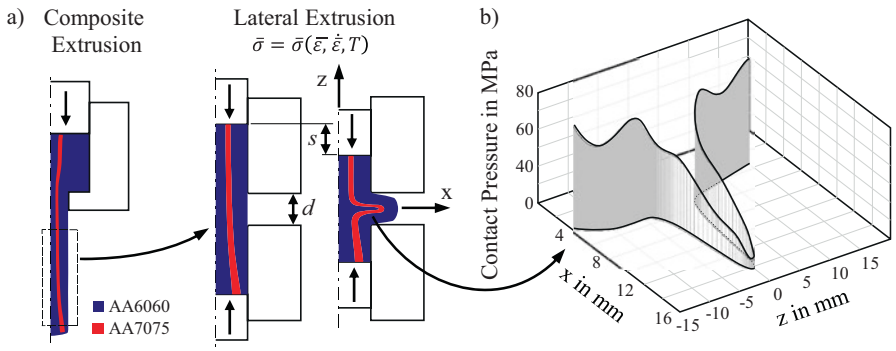


Fig. 7. Lateral extrusion of composite billets a) simulation procedure, b) contact pressure profile at the end of an exemplary lateral extrusion process

The lateral extrusion process parameters punch stroke s and die distance d were found to influence the contact pressure distribution most significantly. Based on this finding, a cold forging tool, featuring a closing device capable to vary the die distance d, was designed. The closing device enables a synchronization of the upper and lower punch movement, resulting into a symmetric material flow.

5 Conclusion and Outlook

A process route for composite hot extrusion with tubular reinforcements and subsequent lateral extrusion was shown. For subsequent processing by cold forging the hot extruded parts it is necessary to generate a sufficient solid state bonding to prevent delamination and formation of failures. The results of finite element simulations provided the following findings:

- Despite the opposite distribution of pressure and strain, it seems possible to achieve solid state bonding during extrusion. A high extrusion ratio and large diameters of the reinforcing tubes are the main drivers of welding.
- However, process limits occur with large tube diameters coupled with high extrusion ratios. This results in a process window in which welding and error-free extrusion seems possible.
- The outside of the tube is more likely to be welded due to the higher strains. Accordingly, it can be assumed that as long as the inner surface has entered into a material bond, the external surface has also exceeded this threshold.
- The proposed simulation methodology is a reasonable to investigate the subsequent cold forging process. Based on the identification of significant process parameters influencing the contact pressure distribution, a cold forging tool was designed.
- However, a criterion for delamination during cold forging due to contact pressure loss is necessary to determine feasible process windows.

Experimental investigations are necessary to prove solid state bonding during composite extrusion with reinforcing elements. The generated profiles could be verified by push-out tests or EDX-images of the boundary layer. In addition, the knowledge generated is to be extended to higher strength reinforcing elements in order to increase the range of components. Cold forging tests with the proposed tooling featuring a closing device provide further information about the interface conditions and serve for the simulation validation in terms of load profiles and metallographic analyses. Based on these findings, feasible cold forging process windows, maintaining the metallic bond of the composite, are to be defined. Static load tests of the forged components are necessary to study the effect of the reinforcement on the product properties.

Acknowledgements. This work is based on investigations of the research project "Production of bead-shaped parts by cold extrusion of semi-finished products fabricated by composite hot extrusion" funded by the German Research Foundation (DFG, Project number TE 508/72-1). Its financial support is greatly acknowledged.

References

1. Sheppard, T.: Extrusion of Aluminium Alloys. Kluwer Academic Publishers, Dordrecht (1999)
2. Kleiner, M., Schomäcker, M., Schikorra, M., Klaus, A.: Production of composite reinforced aluminum profiles for ultralight structures by hot extrusion. Materialwissenschaften Werkstofftechnik **35**, 431–438 (2014)
3. Dahnke, C., Pietzka, D., Haase, M., Tekkaya, A.E.: Extending the flexibility in the composite extrusion process. Procedia CIRP **18**, 33–38 (2014)
4. Dahnke, C., Shapovalov, A., Tekkaya, A.E.: Thermally activated lightweight actuator based on hot extruded shape memory metal matrix composites (SMA-MMC). Procedia Eng. **207**, 1511–1516 (2017)
5. Nakamura, T., Tanaka, S., Hiraiwa, M., Imaizumi, H., Tomizawa, Y., Osakada, K.: Friction-assisted extrusion of thin strips of aluminium composite material from powder metals. Ann. CIRP **41**(1), 281–284 (1992)
6. Chen, Z., Ikeda, K., Murakami, T., Takeda, T.: Extrusion behavior of metal-ceramic composite pipes in multi-billet extrusion process. J. Mater. Process. Technol. **114**, 154–160 (2001)
7. Bay, N.: Cold pressure welding - the mechanisms governing bonding. J. Eng. Ind. **101**(2), 121–127 (1979)
8. Cooper, D.R., Allwood, J.M.: The influence of deformation conditions in solid-state aluminium welding processes on the resulting weld strength. J. Mater. Process. Technol. **214**, 2576–2592 (2014)
9. Kolpak, F., Schulze, A., Dahnke, C., Tekkaya, A.E.: Predicting weld-quality in direct hot extrusion of aluminium chips. J. Mater. Process. Technol. **274**, 116294 (2019)

Experimental Springback Validation of a Finite Element Model of Multi-stage Stator Bending

D. Wüterich[1(✉)], M. Liewald[2], and M. Kopp[1]

[1] SEG Automotive Germany GmbH, Lotterbergstr. 30, 70499 Stuttgart, Germany
david.wueterich@seg-automotive.de
[2] Institute for Metal Forming Technology, University of Stuttgart, Holzgartenstr. 17, 70174 Stuttgart, Germany

Abstract. Today's omnipresent environmental problems have led to an increased demand for electrical machines and its subcomponents – i.e. stators, rotors and control units - that show highest quality/price ratios. The flat-pack technology is a multiple stage method of bending flat, comb-like stator cores with inserted copper windings into rounded rings. Despite the complexity caused by simultaneously acting bending effects, that stator manufacturing technology provides potentials regarding achievable stator performance due to its relatively high copper fill factor. In previous investigations, numerical models of one single bending operation have been validated. In contrast, this investigation focusses on predicting the stator geometry after multiple bending operations. Therefore, simulation models were extended to calculate such bending processes within one single model. Finally, results gained from experiments and simulation runs with same settings are presented. Discussion of results discloses a conformity between numerical and experimental geometries after multiple bending stages beyond 94%.

Keywords: Stator manufacturing · Springback · Multi-material bending

1 Introduction

Current legal regulations concerning pollutant emissions are leading to a fundamental change in vehicle industry with regard to applied driving engine technologies [1]. As a consequence, demands for electric machines grow and requirements on the power density and production costs of their subcomponents such as stators, rotors and control units do increase. Since the stator is one of the main functional subcomponents, its manufacturing technology has a significant impact on performance and overall production costs of the electric motor. Stators mainly consist of electromagnetic coils made of coated copper wires, as well as a so-called stator core, which consists of numerous assembled electrical steel lamellas. One of the main goals for stator

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 202–211, 2021.
https://doi.org/10.1007/978-3-662-62138-7_21

manufacturing is achieving a high mechanical copper fill factor k [2]. The copper fill factor k represents the ratio of the cross section of the wires A_{Cu} to the cross section of the stator core grooves $A_{Gr,2}$, as depicted in Fig. 1b). In addition, a reduction of winding lengths, wire deformations and costs are important for stator manufacturing technology of today. Over the past, this resulted in several manufacturing strategies, for example trickle, linear, pull-in, needle or flyer winding [3]. During recent years, these technologies are gradually replaced by the hairpin technology, which allows for creating stators with highest copper fill factors [4]. An alternative manufacturing method of stators is currently being investigated in collaboration between the Institute for Metal Forming Technology, University of Stuttgart (IFU) and SEG Automotive Germany GmbH. In contrast to many winding technologies, this so-called flatpack technology [5] uses a flat, comb-like stator core body as input material, as shown in Fig. 1a). After inserting pre-bent copper windings, the entire assembly is rounded by bending the steel lamellas in a multi-stage bending process (see Fig. 1b). Subsequently, a welding operation joins both rounded stator core ends, as illustrated in Fig. 1c). The main benefit of the flatpack bending strategy consists in avoiding the welding of windings and still achieving mechanical fill factors beyond 90%, which are comparable to hairpin technology. However, the inner stator geometry shows higher roundness deviations of more than 0.1 mm. This is due to its dependence on mechanical properties of the assembled materials and due to the interaction between the bending operations. In particular, elastic recovery of the workpiece after bending, known as springback behavior, shows significant influence on that inner roundness accuracy. Due to high complexity of the assembled flatpack (multiple materials with nonlinear mechanical behavior and contact situations), a high robustness in production cannot be achieved when defining bending parameters solely based on experience. Therefore, previous investigations on the flatpack technology focused on the development of a nonlinear finite element model for predicting the bending behavior of flatpacks with varying process parameters [6] and workpiece properties [7]. Those studies only dealt with one single bending operation. In this investigation, the FE-model was further developed and experimentally validated to be used for simulating the bending behavior of flatpacks during multiple operations.

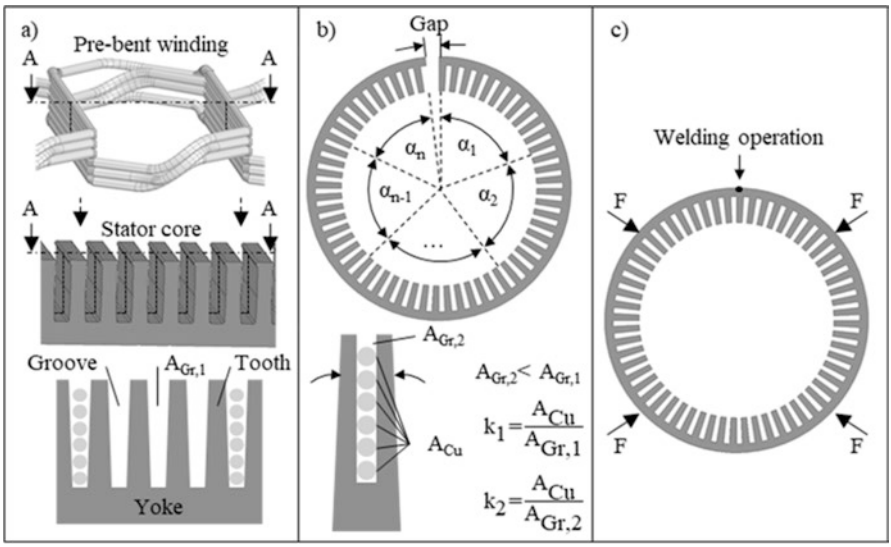


Fig. 1. a) Insertion of pre-bent winding into stator core; b) Reduction of groove section during multi-stage bending operation; c) Welding operation

2 Experimental and Numerical Flatpack Bending Methodology

The fundamental concept of the flatpack technology is based on the application of large stator groove sections $A_{Gr,1}$ allowing windings with large diameters to be inserted (see Fig. 1a). Flatpack technology therefore results into highly efficient stators with mechanical copper fill factors beyond 90%. During bending, movement of the stator teeth reduces the groove section from $A_{Gr,1}$ to $A_{Gr,2}$ and thus increases the copper fill factor from k_1 to k_2, as shown in Fig. 1b). In this respect, SEG has found a specific combination of bending angles α_1 to α_n and bending die diameters enabling the windings to be inserted into the grooves of the stator. In the following, the experimental setup for multi-stage flatpack bending is described. Subsequently, the setup of the FE-model, i.e. input geometry, boundary and contact conditions during bending are presented as well as a new approach for experimentally validating previously gained simulation results.

2.1 Experimental Setup and Digitization of Geometries

The real flatpack bending processes were performed using position-controlled machines. These semi-automatic bending machines ensured the exact setting of the bending angles. To enable the validation of the numerical geometry, first the experimental geometry of processed stators was evaluated according to [7]. For this purpose, digitized measurement data of the stepwise bent stator were created by using a 3D-scanning device having a precision of 0.05 mm based on structured light (see Fig. 2a). Afterwards, cross sections were extracted at different axial locations of the outer surface using GOM Inspect software to obtain two-dimensional representations of the outer geometry. Figure 2b) shows a detailed view of the black cross section containing two measurement errors, a gap on the even surface and a missing sharp edge on the left side. These measurement errors caused by wrong reflection of light occur despite varying settings of the 3D-scanning device and external brightness. For further evaluation, the software was used to automatically distribute surface points (sp) on the outer black cross section of the bent stator core. To compensate for the wrong representation of the edge, the corner point was manually added to the measured set of nodes. Therefore, two straight dashed lines were calculated using the existing surface points. Here, line 1 is defined by the surface points sp_2, sp_3 and sp_4. Line 2 intersects surface point sp_1 and is perpendicular to line 1. The intersection between both lines systematically defines the corner point of the missing sharp edge, the set of measured points remained unchanged. The outcome of this procedure is a cloud of 412 surface points representing the outer geometry of the flatpack from one flatpack end to the other. For comparison, a cloud of numerical node coordinates was generated using a FE-model, as shown in Sect. 2.2. These procedures facilitate the accurate measurement of several characteristic parameters of the virtual and real stator geometries, as shown in Sect. 2.3.

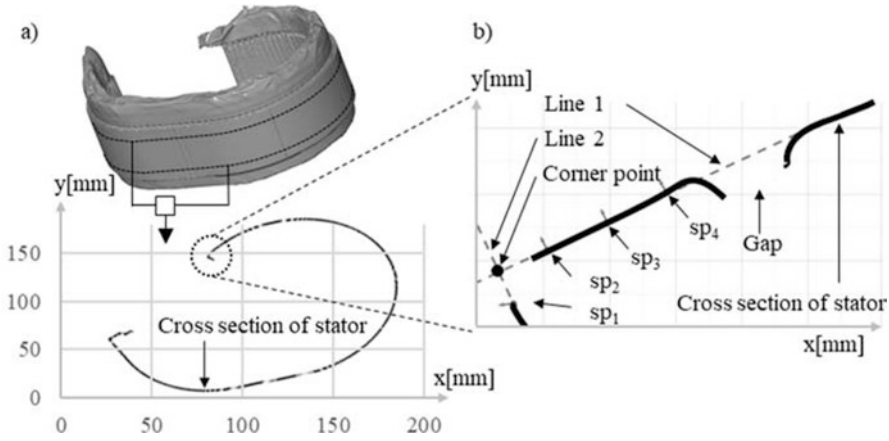


Fig. 2. a) Extraction of outer flatpack geometry into two-dimensional coordinates; b) Manual rework on measurement errors of extracted cross section

2.2 Numerical Setup of the Multi-stage Flatpack Bending Process

The FE-model used for the investigations described in this paper was created in Ansys Mechanical with 3D SOLID 186 elements having quadratic displacement behavior. These elements can be used for calculating nonlinear deformations. The multilinear hardening behavior of steel lamellas was obtained by means of standard uniaxial tensile tests. The mechanical behavior of the copper wires was defined with a standard nonlinear strain-hardening model provided by the software. In future investigations, it is necessary to determine the mechanical behavior of the pre-bent copper windings. For simplification, the lamella structure of the stator was replaced by a solid body and geometrical details such as small chamfers and fillets were eliminated. The flatpack, tool 1, tool 2 and pressure die (PD) as well as the corresponding boundary conditions are shown in Fig. 3. For easier interpretation, only two tool situations are shown, copper windings between the stator teeth are hidden. Performing multiple operations within one simulation model in Ansys Mechanical requires the definition of several load steps. Within one load step, one movement with a set of predefined boundary conditions is calculated. One boundary condition is defined by a "remote displacement" on the external point p_1, which is related to the inner surface s_1 of tool 1. In this load step 1, all degrees of freedom of p_1 are set to zero except the z-rotation, which is defined by the desired bending angle α_1. During this phase, the remote displacement of tool 2 is disabled, meaning that it can move freely depending on the movement of bonded tooth 20. In the real process, there is no relative motion between pressure die and flatpack bottom (FP_bottom), therefore the contact behavior was set to frictionless. In load step 2, tool 1 is fixed to its final z-angle and the pressure die moves in negative y-direction, allowing the elastic springback of the workpiece. Tool 2 changes its location depending on the bending and springback behavior of the workpiece during operation 1.

Load step	Tool 1	Tool 2	PD	
			x	y
1	α_1	-	0	0
2	α_1	-	0	-30
3	-	f	0	-30
4	-	f	0	-30
5	-	D_0	0	-30
6	-	D_0	$-x_1$	-30
7	-	D_0	$-x_1$	0

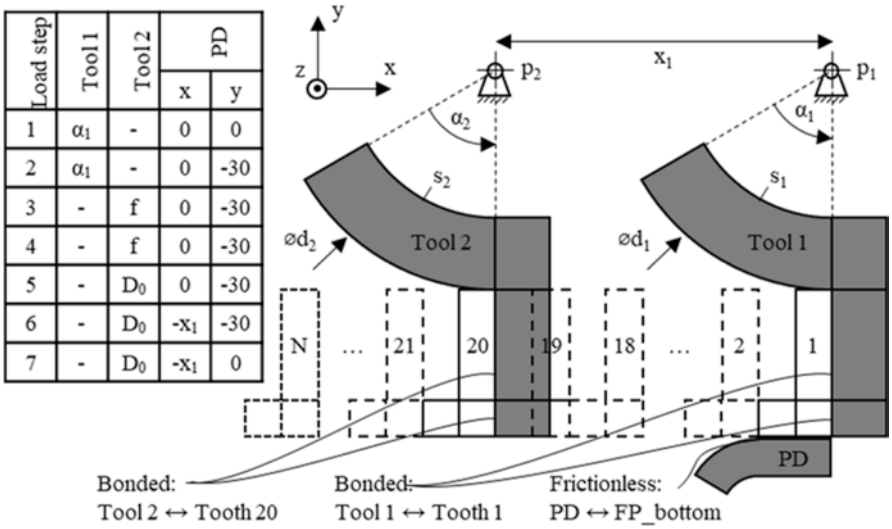


Fig. 3. Simulation setup with two bending tools, one pressure die and the simplified flatpack

Afterwards, a change of boundary conditions is necessary to continue with the next bending operation. In load step 3, the boundary conditions of the tool surface s_2 of tool 2 are fixed (f) using the APDL command "%_fix%". Subsequently, in load step 4, the boundary conditions of tool 1 are disabled, now allowing the whole assembly to be moved in dependence of tool 2 instead of tool 1. In load step 5, bending tool 2 moves to its initial location (D_0). In the next load steps the pressure die moves in $-x_1$ to be located underneath tool 2. Afterwards, PD moves in positive y direction to create the desired contact situation between flatpack bottom and pressure die. Subsequently, the next bending operation starts. After each completed bending operation, the numerical node coordinates of the flatpack bottom are exported to an Excel file, providing a two-dimensional point cloud of the outer flatpack geometry.

2.3 Analysis of Experimental and Numerical Point Clouds

The experimental and numerical results are both point clouds representing the outer flatpack surface. The experimentally measured cloud of surface points is based on the 3D-scan, the numerical point cloud is based on the nodes of the finite element mesh during and after calculation. Figure 4a) shows a 2D coordinate system with a section of the experimental surface nodes. Figure 4b) shows a similar section with the numerical nodes. For reasons of better visibility, the plots only display a section of the flatpack and a reduced number of points. For all calculations, the entire set of points were used. In the following, it is assumed that the outer geometry of the flatpack can be described by a polyline, which is based on straight lines and arcs. That polyline is fully defined by geometric parameters, i.e. the length values of the straight lines (L_1 and L_2) plus the radius r and the angle α of the arcs. While the point clouds remain at their position, the geometric parameters are changed in a way that the squared deviation between the polyline and the surface points is minimized, according to the method of least squares. Therefore, an Excel-based nonlinear optimization algorithm was used. The objective of that optimization is minimizing the squared deviation while the geometric parameters of the polyline are the variables of the optimization. The result is a quantitative description of both point clouds using a small number of geometric parameters (L_1, L_2, r, α).

3 Results

The objective of the research work reported about in this paper was to develop and to validate a numerical model for precisely predicting the springback behavior of flatpacks after multi-stage stator bending. To evaluate the suitability of the simulation model developed in this purpose, numerical and experimental results of three representative process stages are compared. Here, one bending angle of 242°, one bending angle of 76° and the final geometry after all bending operations are considered, since they match the stages of the real SEG process. As mentioned before, a gap between both stator ends is inevitable after the multi-stage bending procedure due to the

springback of the last bending operation. The gap size is relevant for subsequent processes. Figure 4a) and 4b) show a 2D coordinate system containing the point clouds of the outer flatpack surface after an experimental and a numerical bending operation using a small bending angle of 76°. A visual comparison between both point clouds indicates a good congruency. Nevertheless, a quantification is only possible with further evaluation. For this purpose, the nodes serve as anchor points for a polyline based on geometric parameters. Figure 4a) and b) show the polylines of the considered sections. By manipulating the parameters L_1, L_2, r and α, the sum of all quadratic deviations between the polyline and the point cloud is minimized leading to a set of parameters that optimally fit the point cloud.

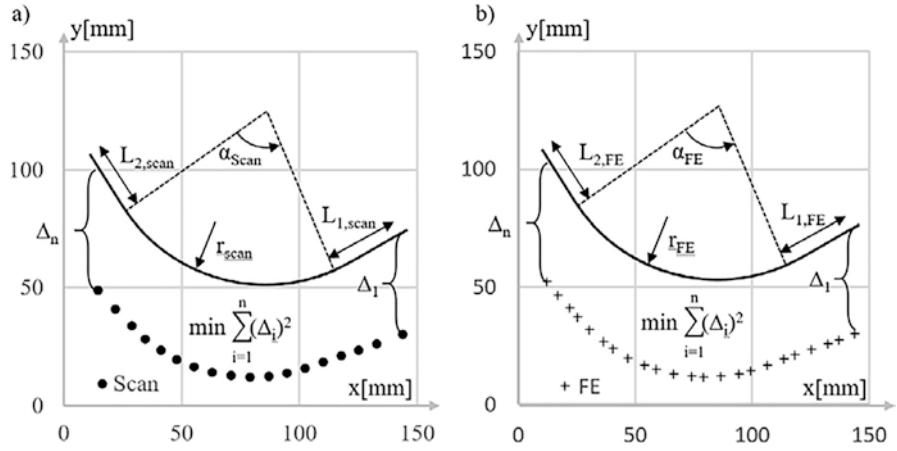


Fig. 4. Polyline based on four parameters a) Surface nodes of scanned experiment; b) Surface nodes of finite element simulation

Table 1 summarizes the quantified results of the experimental evaluations. Columns of the table contain the average values and the standard deviations based on the corresponding bending operation. Based on geometry measurements of four different stator samples, an average value was calculated with the standard deviation as an indicator for the reproducibility of the bending process. During each bending operation, the elongation of the outer layer causes a change in length. Starting with a length of 422.01 mm after the first bending operation, the length evolves to a value of 433.17 mm after the final bending operation. The average angle α_1 as well as the radius r_1 after bending and subsequent springback does not change after bending operation 1. Same applies to the angle α_2 and the radius r_2 after bending operation 2. For this reason, the values are indicated with "=". The average gap dimension between both stator ends only emerge relevant after the final bending operation.

Table 1. Experimental evaluation of four outer flatpack geometries after bending operations

Parameters	Bending operation 1		Bending operation 2		Final bending operation	
	Average	Std. dev.	Average	Std. dev.	Average	Std. dev.
Length [mm]	422.01	0.32	431.29	0.16	433.17	0.22
α_1 [°]	113.85	0.11	=	=	=	=
r_1 [mm]	71.39	0.06	=	=	=	=
α_2 [°]	x	x	224.75	1.10	=	=
r_2 [mm]	x	x	71.28	0.35	=	=
Gap [mm]	x	x	X	x	7.51	0.41

A comparison between the quantified results of the numerical and experimental investigation is shown in Table 2. The structure of the table is similar to the structure of the experimental evaluation, given that the rows show the calculated parameters and the columns display the bending operations. Every parameter was calculated for the numerical and for the experimental results, hence the absolute difference allows for a first quantitative comparison. The absolute difference of the total outer flatpack lengths was calculated between 0.01 mm and 0.25 mm. The deviations of the angles after springback for α_1 and α_2 of the corresponding section differ between 1.16° and 2.16°. The difference in radius for r_1 and r_2 after springback emerge less than 0.75 mm. Finally, the target value is the gap between both stator ends, which is 0.45 mm larger in the simulation compared to the experimental measurement.

Table 2. Comparison between calculated parameters of numerical and experimental evaluation

Parameters	Bending operation 1			Bending operation 2			Final bending operation		
	FEM	Exp.	Δ	FEM	Exp.	Δ	FEM	Exp.	Δ
Length [mm]	422.00	422.01	0.01	431.04	431.29	0.25	432.94	433.17	0.23
α_1 [°]	112.69	113.85	1.16	=	=	=	=	=	=
r_1 [mm]	70.90	71.39	0.47	=	=	=	=	=	=
α_2 [°]	x	x	x	222.59	224.75	2.16	=	=	=
r_2 [mm]	x	x	x	70.75	71.28	0.53	=	=	=
Gap [mm]	x	x	x	X	x	x	7.96	7.51	0.45

4 Discussion

Results presented in Table 1 allow for concluding that the semi-automatic bending process and the geometry evaluation method show acceptable reproducibility for further validation procedures. The relative standard deviations, i.e. the ratio between standard deviation and average value, are highest for the final gap between both stator ends and lowest for all total length values. The relative standard deviation of the final

gap is 5.5 and of the length after bending operation 2 it is 0.08. Of course, a statement on the process robustness is only possible with a higher number of samples and measurements. When comparing the numerical and the experimental results, it however becomes clear that the multi-stage simulation model can predict the bending behavior of the highly nonlinear material constellation of flatpack. The relative deviations of all parameters except the gap length differ less than 3%. The final gap of rounded stator ring adds up as a result of multiple bending operations, thus the minor deviations of each bending operation contribute differently. Using the simulation setup as described, the relative difference between the measured and the simulated gap length is 6%. As shown in Table 2, that relative value is based on the absolute of 0.45 mm.

Using the multi-stage simulation model as described, now it is possible to calculate the influence of several process and dimensional parameters of the stator on its final geometry after multiple flatpack bending process steps. That includes the possibility of performing a numerical optimization of the bending angles and bending tool diameters with an accuracy of more than 94%. Especially the elastic springback of the last bending operation, which leads to the final gap between both stator ends, appears significant for the quality of bending process sequence. If the gap is comparatively large, the welding operations subsequent to the bending process are more complicated or impossible. If the gap occurs too small, both stator ends collide during the last bending operation, which at the end leads to higher scrap rates.

The setup of the numerical model of the multi-stage stator bending process offers several potentials to improve the reliability of predicting the bending results during future research work. First, the material behavior of copper windings was set to standard nonlinear copper. The mechanical behavior of that material has to be investigated to improve its influence on the forming behavior of the entire assembly. Additionally, the same multi-stage simulation setup could be validated with further flatpack designs and process parameters with a higher number of samples to expand its validity.

5 Summary

Being one of the leading-edge technologies regarding achievable copper fillings of stators, the main challenge of the flatpack bending technology lies in accurately predicting the forming behavior and thus the resulting stator geometry. Therefore, the investigation reported about in this paper presents the validation of a multi-stage stator bending simulation. A numerical model was created in Ansys Mechanical, which allows for calculating multiple bending and subsequent springback operations by defining multiple load steps. The objective of this procedure is processing a flat, comb-like assembly of lamellas and copper windings into a circular stator for electric machines. The highly nonlinear bending behavior of the workpiece was evaluated regarding geometrical parameters and validated by experiments. The accuracy of predicting single geometrical parameters numerically was higher than 97% and the gap between both ends of the circular workpiece was calculated to be 6% larger than the experimental result.

Reference

1. Helms, H., Jöhrens, J., Kämper, C., Giegrich, J., Liebich, A., Vogt, R., Lambrecht, U.: Weiterentwicklung und vertiefte Analyse der Umweltbilanz von Elektrofahrzeugen. Texte 27. Umweltbundesamt, Dessau-Roßlau (2016)
2. Hagedorn, J., Sell-Le Blanc, F., Fleischer, J.: Handbook of Coil Winding. Springer Vieweg, Berlin (2018)
3. Halwas, M., Sell-Le Blanc, F., Jux, B., Doppelbauer, M., Wirth, F., Hausmann, L., Hofmann, J., Fleischer, J.: Coherences between production technology and performance of electric traction drives. In: 9th International Electric Drives Production Conference (EDPC), pp. 1–9. IEEE, Esslingen, Germany (2019)
4. Wirth, F., Kirgör, T., Hofmann, J., Fleischer, J.: FE-based simulation of hairpin shaping processes for traction drives. In: 8th International Electric Drives Production Conference (EDPC), pp. 1–5. IEEE, Schweinfurt, Germany (2018)
5. Didra, S., Laszlo, T., Gremmel D.: Flat package for manufacturing a stator. Patent DE201310222643 (2013)
6. Wüterich, D., Liewald, M., Kopp, M.: Numerical investigations on the influence of material properties on bending behavior of comb shaped stator bodies. AIP Conf. Proc. **2113**(1), id. 110011. AIP Publishing LLC (2019)
7. Wüterich, D., Liewald, M., Kopp, M.: Validation of bending simulation models based on yield strength influences of electrical steel sheets on stator core radius. In: 9th International Electric Drives Production Conference (EDPC), pp. 1–6. IEEE, Esslingen, Germany (2019)

Cutting Machine Tools and Manufacturing Processes

Concept of a Mechatronic System for Targeted Drill Head Direction and Angular Alignment Control in BTA Deep Hole Drilling

J. F. Gerken(✉) and D. Biermann

Institute of Machining Technology, TU Dortmund University, Baroper Str. 303, 44227 Dortmund, Germany
julian.gerken@tu-dortmund.de

Abstract. The compensation of straightness deviation in BTA deep hole drilling represents a major challenge caused by the inaccessibility of the cutting zone and the multitude of influencing factors. It is required for a variety of drilling applications with length-to-diameter-ratios larger than l/D > 10 and high quality features such as roundness, diameter accuracy and surface quality. In addition to these classical requirements, the straightness deviations represents one of the most important quality features. This paper presents the conceptual idea of an innovative tool system for in process compensation of straightness deviation. It can be retrofitted for existing deep drilling machines and carries out the straightness deviation correction without process interruption. Fundamental idea is the process parallel measurement of the wall thickness with an ultrasonic measurement device, which is connected to an actuator system by a control loop and allows the drill head to be slightly tilted in a targeted manner.

Keywords: BTA deep hole drilling · Straightness deviation · Compensation unit

1 Introduction

Deep drilling methods allow the production of bores with a large bore depth compared to the diameter and they are economically used for a variety of drilling applications with a length-to-diameter-ratio larger than l/D = 10. The different tool designs are used for the production of deep bores based on specific aspects of the mechanical processes. For one thing, there are tools with an asymmetrical single-edged design and secondly, there are tools with two symmetrically arranged cutting edges [1]. Symmetrical tool designs are used in twist and double-lip drilling. The classical deep hole drilling methods with an asymmetrical design are single-lip drilling, ejector drilling with a double-tube system and BTA deep drilling with a single-tube system. The

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 215–224, 2021.
https://doi.org/10.1007/978-3-662-62138-7_22

asymmetrical design leads to self-centering of the tool in the bore by guide pads [2]. In this paper the BTA deep drilling is considered in more detail. The two main parts of a BTA tool system are the drill head with exchangeable cutting inserts and exchangeable guide pads and the cylindrical drilling tube. BTA deep drilling is used in the bore diameter range of D approx. 10…1500 mm. Characteristic features of this process are the special cooling lubricant supply with a high volume flow and a pressure of p = 4…10 bar. The cooling lubricant enters the annular gap between the drilling tube and the drill guide via the oil pressure head and not through a special coolant channel [3]. During the machining process, the cooling lubricant passes through the produced bore to the drill head, flows around the guide pads and cutting edges and flushes the chips inside the drilling tube out of the bore. This results in an improved surface quality, because the chips have no contact with the bore wall [4, 5]. It is often used at the end of the value chain for machining expensive components such as hydraulic cylinders, transmission shafts and aircraft landing legs [1].

In addition to classical drilling requirements, such as roundness, diameter accuracy and surface quality, the straightness deviation represents an important quality feature of deep bores. It reflects the radial position error of the current drilling axis relative to the ideal drilling axis and is decisive for the functional capability of the machined workpiece [6, 7]. Influencing factors are the modes of operation (rotating tool, rotating workpiece or counterrotating), cutting parameters, tool construction, workpiece design and material and temperature distribution in the workpiece [8, 9]. Complete avoidance of straightness deviation is not possible due to the complex interactions between the influencing parameters. Due to this fact a tooling system for in-process compensation of straightness deviation is highly needed [10, 11].

2 Approaches to Influence Straightness Deviation

Due to the high technical and economic importance of the reduction of the straightness deviation, different investigations have already been carried out in this research area. On the one hand, an improved chip breaking behavior could be achieved through the pulsation of cooling lubricant, which resulted in an increased process stability. In cooperation with the TU Wien, this procedure was further developed and transferred to industrial application in collaboration with the company *Schoeller-Bleckmann Oilfield Technology GmbH* [12]. To reduce already existing straightness deviation in deep bores, the US Navy developed a concept, which records the deviation with a laser measuring system through special channels and compensates this deviation by adjusting the drilling direction [13]. The drilling direction is adjusted hydraulically. Further revisions of the actuator technology were made with piezoelectric actuators. Both the hydraulic solution as well as the piezoelectric based development require a pilot bore and are only available within the range of diameters d = 100…130 mm. In the field of ejector deep drilling, a mechanical concept was developed by *Schwäbische Hüttenwerke GmbH*. In this concept, a non-rotating thrust piece is positioned over a pipe between the bore wall and the drilling tube so that the drilling tube is tilted in the desired direction. This process is used to produce cooling channels in calender rolls.

The detection of the straightness deviation, which is caused by differences in microstructure and hardness, is done with an ultrasonic measuring system. Kessler investigated a further possibility of reducing the straightness deviation by inductive heating of the component [9, 14, 15]. Iovkov developed a method for straightness deviation compensation based on a radial traverse of the machine spindle superimposed on the axial feed. The drill guided in the bore is slightly tilted, and the drilling direction is corrected by the caused deflection [16]. The different approaches show that a reduction of the straightness deviations is of great economic interest. Especially in the field of BTA deep drilling, a reduction of the straightness deviation is highly needed in demand due to the expensive and complex workpieces.

3 Conceptual Design of the System

The concept of the automated system for compensation of straightness deviation in BTA deep drilling is shown in Fig. 1. The developed system can be retrofitted for existing deep drilling machines and carries out the straightness deviation correction without a process interruption. The structure of the entire system consists of four main components, which are adapted for the experimental investigations on a Giana GGB 560 deep drilling machine. In the following, the four components are presented in detail. The concept design was developed according to the procedure of VDI guideline 2221 "Design of technical products and systems – Model of product design" [17, 18].

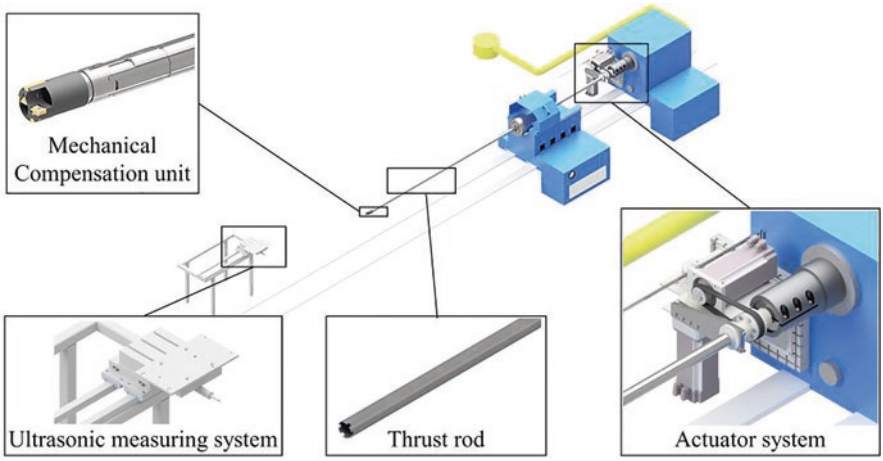


Fig. 1. Experimental deep drilling machine with the adapted correction device

3.1 Mechanical Compensation Unit

The design of the mechanical compensation unit comprises four main components. The assembly of the complete mechanical compensation unit is shown in Fig. 2 in half-section. A drill diameter of D = 60 mm is specified for the compensation system. This corresponds to typical industrial application. It is developed in such a way that it enables the drill head to be tilted in a targeted manner by means of low forces and can be screwed in between a conventional BTA drill head and drilling tube.

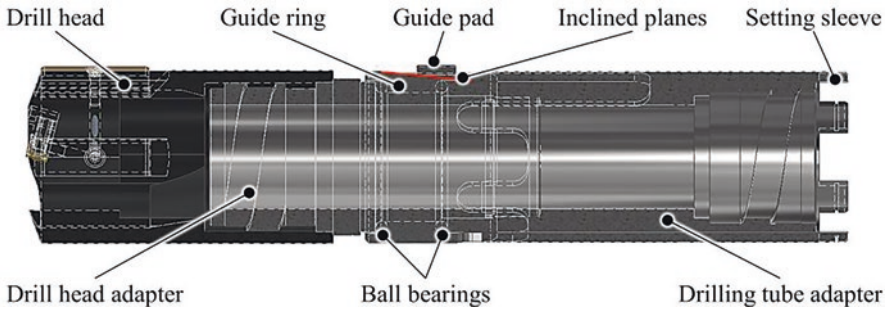


Fig. 2. Sectional view of the mounted mechanical compensation unit

The functions and specific properties of the individual components are explained in more detail below. The technical design of the mechanical compensation unit enables two main functions. On the one hand, the coupling function ensures the transmission of the drilling torque M_B between the drilling tube and drill head. The drilling tube and drill head adapter realizes the coupling function. On the other hand, the technical design allows a rotational positioning movement of the guide ring by 360 degrees and a translational positioning movement of the setting sleeve. The interaction of the guide ring and the translational movement of the setting sleeve results in a radial extension movement of the guide pad due to the inclined planes.

The assembly positions and the simulations of the mechanical compensation unit are shown in Fig. 3. The red lines outline the assembly positions of each component. Hardened bearing balls are additionally required for complete assembly. The drill head adapter and the drilling tube adapter are manufactured with inner running surfaces for the ball bearings. The opposite two running surfaces are fitted in the guide ring. The drill head adapter enables the coupling of the drill head to the drilling tube adapter on the drilling tube side as well as the axial fixing of the guide ring in feed direction by the bearing running surface.

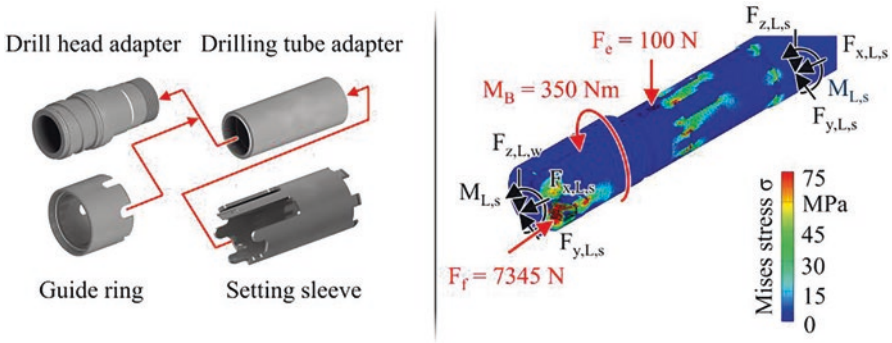


Fig. 3. Left: Mechanical components of the compensation unit; Right: Static load FE-analysis of the compensation unit with contact to the workpiece bore wall under process loads

The drilling tube adapter enables the coupling of the drill head adapter to the drilling tube as well as the axial fixing of the guide ring against the feed direction by the bearing surface. With the aid of an inclined plane, the guide ring converts the axial displacement of the setting sleeve into a radial extension force against the bore wall. The setting sleeve transmits the axial movement of the actuator to the inclined plane of the guide ring. In addition, the rotary movement of the actuator is transmitted via the toothed parts to the guide ring. Small guide pads can be screwed into the threaded bore in the setting sleeve. This design enables the decoupling of the tool rotation from the compensation unit rotation due to the ball bearing suspension. The connection of the compensation unit to the actuator system is realized with a thrust rod.

The FE-modeling of the compensation unit is shown on the right side in Fig. 3. It shows the drill head in the fixed workpiece. The executed static analysis shows the structural response of the unit as a result of an external force $F_e = 100$ N applied to the guide pad. Real process loads were also taken into account. The drilling torque $M_B = 350$ Nm, resulting from the cutting force F_c and the friction forces F_R on the guide pad, and the feed force $F_f = 7345$ N were measured in reference experimental investigations. The mesh was generated automatically by the used software, but refined into smaller elements at the guide pad and the contact zones between the setting sleeve and the guide ring. This improves the accuracy of the simulation. A simplified material model for alloyed hardened steel is used. Based on the simulations, the design was optimized and the position of the small guide pad on the setting sleeve was adjusted. Furthermore, it could be shown that the targeted tilting of the drill head under process loads is possible.

3.2 Thrust Rod

Another main component of the system is the thrust rod, which is required to realize the power transmission form the actuator system to the compensation unit. The length of the thrust rod depends on the drilling depth. For the experimental tests a thrust rod length of $l = 4500$ mm is used. The inner diameter of the thrust rod has to be larger

than the outer diameter of the drilling tube. In addition, the outer diameter of the thrust rod has to be smaller than the final bore diameter D = 60 mm because the thrust rod is positioned in the annular gap between the drilling tube and the bore wall. A special solution is used to connect the thrust rod to the setting sleeve of the compensation unit. The torque is transmitted with a frontal gearing. The translational movement is ensured by a groove in the frontal gearing teeth and an internal retaining ring. Due to the small space in the annular gap, a maximum wall thickness of the thrust rod of s = 2 mm is possible.

3.3 Actuator System

In the next step, the structure and connection of the actuator system is described in more detail. The conceptual design of the actuator system and its components are shown in Fig. 4. The entire actuator system is bolted to the tool headstock with a motor bracket. This enables a coupling of the system to the process feed. The two movement possibilities of the system are marked in red. Two electric motors are required for the two types of movement.

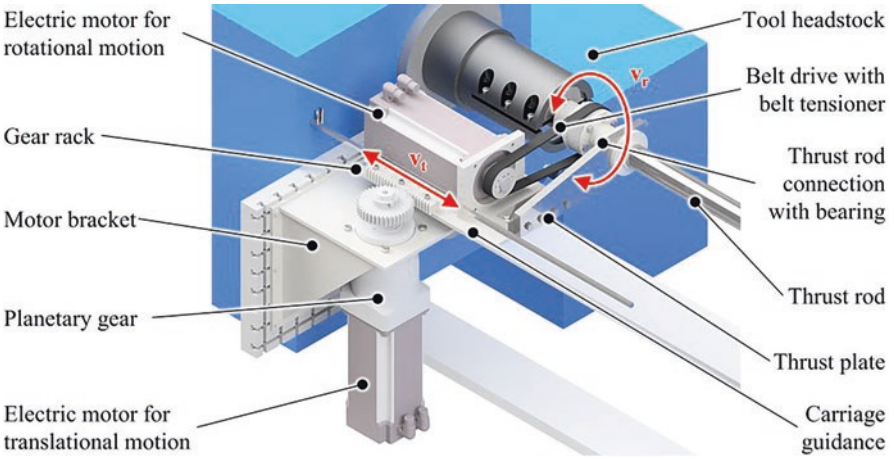


Fig. 4. Design of the actuator system

The translational movement v_t is realized by the vertically screwed motor, which is necessary for the extension movement of the guide pad of the compensation unit. A planetary gear is necessary for the finer translational movement. A gear rack realizes the conversion of the reduced rotational motion of the motor into a translational motion. The gear rack is screwed to a second motor plate with a carriage guidance. A thrust plate ensures the transmission of the movement from the second motor console to the thrust rod. The second necessary movement component v_r is required for the rotational orientation of the compensation unit. The rotational motion is transmitted to

the thrust tube by a toothed belt. It is necessary that the entire actuator system for the rotational movement v_r also performs the translational extension movement v_t.

The detailed view of the actuator system is shown in Fig. 5. As described in Sect. 3.2, the maximum wall thickness of the thrust rod is $s = 2$ mm. A clamping connection between the thrust rod and the thrust plate is not possible due to the thin-walled design. The power transmission is only possible by a soldered connection. Furthermore, a brass case is screwed to the soldering component, which realizes the power transmission. The assembly in Fig. 4 shows that all screwed components of the system carry out the rotational movement of the thrust rod. For the final assembly the counterpart of the thrust plate has to be screwed onto the thrust plate.

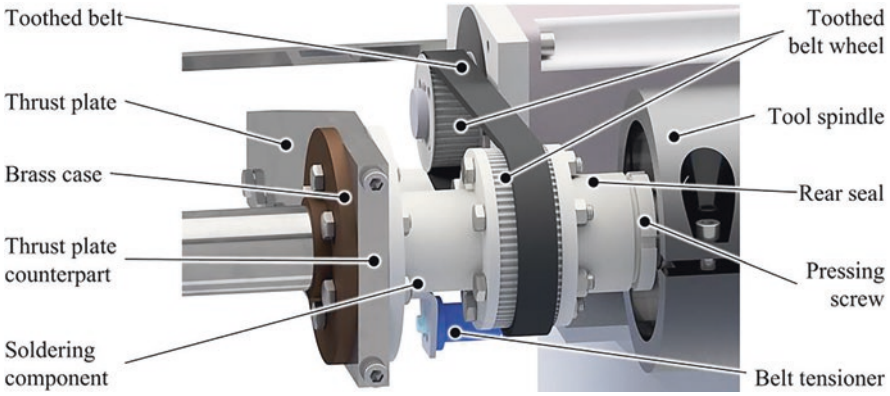


Fig. 5. Detailed view of the actuator system for rotational and translational motion transmission

The thrust rod bearing and sealing system is described in the sectional view in Fig. 6. Due to the large overhang of the thrust rod ($l = 4500$ mm), a bearing system is required for the later operation. The red marking shows the solder connection. The design ensures that the thrust rod is supported in the clamping shells of the tool headstock. The reaction force is transferred through the solder joint into the toothed belt wheel and the rear seal and is transferred via the bronze rings to the drilling tube and the clamping shells.

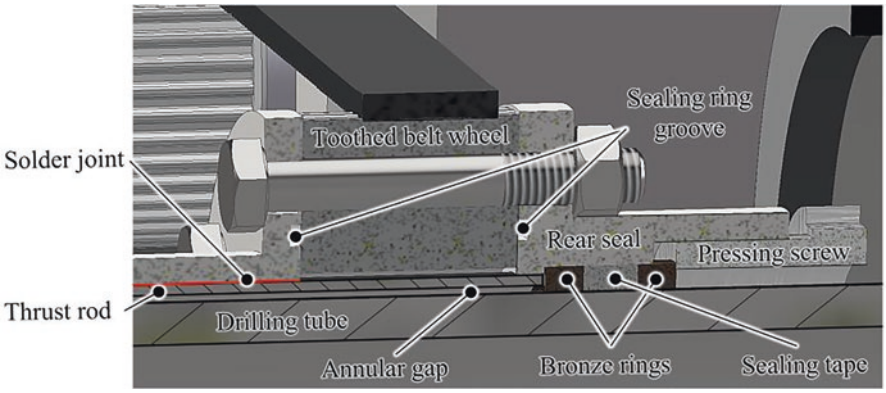


Fig. 6. Sectional view of the thrust rod bearing and the sealing system

Furthermore, a sealing of the system is necessary, because the standard sealing by the oil pressure head does not work because of the thrust rod. Due to the lack of sealing by the oil pressure head, the deep drilling oil would be led to the actuator system through the annular gap between drilling tube and the thrust rod and run out. That's why an an axial and radial seal is provided. The radial seal is ensured by sealing rings on the toothed belt wheel. Axial sealing is provided by sealing tape and a pressing screw.

3.4 Ultrasonic Measuring System

In order to realize an in-process compensation of the straightness deviation it is necessary to be able to determine deviation reliable and with sufficient accuracy. As described in Sect. 2, such a measuring system is currently not available. The ultrasonic measuring system was developed by the cooperating institute IFW in Hannover. The developed measuring system for the compensation system consists of three ultrasonic sensors, which are arranged around the workpiece with radial shift by 30°. The sensors are coupled acoustically with a sleeve. Deep drilling oil is used as contact medium, which is passed through the sleeve to ensure a better connection between the sensors and the workpiece surface. The wall thickness of the workpiece at the measuring position is carried out using the pulse-echo method. Based on the three measured wall thicknesses, three points on the inner wall of the bore are calculated and a compensation circle is used to determine the displacement of the inner circle in relation to the outer circle. The functionality of the measuring concept has already been confirmed and verified by coordinate measuring machine measurements. The maximum deviation between the ultrasonic and the coordinate machine measurements is $W = 0.033$ mm. In order to measure the straightness deviation parallel to the deep drilling process, the sensor sleeve of the measuring system is coupled to the feed slide and carried along on the workpiece surface parallel to the drill head.

4 Conclusion and Outlook

This paper deals with the minimization of straightness deviation in BTA-deep hole drilling. A concept design of an automated in-process compensation of straightness deviation, which can be retrofitted on existing BTA deep drilling machines was designed and constructed. The compensation unit influences the straightness deviation by means of a targeted tilting of the drill head. The functionality of the system could be proven by simulations. In the next step, the components are manufactured and assembled that initial field tests can be carried out. Different evaluation criteria have to be taken into account in the analysis of the experimental test series. The influence of the guide pad positioning movement as well as the guide pad orientation on the resulting straightness deviation must be investigated. Furthermore, the guide pad, as a wear element, represents another important component whose influence must be analyzed. The main focus is on the resulting guide pad wear and the influence on the surface quality of the bore wall as a function of the radial positioning movement. In addition, different guide pad designs with different coatings are taken into account. After evaluation of the experimental investigations, the measuring system is coupled to the actuator system by means of a control loop. The final system should allow a process-simultaneous compensation of the straightness deviation with an accuracy of $acc = 0.1$ mm.

Acknowledgements. The research project was carried out in the framework of the industrial collective research programme (IGF no. 20422 N). It was supported by the Federal Ministry for Economic Affairs and Energy (BMWi) through the AiF (German Federation of Industrial Research Associations e.V.) based on a decision taken by the German Bundestag. The ISF, the IFW and the project-accompanying committee are grateful for the financial support.

References

1. Biermann, D., Bleicher, F., Heisel, U., Klocke, F., Möhring, H.C., Shih, A.: Deep hole drilling. CIRP Ann. **67**, 673–694 (2018)
2. VDI-Richtlinie 3210 – Blatt 1: Deep-hole boring. VDI-Gesellschaft Produktion und Logistik (2006)
3. Hagemeyer, C.: Stütz- und Führungsleistennormalkräfte beim BTA-Aufbohren. WGP-Berich, pp. 38–43 (1990)
4. Dinglinger, E.: Tieflochbohren mit Hartmetall-Bohr- und Reibwerkzeugen bei umlaufendem Werkstück. Werkstatttech. Maschinenbau **42**(4), 122–128 (1952)
5. Dinglinger, E.: Neue Erfahrungen mit Tieflochbohrwerkzeugen. Werkstatttech. Maschinenbau **45**(8), 361–367 (1955)
6. Katsuki, A., Sakuma, K., Tabuchi, K., Onikura, H., Akiyoshi, H., Nakamuta, Y.: The influence of tool geometry on axial hole deviation in deep drilling: comparison of single- and multi-edge tools. JSME Int. J. **30**(265), 1167–1174 (1987)
7. Katsuki, A., Onikura, H., Sakuma, K., Chen, T., Murakami, Y.: The influence of workpiece geometry on axial hole deviation in deep hole drilling. JSME Int. J. Ser. 3 Vib. Control Eng. Eng. Ind. **35**(1), 160–167 (1992)

8. Sakuma, K., Taguchi, K., Katsuki, A., Takeyama, H.: Self-guiding action of deep-hole-drilling tools. CIRP Ann. **30**(1), 311–315 (1981)
9. Kessler, N.: Thermische Mittenverlaufsbeeinflussung beim BTA-Tiefbohren. Dissertation, Technische Universität Dortmund. Vulkan, Essen (2011)
10. Stürenberg, H.O.: Zum Mittenverlauf beim Tiefbohren. Teil 1. TZ Metallbear. **77**(6), 34–37 (1983)
11. Stürenburg, H.O.: Zum Mittenverlauf beim Tiefbohren: Berichte aus dem Institut für Werkzeugmaschinen der Universität Stuttgart. Dissertation, Technischer Verlag Günter Grossmann Stuttgart-Vaihingen (1983)
12. Bleicher, F., Metzner, K., Radkowitsch, W.: Verfahren zum Bohren. Patentschrift EP2730358 B1 (2004)
13. Prince, T.E.: Laser helps bore straight holes: a self-correcting laser guidance system hooked to a boring lathe is helping the U.S. Navy bore straighter holes in gun-barrel forgings for 5 in. guns. Here's how it works. Am. Mach. **9**, 60–61 (1972)
14. Griffiths, B.J.: Axial hole runout during deep drilling (self piloting drilling). In: Proceedings of the Second International Conference on Deep Hole Drilling and Boring, pp. 1–8 (1997)
15. Weinert, K., Hagedorn, M., Kessler, N.: Mittenverlauf beim BTA-Tiefbohren – Untersuchungen zeigen Einflussfaktoren auf die Bohrungsqualität. VDI-Z. Int. Produkt. **147**(11/12), 35–38 (2005)
16. Iovkov, I.: Technologische Analyse des Tiefbohrens mit Minimalmengenschmierung und simulationsbasierte Kompensation des Mittenverlaufs. Dissertation, Technische Universität Dortmund. Vulkan, Essen (2016)
17. VDI-Richtlinie 2221 – Blatt 1: Design of technical products and systems – model of product design. VDI-Gesellschaft Produkt- und Prozessgestaltung (2019)
18. VDI-Richtlinie 2221 – Blatt 2: Design of technical products and systems – configuration of individual product design. VDI-Gesellschaft Produkt- und Prozessgestaltung (2019)

Influence of a Two-Cut-Strategy on Tool Wear in Gear Hobbing

N. Troß[1(✉)], J. Brimmers[1], and T. Bergs[1,2]

[1] WZL of RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
n.tross@wzl.rwth-aachen.de
[2] Fraunhofer IPT, Steinbachstraße 17, 52074 Aachen, Germany

Abstract. This report addresses both the empirical and theoretical analysis of a two-cut strategy for gear hobbing with conventional hobs. Tool life and wear behavior of PM-HSS S390 tools are analyzed in fly-cutting trials when machining 20MnCr5 case-hardening steel. By varying the axial feed rate in the first and the second cut, the influence of the interaction between both cuts on the tool wear is investigated. The results are compared to wear measurements from trials conducted with a one-cut strategy. It was shown that at the underlying test conditions, the average proportion of wear in the second cut was about 4%. For a better understanding of the interaction between the two cuts, load collectives were formed based on characteristic chip values calculated with SPARTApro. The collectives were evaluated with regard to the tool life.

Keywords: Gear hobbing · Tool wear · SPARTApro · Load collective

1 Introduction

Gear hobbing is one of the most relevant manufacturing processes for gears and subject of many research activities [1–5]. A two-cut strategy (roughing and finishing) is used on the one hand for soft fine machining (*finish hobbing*) and on the other hand for general gear soft machining (*hobbing*) [6]. *Finish hobbing* is used to generate the final gear geometry and thus represents the last step in the process chain before heat treatment [7]. Furthermore, without the necessity of a final grinding operation, a complete dry process chain can be realized, which is both economically and ecologically beneficial. In *hobbing*, two cuts are used to manufacture gears with high demands on geometry and surface quality for subsequent process steps [8]. This is especially the case, when a second cut removes surface defects, such as chip welds [9]. Despite the industrial relevance, there is currently no scientifically substantiated process design taking into account the interaction of the individual cuts for conventionally designed hobs. Only an optimized process allows a balanced design of tool life (lower tool costs and efficient use of resources), component quality and productivity (lower costs per part).

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 225–234, 2021.
https://doi.org/10.1007/978-3-662-62138-7_23

2 Objective and Approach

The objective of this report is the analysis of the interaction of the feed rate in the first and the second cut on tool wear in a two-cut strategy in gear hobbing. Wear investigations are performed in fly-cutting trials, using both a two-cut and a single-cut strategy as reference. The feed rate of the first and second cut is varied, while all other parameters are kept constant. Wear behavior and tool life are compared. For a deeper understanding of the interaction between the individual cuts and the overall process, load collectives are derived based on characteristic chip values.

3 Materials and Methods

The experimental setup as well as the gear and tool data are shown in Fig. 1. To investigate the influence of a two-cut strategy on the tool wear, the fly-cutting trial was used as an analogy process for gear hobbing. In the fly-cutting trial, a single hob tooth (fly-cutter) reproduces all of the generating positions of the corresponding hob by a continuous tangential shifting process. The tool is first positioned in z-direction by the approach distance to the workpiece face and in x-direction at the target cutting depth T. Under a continuous rotational speed n_0, the tool then moves tangentially to the workpiece with a tangential velocity v_t. The workpiece is rotating at the same time, while its rotational speed n_2 is synchronized to the movement of the tool. After a shift loop is completed, the tool returns to its starting position, moves parallel to the workpiece axis (z-direction) by the amount of the feed rate f_a and starts the next shift loop. This process is repeated until the gear is completely profiled [10].

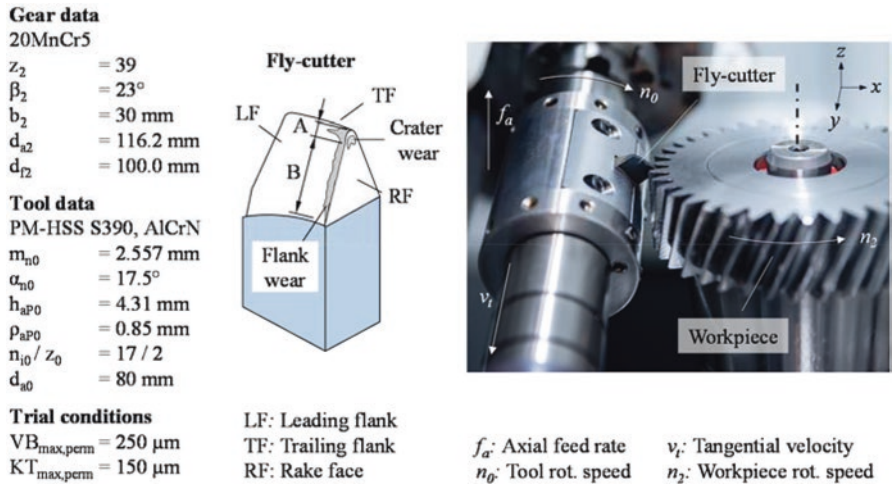


Fig. 1 Experimental setup and conditions

The material used was a case-hardening steel 20MnCr5. The fly-cutters were separated from powder metallurgical high speed steel (PM-HSS) S390 hob with an AlCrN-based coating. The trials were carried out on a LIEBHERR LC180 gear hobbing machine. Between the first and the second cut and after each completely profiled workpiece, the fly-cutter was removed from the machine and the wear was analyzed. The maximum wear mark width VB_{max} was measured manually on the trailing (TF) and leading (LF) flank in areas A and B by means of a KEYENCE VHX-5000 digital microscope. Area A marks the area that is curved due to the tool tip radius ρ_{aP0}. The curvature hinders chip flow, which means that different wear phenomena occur in this area than in the unbent area B. In addition to the flank wear, the maximum crater depth KT_{max} on the rake face (RF) was documented manually by means of a MIKROCAD 3D surface metrology system. As wear criteria a maximum permissible wear mark width $VB_{max,perm} = 250$ µm and a maximum permissible crater depth $KT_{max,perm} = 150$ µm were defined.

Two strategies were investigated: A one-cut strategy, which serves as reference, and a two-cut strategy. For the one-cut strategy, the feed rate was chosen with $f_a = 1.6$ mm (which corresponds to a maximum HOFFMEISTER chip thickness of $h_{cu,max,Hoff} = 0.16$ mm, [11]), $f_a = 2.5$ mm ($h_{cu,max,Hoff} = 0.20$ mm) and $f_a = 3.5$ mm ($h_{cu,max,Hoff} = 0.24$ mm). To achieve the required root diameter $d_{f2} = 100$ mm, a cutting depth of $T = 8.1$ mm was used. For the two-cut strategy, the same feed rates as for the reference trials were defined for the first cut. For the second cut, the feed rates $f_{a2} = 1.0$ mm, $f_{a2} = 2.0$ mm and $f_{a2} = 3.0$ mm where chosen. The examined feed rate combinations can be taken from Fig. 2. The cutting depths were determined with $T_1 = 7.9$ mm for the first cut and with $T_2 = 8.1$ mm for the second cut. As a result, there is a radial stock allowance of $\Delta T = 0.2$ mm between the first and second cut. The cutting velocity was kept constant at $v_c = 200$ m/min for all variants and cuts. All experiments were carried out in climb cutting without the use of cutting fluid.

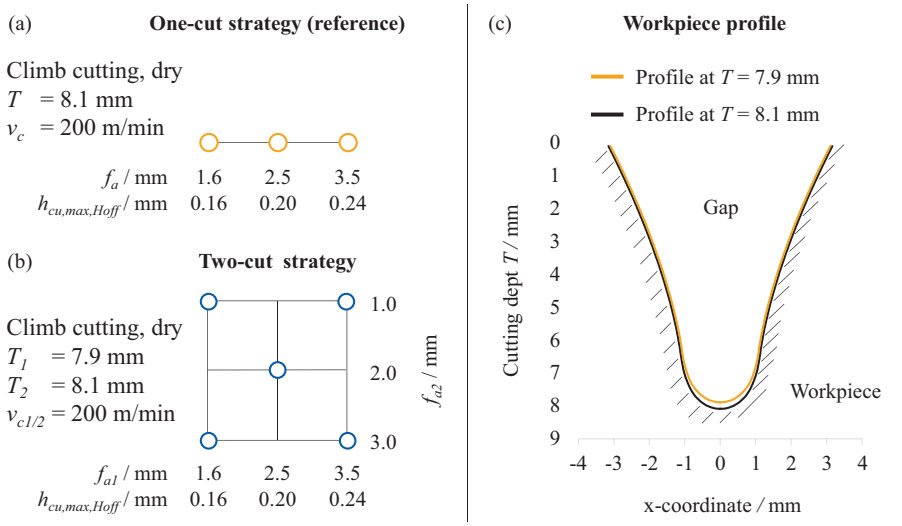


Fig. 2 Design of experiments (a) one-cut strategy (b) two-cut strategy (c) workpiece profile

4 Analysis of the Tool Wear Behavior

The wear curves of the reference trials (one-cut strategy) are shown in Fig. 3. In all cases, the wear criterion was first met by crater wear. Therefore, a representation of the flank wear curves was omitted. The course of the maximum crater depth KT_{max} is plotted over the machined length L. A machined length of $L = 1.3$ m corresponds to one machined gear $N = 1$. The wear curves are supplemented by light microscopic images of the rake face after the tool life was reached. For the reference trials, the highest tool life was achieved at an axial feed rate of $f_a = 3.5$ mm with $L_{KT150} = 24$ m, followed by $L_{KT150} = 21.6$ m at $f_a = 1.6$ mm. The shortest tool life $L_{KT150} = 19$ m was reached at a feed rate of $f_a = 2.5$ mm. In all three cases, crater wear occurred on the rake face on the trailing side in the area of the tool tip radius (TF-A).

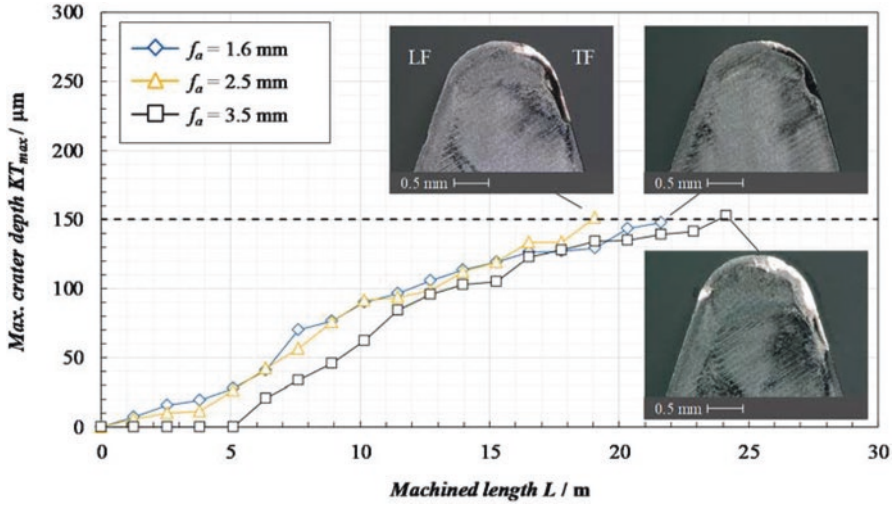


Fig. 3 Wear curves for the reference trials (one-cut strategy)

The tools used in the trials with a two-cut strategy also reached the maximum permissible crater depth $KT_{max,perm}$ in the area TF-A on the rake face. The courses of the maximum crater depth after the second cut KT_{max2} are shown in Fig. 4. A feed rate $f_{a1} = 3.5$ mm in the first cut resulted in the highest ($L_{KT150} = 28$ m) and a feed rate $f_{a1} = 2.5$ mm in the lowest ($L_{KT150} = 18$ m) tool life. Using a feed rate $f_{a1} = 1.6$ mm achieved an intermediate tool life. Compared to the first cut, the second cut seemed to have a secondary influence. Respectively, the highest tool life was achieved with the higher feed rate $f_{a2} = 3.0$ mm in the second cut. The most severe degree of wear on the rake face was observed when using a feed rate of $f_{a1} = 3.5$ mm in the first cut. The similarities in tool life and wear behavior of the reference and the two-cut variants indicate a dominant influence of the first cut on tool wear.

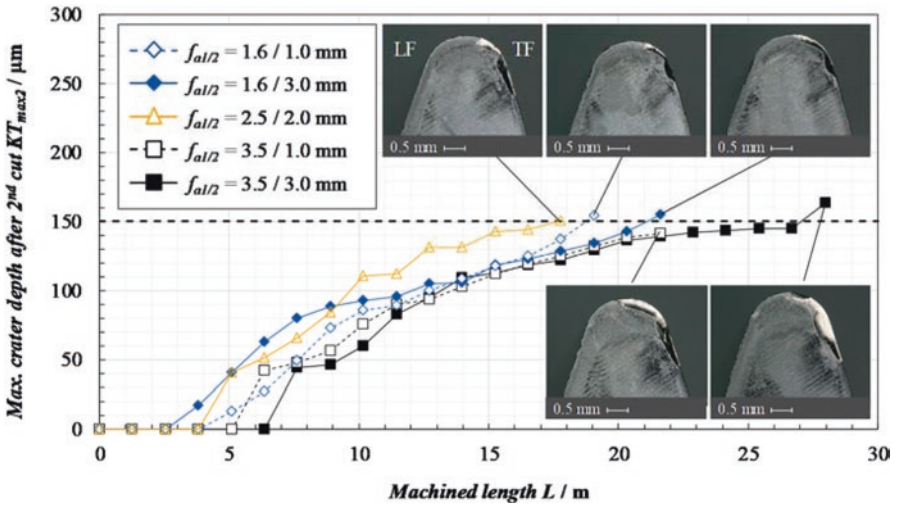


Fig. 4 Wear curves for the two-cut trials

To analyze the share in wear generation of the individual cuts, the measured values of the maximum crater depth after the first (KT_{max1}) and the second (KT_{max2}) cut were compared. For the two-cut trial with $f_{a1} = 2.5$ mm and $f_{a2} = 2.0$ mm, the measured values are represented in Fig. 5.

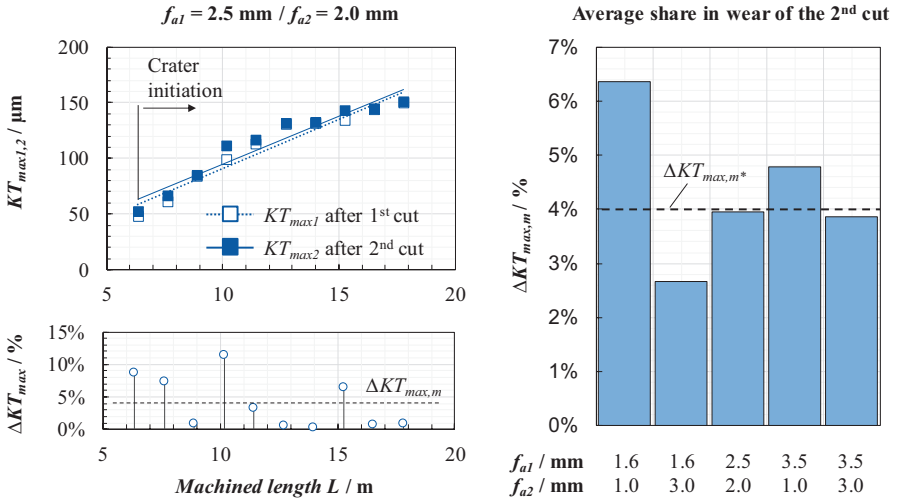


Fig. 5 Average share in wear of the first and second cut

The trend of the two curves is represented by a straight line. Both the measured values and the trend lines differ only slightly from each other. The differences become

smaller towards the end of the trial and the trend lines are approximating. By subtracting KT_{max1} from $KT_{max2,}$ the share in wear of the individual cuts can be quantified. In percentage terms, the average share in wear of the second cut is $\Delta KT_{max,m} = 4\%$. Comparable average values were also determined for the other trials. The average value observed over all tests was calculated with $\Delta KT_{max,m^*} = 4\%$. The results suggest, that, when designing a gear hobbing process, a design of the first cut is sufficient for tool life prediction, while the second cut can be optimized with regard to the required part quality.

5 Load Collectives in Gear Hobbing

In gear hobbing, thermal and mechanical loads occur along the cutting edge depending on the local and temporal cutting conditions. The sum of the individual loads results in an individual load collective for each position x of the tool profile, which consequently causes different degrees of wear along the cutting edge. For a better understanding of the process and a consideration of the locally and time-resolved cutting conditions, the process must therefore be understood as a collective of all individual loads and analyzed as such. For fatigue life calculations, load collectives are usually derived from load-time functions [12]. In the case of gear hobbing, the load can be estimated local- and time-resolved using manufacturing simulations.

5.1 Derivation of Load Collectives

In Fig. 6, the derivation of a load collective based on characteristic chip values is described. With the help of the manufacturing simulation SPARTAPRO, all chip geometries occurring in the process were determined on the basis of a geometric penetration calculation. The software abstracts the workpiece by a defined number of parallel planes, which are penetrated by the tool profile while simulating the machine kinematics. The intersection results in the penetration area, which corresponds to the chip geometry of the respective generating position in the process. Based on the chip geometry, characteristic chip values (e.g. maximum chip thickness $h_{cu,max}$) can be deducted, which allow an evaluation of the load acting on the hob during the cutting process [10, 13].

Using the example of an undeformed chip from an arbitrary generating position j in an arbitrary revolution i, the calculation of the chip values relevant for the load collective is described. The revolution i indicates how often the axial feed f_a was performed, thus, it represents the axial position of the tool at the time of observation. At a position x of the unrolled tool profile, the course of the chip in the cutting direction y is analyzed more closely. The course of the chip thickness $h_{cu}(y)$ can be quantified by the maximum chip thickness $h_{cu,max}$, the mean chip thickness $h_{cu,m}$, the cutting length l_{cu} and the specific chip volume V'_{cu}.

For describing the tool load in gear hobbing, the mean chip thickness $h_{cu,m}$ is an appropriate parameter, while a consideration of the maximum chip thickness is less meaningful due to its unique occurrence [14]. Instead of a time value, the load can be

applied over the travelled distance. Consequently, the cutting length l_{cu} can be used, as it can be cumulated over all cuts to the cumulated cutting length $l_{cu,\Sigma}$. To form the collective, the mean chip thicknesses are sorted into defined classes and plotted over the respective cutting length. The collective is arranged in such a way that the highest class is first and the lowest class last.

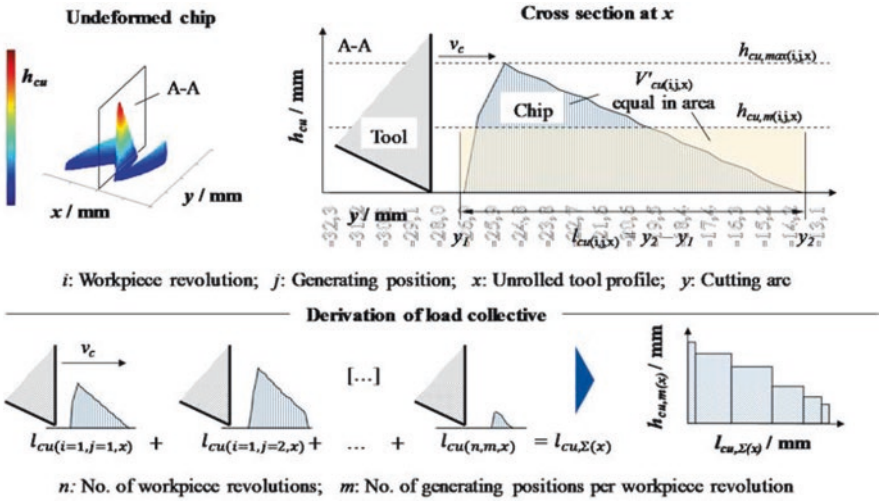


Fig. 6 Derivation of load collectives based on characteristic chip values

5.2 Analysis of Load Collectives

For the investigated variants, the load collectives were determined at the position $x = 1.3$ mm of the unrolled tool profile. At this position, the most severe degree of crater wear was observed. The mean chip thicknesses $h_{cu,m}$ were divided into a total of ten equidistant classes and plotted over the cumulated cutting length $l_{cu,\Sigma}$. The highest class was defined based on the highest occurring mean chip thickness $h_{cu,m}$ of all variants. The same classes were used for all variants.

For the reference variant with $f_a = 2.5$ mm and the two-cut variant with $f_{a1} = 2.5$ mm and $f_{a2} = 2.0$ mm, the determined load collectives are shown in Fig. 7. For the two-cut variant, the collective for the first and the second cut is shown individually and in combination. The load collective for the first cut (a) of the two-cut strategy is almost identical to the collective of the reference variant (d). The classes seem virtually identical, and only slight differences can be observed. The cumulated cutting length $l_{cu,\Sigma}$ is lower for (a) than for (d), while the mean values $h_{cu,m*}$ are identical.

The second cut (b) of the two-cut strategy only forms chips in classes 8 to 10 and has a relatively low cumulated cutting length $l_{cu,\Sigma}$ and mean chip thickness $h_{cu,m*}$. This corresponds to the experimental results, which showed that the second cut had a low influence on wear. Consequently, the combined collective (c) of the first and second cut does not differ much from the collective of the first cut (a) or the reference (d). However, combining the collectives results in a higher cumulated cutting length

$l_{cu,\Sigma}$, due to the increase in classes 8 to 9. The mean chip thicknesses $h_{cu,m*}$ between the reference (d) and the combined collective (c) deviate slightly from each other. In this case, a consideration of the second cut in the combined collective reduces the average load on the one hand, but increases the traveled distance on the other.

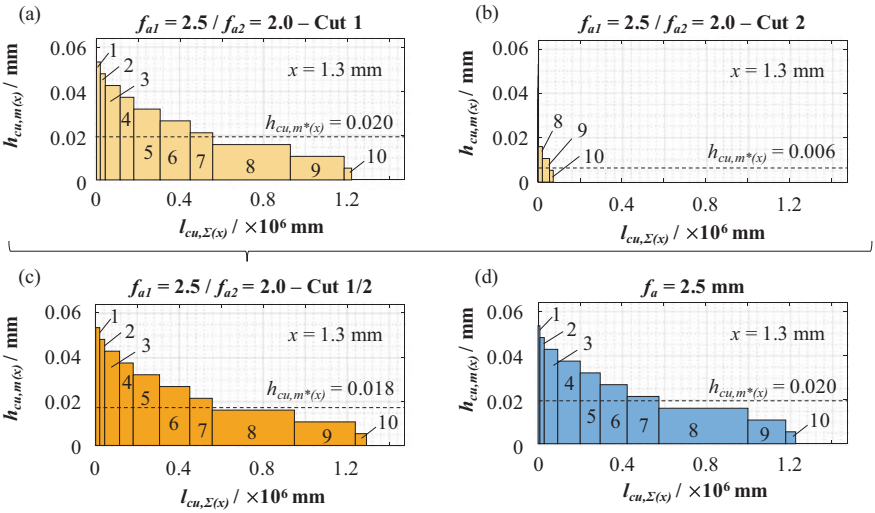


Fig. 7 Comparison of the load collectives at $x = 1.3$ mm. Two-cut strategy: (a) First cut, $f_{a1} = 2.5$ mm (b) Second cut, $f_{a1} = 2.5$ mm (c) Combined. Reference: (d) $f_a = 2.5$ mm

5.3 Comparison of Characteristic Load Collective Values

The mean chip thicknesses $h_{cu,m*}$ and the cumulated cutting lengths $l_{cu,\Sigma}$ were determined for all variants at the position $x = 1.3$ mm of the unrolled tool profile. By multiplying $l_{cu,\Sigma}$ with the respective number of machined gears N, the tool life can be described based on characteristic chip values as the bearable cumulated cutting length $l_{cu,\Sigma,lim}$. In Fig. 8, $h_{cu,m*}$ is plotted over $l_{cu,\Sigma,lim}$. For the two-cut strategy, the collective values were calculated for the combined collective (a) and considering only the first cut (b). The data points are labeled with regard to the used axial feed rate.

When both cuts are considered (a), the respective data points deviate from the reference values in terms of the mean chip thickness $h_{cu,m*}$. A trend can be observed, which indicates a decrease in the bearable cumulated cutting length $l_{cu,\Sigma,lim}$ with increasing mean chip thickness $h_{cu,m*}$. However, the coefficient of determination is $R^2 = 0.41$, which is fairly low. By taking only the first cut into account (b), the data points of the first cut and of the reference are almost identical with regard to the mean chip thickness $h_{cu,m*}$. R^2 increases to $R^2 = 0.63$ and the different load levels become more evident. This reinforces the assumption, that under the present boundary conditions, the second cut has minor or even negligible influence on tool wear. If this correlation is verified in further experiments, it could be used as a modeling approach for tool life predictions. Besides a linear approximation, different model functions,

e.g. power or exponential functions, have to be checked and verified in future investigations.

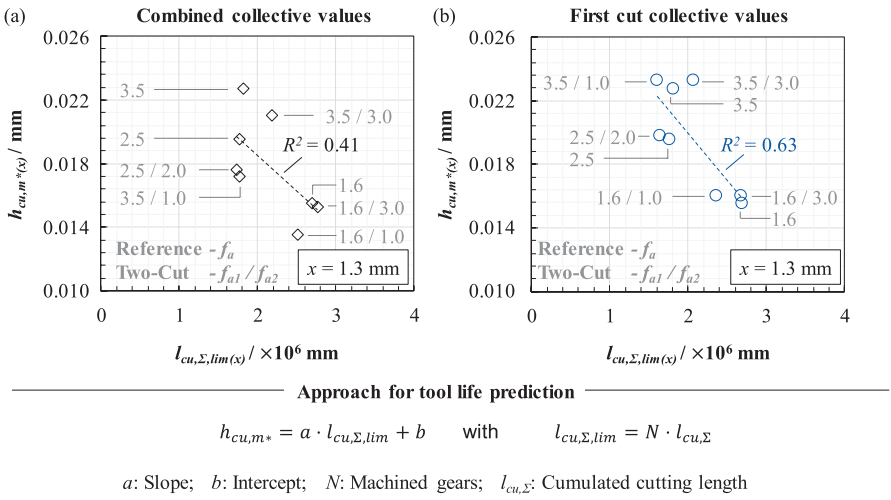


Fig. 8 Correlation of mean chip thickness $h_{cu,m}$ and bearable cumulated cutting length $l_{cu.\Sigma,lim}$ (a) combined collective values (b) first cut collective values

6 Conclusion

The influence of the interaction of the first and second cut on the tool wear in gear hobbing with a two-cut strategy was investigated. Experimental studies were performed in fly-cutting trials. At the underlying test conditions, the share in wear of the second cut was 4% in average. Consequently, the wear behavior and the tool life of tools used with a two-cut strategy was similar to tools used with a one-cut strategy (reference trials). The results suggest, that, when designing a gear hobbing process, a design of the first cut is sufficient for tool life prediction, while the second cut can be optimized with regard to the required part quality.

For a deeper understanding of the interactions between the first and second cut and the process itself, load collectives were derived based on characteristic chip values. An analysis of the collectives at the cutting edge position with the most severe tool wear showed, that only relatively low mean chip thicknesses are generated over a simultaneously short cumulated cutting length. A comparison of characteristic values derived from the load collectives showed a correlation of the load ($h_{cu,m*}$) and the traveled distance ($l_{cu,\Sigma,lim}$). The correlation was more concise when only the first cut of a two-cut strategy was taken into account. This correlation could be used for predicting tool life and has to be verified in further experiments.

Acknowledgements.

The IGF project 20391 of the VDW-Forschungsinstitut e.V. is funded via the AiF as part of the program to promote joint industrial research (IGF) by the Federal Ministry of Economics and Energy based on a resolution of the German Bundestag.

The authors gratefully acknowledge financial support by the WZL Gear Research Circle for the achievement of the project results.

References

1. Rech, J.: Influence of cutting edge preparation on the wear resistance in high speed dry gear hobbing. Wear **261**, 505–512 (2006)
2. Karpuschewski, B., Knoche, H.J., Hipke, M., Beutner, M.: High performance gear hobbing with powder-metallurgical high-speed-steel. CIRP Conf. High Perform. Cut. **1**, 196–201 (2012)
3. Karpuschewski, B., Beutner, M., Köchig, M., Wengler, M.: Cemented carbide tools in high speed gear hobbing applications. CIRP Ann. Manufact. Technol. **66**, 117–120 (2017)
4. Bouzakis, K.D., Kombogiannis, S., Antoniadis, A., Vidakis, N.: Gear hobbing cutting process simulation and tool wear prediction models. J. Manuf. Sci. Eng. **124**(1), 42 (2002)
5. Karpuschewski, B., Beutner, M., Köchig, M., Wengler, M.: Gear hobbing – research activities and state of the art. Adv. Mat. Res. **1018**, 3–12 (2014)
6. Sari, D., Troß, N., Löpenhaus, C., Bergs, T.: Development of an application-oriented tool life equation for dry gear finish hobbing. Wear **426–427**, 1563–1572 (2019)
7. Winkel, O.: Fertigfräsen - heute und morgen. Leistungsfähigkeit und Potenziale der Weichfeinbearbeitung. Seminar “Feinbearbeitung von Zahnrädern”, WZL, Aachen (2007)
8. Sari, D.: Leistungssteigerung des Fertigwälzfräsens unter Berücksichtigung der fertigungsbedingten Bauteileigenschaften. Ph.D. Thesis, RWTH Aachen University, Aachen (2016)
9. Krömer, M.: Entstehung von Spanaufschweißungen beim Trockenwälzfräsen. Ph.D. Thesis, RWTH Aachen University, Aachen (2019)
10. Klocke, F., Brecher, C.: Zahnrad- und Getriebetechnik. Auslegung – Herstellung – Untersuchung – Simulation. Carl Hanser, München (2017)
11. Hoffmeister, B.: Über den Verschleiß am Wälzfräser. Ph.D. Thesis, RWTH Aachen University, Aachen (1970)
12. FKM: Rechnerischer Festigkeitsnachweis für Maschinenbauteile, Frankfurt am Main (2003)
13. Krömer, M., Brecher, C., Löpenhaus, C., Weber, G.-T.: Validierung von spanenden Fertigungssimulationen am Beispiel Wälzfräsen. In: GETPRO. Kongress zur Getriebeproduktion, pp. 128–139. Eigen, Frankfurt a. M. (2015)
14. Mundt, A.: Modell zur rechnerischen Standzeitbestimmung beim Wälzfräsen. Ph.D. Thesis, RWTH Aachen University, Aachen (1992)

Application Potential of Thermoelectric Signals for Temperature Monitoring in Turning of Aluminum Alloys

T. Junge(✉), A. Nestler, and A. Schubert

Chemnitz University of Technology, Professorship Micromanufacturing Technology, Reichenhainer Straße 70, 09126 Chemnitz, Germany
thomas.junge@mb.tu-chemnitz.de

Abstract. Temperature measurement close to the point of its origin is of great importance during chip removal. Therefore, two different temperature measurement methods are used and compared with each other for turning of the aluminum alloy EN AW-2017. On the one hand, cutting temperatures are measured by three thermocouples embedded in the indexable insert. On the other hand, the temperature in the contact area of the tool and the workpiece is detected by a tool-workpiece thermocouple measuring the thermoelectric voltage resulting from the Seebeck effect. For the experiments the cutting speed remains constant while the depth of cut and the feed are varied.

The results show a rise of the cutting temperature with increasing cross-section of the undeformed chip. In comparison, the tool-workpiece thermocouple offers a higher sensitivity while the embedded thermocouples measure higher temperatures for large depths of cut. Hence, the suitability of the methods is affected by the cross-section of the undeformed chip.

Keywords: Seebeck effect · Temperature measurement · Tool-workpiece thermocouple

1 Introduction

Machining processes are affected by the generation of heat during chip removal. The emerging cutting temperature influences the tool life, the degree of the built-up edge formation, and the surface layer properties. Consequently, temperature measurement close to the point of its origin is of great importance for a fast characterization and analysis of process changes. According to Vieregge [1, S. 34] the maximum temperature values occur at the interface of the chip underside and the tool rake face, where the mechanical energy is almost completely converted into thermal energy by chip formation and because of the friction between chip and tool.

Davis et al. [2] and Barrow et al. [3] described different sensor principles for temperature measurement that can be classified in thermochemical, optical and thermoelectric effects. Optical measurement is a contactless approach that records the emitted

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 235–245, 2021.
https://doi.org/10.1007/978-3-662-62138-7_24

radiation during machining with a high temporal resolution by pyrometers [4] or infrared cameras [5]. The problems with the acquisition of the cutting edge temperature are the limited spatial resolution and the impaired local accessibility of the tool area where the maximum temperatures occur. In order to measure temperatures close to its point of origin optical fibers can be placed as near as possible to the contact point between the tool and the workpiece by means of bores in the tool [4].

Other methods for temperature monitoring use thermoelectric effects. Thermocouples are one of the most widespread applications for temperature measurement during machining. The operational principle is based on the Seebeck effect. Two dissimilar electroconductive materials connected in an electric circuit generate an electric voltage if their contact point (hot junction) is heated. The potential difference between the hot and the cold junction is a function of the temperature. The value of the induced voltage depends on the combination of the materials in the thermocouple and is determined by the temperature dependent Seebeck coefficients. It can be calculated by Eq. (1) integrating the resulting Seebeck coefficient S of both active partners (A, B) in the temperature limits (T_1, T_2) between the hot and the cold junction.

$$U_{Th}(T) = U_{Th,B} - U_{Th,A} = \int_{T_1}^{T_2} (S_B(T) - S_A(T))dT \tag{1}$$

Khajehzadeh et al. [6] embedded a thermocouple in the indexable insert to measure the temperature close to the tool rake face during turning of the aluminum alloy EN AW-7075 with and without ultrasonic vibration assistance in the cutting direction. The thermocouple showed a low sensitivity while turning without ultrasonic vibration assistance ($v_c = 30$ m/min, $a_p = 1$ mm, $f = 0.14$ mm) reaching a steady state of about 240 °C just after 20 s. Vibration amplitudes of 8 µm and 10 µm resulted in a higher sensitivity and lower temperatures. In contrast to Khajehzadeh et al., O'Sullivan and Cotterell [7] integrated a pair of thermocouples into the workpiece to measure the temperature during turning of EN AW-6082 aluminum alloy. As a result of increasing tool wear the temperatures rose from 70 °C up to 90 °C ($v_c = 165$ m/min, $a_p = 1$ mm, $f = 0.1$ mm). According to Vieregge [1, S. 34] just a very small percentage of the overall heat is transmitted to the workpiece. Therefore Ay et al. [8] measured higher temperatures up to 100 °C with thermocouples embedded into the indexable insert at nine different positions despite machining with lower cutting parameters ($v_c = 115.3$ m/min, $a_p = 0.25$ mm, $f = 0.1$ mm). The steady-state temperatures distinguished in relation to the sensor position. However, Attia et al. [9] showed in simulative and experimental studies that embedded thermocouples have a strong influence on the heat conduction and distribution within the tool.

An alternative principle to avoid the spatial influence on the heat flux is the tool-workpiece thermocouple that has been the subject of early investigations by Gottwein [10], Shore [11] and Herbert [12]. They formed a dynamic thermocouple measuring the thermoelectric voltage, which was generated by the dissipated heat between the electroconductive tool and the workpiece.

For an exact determination of the temperature a reference calibration is required by heating the tool-workpiece thermocouple and comparing the resulting thermoelectric voltage with a reference thermocouple. Despite the fact that the thermoelectric

voltage corresponds just to the average temperature of the interface of the tool and the workpiece [13, 14] it still provides a temperature measurement close to its point of origin. Consequently, it should result in more accurate temperature values than embedded thermocouples if the generation of secondary thermoelectric voltages can be suppressed. They emerge when the cold junction points at the tool and at the workpiece warm up because of process heat [3]. Thus, the tool and the workpiece have to be electrically insulated from the machine.

Many studies used the tool-workpiece thermocouple method for the examination of process temperatures during metal cutting operations. Santos et al. [15] utilized the tool-workpiece thermocouple to measure the interface temperature in turning of the aluminum alloys EN AW-1350 and EN AW-7075 with a cemented carbide indexable insert. They measured relatively high temperatures of approximately 420 °C (EN AW-1350) and 650 °C (EN-AW 7075) during machining with a cutting speed of 600 m/min, a depth of cut of 1 mm and a feed of 0.3 mm. In their design of experiment they concluded that the cutting temperature rises with increasing cutting speed and feed while it decreases with rising depth of cut. Furthermore, they demonstrated the usability of contact brushes (aluminum alloy) during machining with flood cooling and minimum quantity lubrication. Byrne [16] also found an increase of the average interface temperature up to 330 °C with increasing feed ($v_c = 172$ m/min, $a_p = 1.1$ mm, $f = 0.06$ mm–0.48 mm). However, this rise was nonlinear and decreased for higher feeds.

There are several methods for temperature measurement during machining distinguishing themselves by their temperature range, spatial and temporal resolution, measurement dynamics, calibration procedure, accessibility to the heat source etc. Despite the fact that embedded thermocouples are a widely used, low cost device for measuring local temperatures their measurement values are still delayed depending on the distance to the heat source. Tool-workpiece thermocouples could overcome this disadvantage. Nevertheless, the results of both methods in the current state-of-the-art of science and technology are difficult to compare due to the different tool-workpiece material combinations and process conditions. Therefore, this paper examines both methods simultaneously by implementing them into the experimental setup. Equal process conditions should allow for a comparison regarding the above-mentioned characteristics.

2 Measurement Methods and Experimental Setup

2.1 Specimens and Cutting Tool

For the experimental investigations rotationally symmetric specimens consisting of the aluminum alloy EN AW-2017 (T4) were used. They had a total length of 80 mm and were stepped. For a length of 60 mm the diameter was 37 mm and for the rest of the length the diameter amounted to 25 mm. The shorter side was utilized for clamping using a dead length collet chuck.

Finish machining experiments were conducted with uncoated cemented carbide (HTi10) indexable inserts from the company Mitsubishi Materials. The tools of the type CCGT 09T304-AZ were made for machining of aluminum materials. A chip breaker on the rake face enhanced chip control. Additionally, the polished rake face

decreased built-up edge formation. The tools were characterized by a tool included angle of 80°, a clearance angle of 7°, and a rake angle of 26° near the cutting edge. In connection with the tool holder used the nominal tool cutting edge angle was 95°. To describe the cutting edge microgeometry it was optically recorded by a 3D laser scanning microscope and analyzed at ten orthogonal profile sections along the cutting edge. The mean values of the measurements resulted in a cutting edge radius of 1.6 µm, an edge rounding of S_α 3.35 µm and S_γ 3.39 µm ($K_S = 1$).

2.2 Measurement Principles and Implementation

Two different measurement principles were applied to measure the generated heat close to its point of origin and to analyze their potential for temperature monitoring in turning of aluminum alloys. Furthermore, the setup was extended to measure the components of the resultant force during turning.

In order to predict not only the temperature but also the heat distribution inside the indexable insert three sheath thermocouples (TCs) of the type K were embedded into the tools. Previous investigations revealed a low sensitivity of thick thermocouples with a diameter of 0.5 mm. Consequently, thinner thermocouples with a sensor diameter of 0.245 mm were used and placed as close as possible to the corner of the indexable inserts. The positioning bores perpendicular to the tool reference plane were manufactured by electro-discharge machining and had a diameter of approximate 0.25 mm. Figure 1 shows the arrangement of the bores in the indexable insert. The resulting shortest distance of the sensor tips to the chip breaker geometry was about 95 µm. In combination with a clearance angle of 7° the distance of the bore openings was around 620 µm away from the cutting edge. The three connection sleeves of the sheath thermocouples were held in position by a designed holder device shown in Fig. 1. Additionally, the flexible thermocouple probes with a length of 25 mm (TC3) and 40 mm (TC1, TC2) were adhered to the tool holder.

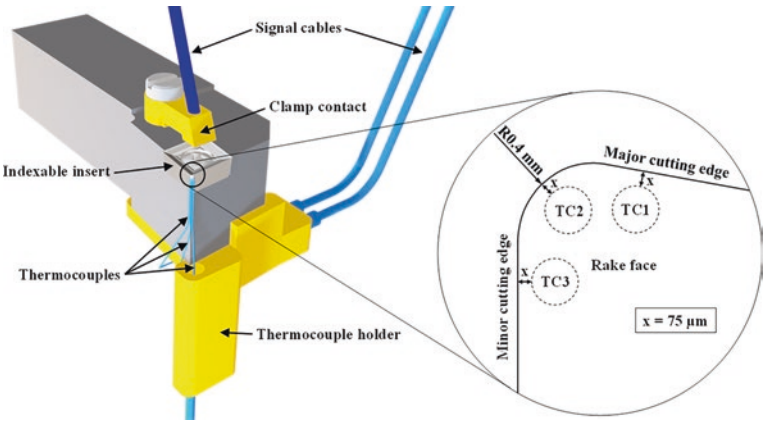


Fig. 1. Tool-side setup for temperature measurement by embedded thermocouples and tool-workpiece thermocouple

The signal cables of the embedded thermocouples were connected to a SCB-68 connector block from National Instruments which represented the cold junction of the setup.

Additionally to the temperature measurement setup, the tool holder was mounted on a three-axis dynamometer of the type 9257 A (Kistler) which was connected to a charge amplifier of the type 5070A (Kistler). Thus, thermomechanical interactions during machining could be investigated.

Furthermore, a tool-workpiece thermocouple was implemented into the experimental setup to investigate the potential of thermoelectric signals for process monitoring. Therefore, the tool and the specimens were electrically connected at thermally independent points. On the one hand, an insulated copper wire was clamped on the indexable insert on the opposite corner of the engagement point of the cutting edge (see Fig. 1). On the other hand, the rotating specimens were connected by two copper-graphite brushes which were positioned by a designed brush holder device close to the dead length collet chuck. The impact of wo tension springs pressed the brushes against the running surface to ensure consistent contact conditions.

The tool holder was electrically insulated by an insulation foil. On the workpiece side, a clamping ring consisting of electrically non-conductive polyoxymethylene was placed between the specimen and the dead length collet chuck. Furthermore, the brush holder was insulated by mounting it onto a bolt that was covered with phenolic paper.

The resultant thermoelectric signals were recorded by a digital multimeter (DAQ6510) from Tektronix with an integrated multiplexer card (type 7700) that enables the measurement of multiple electrical parameters. All signals were subsequently received and processed via the graphical software LabVIEW.

For the conversion of the thermoelectric voltages into the average temperature of the contact surface the tool-workpiece thermocouple had to be calibrated. The Seebeck coefficients of the thermoelectric materials were determined before turning. Therefore, for both materials a small specimen was clamped between two electrodes into the LSR-3 Seebeck measuring device from Linseis. Each specimen was contacted with two thermocouples at a distance of 8 mm between each other. The measurement arrangement was placed into a furnace where the air was replaced by helium 4.7. Under constant thermal ambient conditions, the specimens were heated from room temperature to 600 °C (EN-AW 2017) or 800 °C (cemented carbide), respectively, in steps of 100 °C by a heating coil in the lower electrode. The defined temperature difference over the specimen and the resulting thermoelectric voltage were measured by two thermocouples. Consequently, the temperature dependent Seebeck coefficient could be determined using Eq. (2).

$$S = \frac{U_{Th}}{T_1 - T_2} \tag{2}$$

The Seebeck coefficients calculated from the measured values are displayed as a function of the temperature by a third degree polynomial fit using the method of least squares based on a bi-square weighting (see Fig. 2(a)). The difference between the two curves provided the resulting Seebeck coefficient of the tool-workpiece thermocouple which was applied to Eq. (1). Thus, the integral was solved with the condition of an ambient temperature of $T_1 = 296.15$ K. These constant climatic conditions

were guaranteed by an air conditioning system in the laboratory. Figure 2(b) shows the resulting mathematical relationship between the measured voltage and the interface temperature.

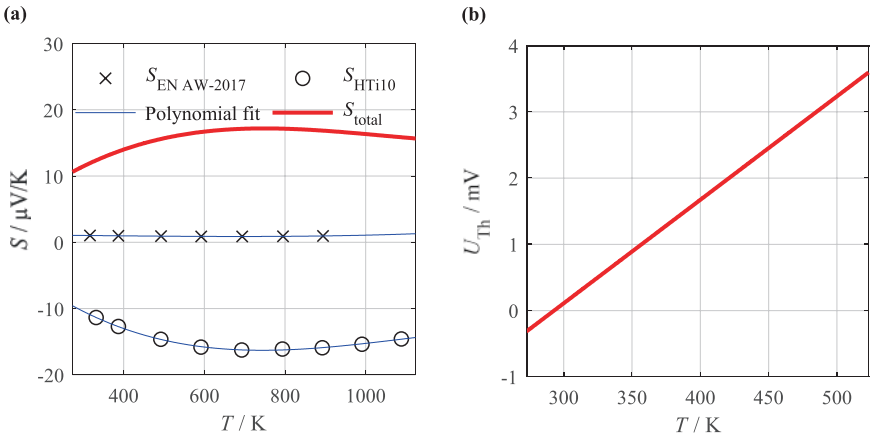


Fig. 2. (a) Determined Seebeck coefficients of the aluminum alloy EN AW-2017, cemented carbide HTi10 and the tool-workpiece thermocouple; (b) Calibration function for the interface temperature between tool and specimen in relation to the thermoelectric voltage

2.3 Machining Experiments

The machining experiments were conducted on a precision lathe of the type SPINNER PD 32. During pre-machining the feed (0.1 mm), the depth of cut (0.5 mm), and the cutting speed (150 m/min) were kept constant on a machining length of 50 mm keeping a cylindrical section of 10 mm as cold junction for the electrical brush contact. To avoid a significant heating of the specimen, the specimen was cooled down with a cold air nozzle.

After pre-machining the specimens had a diameter of 31 mm. In the finish machining experiments, the influence of the feed and the depth of cut on the temperature development close to the shear zone were investigated. Therefore, the feed was varied between 0.04 mm and 0.2 mm and the depth of cut was changed between 0.4 mm and 2 mm. Both parameters were increased in four equal-sized steps within this range. All levels of both parameters were combined using a full factorial design of experiments. In earlier investigations [17] the influence of the cutting speed on the thermoelectric signals was already examined. Thus, the cutting speed was kept constant at 300 m/min. Due to the measurement setup the experimental investigations were performed without any lubrication. For each combination of feed and depth of cut, three specimens were machined.

3 Results and Discussion

Within the scope of the experiments general statements can be made about the course of the temperature signals acquired by the embedded thermocouples and the tool-workpiece thermocouple during one machining operation. The measurement curves of the embedded thermocouples show a degressive increase towards a static-state temperature (see Fig. 3). As the feed and the depth of cut increase, the rise of the curve becomes steeper, so that a state of equilibrium between the supplied and the dissipated heat is reached after approximately one second.

In contrast, the tool-workpiece thermocouple measures the temperature at the interface of the tool and the workpiece where most of the heat is generated. Hence, there is no time delay whereby the thermoelectric signal rises with a steep edge at the point of the first tool engagement and shows a steady course during machining (see Fig. 3). Nevertheless, for a combination of feeds smaller than 0.2 mm and depths of cut between 1.2 mm and 2.0 mm almost no chip breaking occurs. Consequently, the resulting long and hot chips can touch the brush device holder at the end of the machining and therefore reduce the temperature signal due to the heat supply at the cold junction point.

Against the odds, TC1 does not measure the highest temperatures during machining despite being located closest to the shear zone. Instead, TC3 often measures the highest temperatures while being located next to the minor cutting edge (see Fig. 3). These results can be explained by the rise of the components of the resultant force with an increasing cross-section of the undeformed chip. Therefore, the cutting edge is exposed to high mechanical loads that partly lead to the cutting edge breakage at the position of TC1 (see Fig. 4). Hence, the bores in the indexable insert weaken the mechanical stability of the tool. Additionally, TC1 and TC2 are sometimes ripped out of their bores by long chips despite being adhered to the tool holder. Because of these unstable process conditions only the temperature values of TC3 can be evaluated in conjunction with all process parameter settings.

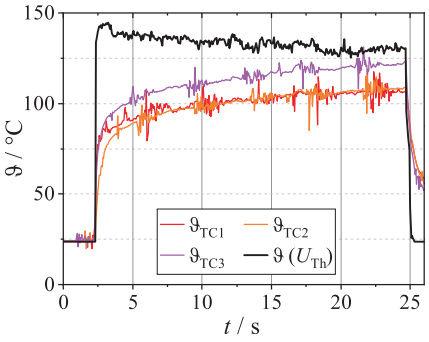


Fig. 3. Signal curves of the temperature ($a_p = 1.2$ mm, $f = 0.04$ mm) measured by the embedded thermocouples (ϑ_{TC1}, ϑ_{TC2}, ϑ_{TC3}) and the tool-workpiece thermocouple ($\vartheta(U_{Th})$) during machining

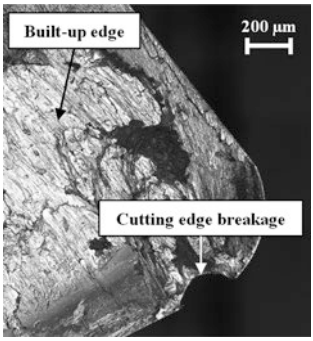


Fig. 4. Image of cutting edge breakage at the position of TC1 and the built-up edge on the rake face

Figure 5 displays the mean values of the temperature over the machining time for the tool-workpiece and the embedded thermocouple TC3 considering only the time range of the static-state temperatures. The error bars consider the range of the mean values of the three tests. The results show higher temperature fluctuations for the tool-workpiece thermocouple that could result from the formation of a built-up edge on the rake face (see Fig. 4) that might change the contact conditions and the heat distribution between the tool and the workpiece.

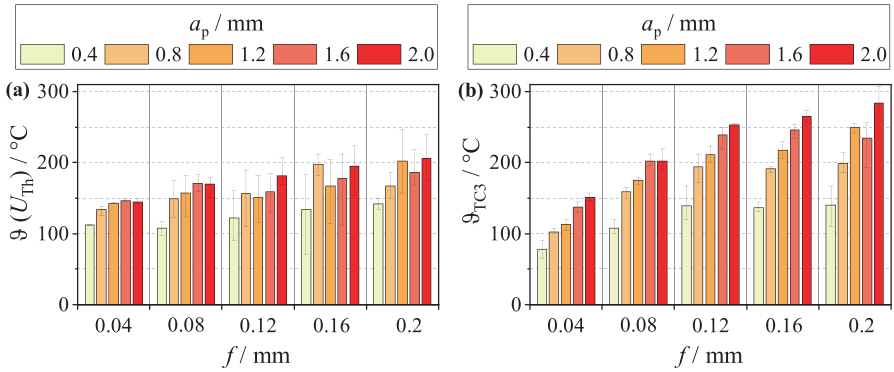


Fig. 5. Influence of the feed and the depth of cut on the temperature measured by (a) the tool-workpiece thermocouple ($\vartheta(U_{Th})$) and (b) the embedded thermocouple TC3 (ϑ_{TC3})

For both measurement methods the temperature rises with increasing feed and depth of cut. This effect can be explained by the increase of the cutting power P_c which is almost completely converted into thermal energy (see Fig. 6(a)). Despite an approximately linear growth of the cutting power the measured temperatures show a degressive growth with increasing feed due to a higher heat dissipation over the outflowing chip.

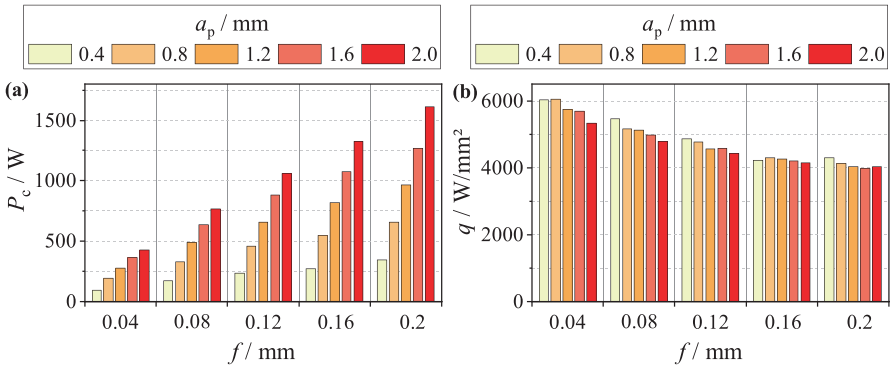


Fig. 6. Influence of the feed and the depth of cut (a) on the cutting power P_c and (b) the heat flux q

In connection with small cross-sections of the undeformed chip the temperatures measured by the tool-workpiece thermocouple are higher or equal to the temperatures measured by TC3. However, with increasing depth of cut the local temperature near the cutting edge rises stronger than the thermoelectric values. Their results differ from each other due to the different measurement principles. The tool-workpiece thermocouple measures a temperature at the contact surface and the embedded thermocouple measures a local temperature at one particular point. The heat flux q shown in Fig. 6(b) represents the rate of thermal energy that passes through the cross-section of the undeformed chip. With an increasing cross-section of the undeformed chip there is a decline of the heat flux. This decrease in thermal power per surface area results in a lower mean temperature at the interface of the tool and the workpiece compared to the local temperature near the cutting edge.

4 Summary and Conclusion

The results of the experiments show that the temperature measurement with the tool-workpiece and the embedded thermocouple are both affected by the cross-section of the undeformed chip. However, the temperature changes differ due to the specific measurement principles.

Although embedded thermocouples can measure higher temperatures under certain process conditions than the tool-workpiece thermocouple they have to be placed very close to the cutting edge. The experimental results show that the reduction of the spatial distance to the cutting edge can affect the stability of the cutting edge as well as the measurement. However, an increasing distance to the area of heat generation would further reduce the sensitivity of the embedded thermocouples. Thus, the method is probably not as suitable as the tool-workpiece thermocouple for monitoring large temperature changes in a short time.

The tool-workpiece thermocouple seems well suited for a real-time process monitoring due to its high sensitivity. Especially in the finish machining of parts where the depths of cut and feeds are usually lower the thermoelectric signals can represent very accurately the process temperatures close to their point of origin without changing the heat flux inside the tool. However, the application of the tool-workpiece thermocouple demands the elimination of secondary thermoelectric voltages and short circuits which is challenging if long chips contact the brush holder device. As a result, the copper-graphite brushes could be replaced by fiber brushes which contact the specimen in a groove at the external diameter. This setup is more compact and avoids obstructions of the chip flow. Further investigations should exhibit the functionality of this setup and the influence of changing contact conditions introduced by wear on the resulting thermoelectric signals.

The findings provide important information for the application of thermoelectric signals for a precise and highly sensitive temperature measurment that could enable a real time monitoring of the process changes and surface layer properties.

Acknowledgements. The scientific work has been supported by the DFG within the research priority program SPP 2086. The authors thank the DFG for this funding and intensive technical support.

References

1. Vieregge, G.: Zerspanung der Eisenwerkstoffe, 2nd edn. Stahleisen, Düsseldorf (1970)
2. Davies, M.A., Ueda, T., M'Saoubi, R., et al.: On the measurement of temperature in material removal processes. CIRP Ann. **56**(2), 581–604 (2007)
3. Barrow, G.: A review of experimental and theoretical techniques for assessing cutting temperatures. CIRP Ann. **22**(2), 203–211 (1973)
4. Ueda, T., Sato, M., Hosokawa, A., Ozawa, M.: Development of infrared radiation pyrometer with optical fibers - two-color pyrometer with non-contact fiber coupler. CIRP Ann. **57**, 69–72 (2008)
5. Vernaza-Peña, K.M., Mason, J.J., Li, M.: Experimental study of the temperature field generated during orthogonal machining of an aluminum alloy. Exp. Mech. **42**, 221–229 (2002)
6. Khajehzadeh, M., Akhlaghi, M., Razfar, M.R.: Finite element simulation and experimental investigation of tool temperature during ultrasonically assisted turning of aerospace aluminum using multicoated carbide inserts. Int. J. Adv. Manuf. Technol. **75**, 1163–1175 (2014)
7. O'Sullivan, D., Cotterell, M.: Temperature measurement in single point turning. J. Mater. Process. Technol. **118**, 301–308 (2001)
8. Ay, H., Yang, W.J.: Heat transfer and life of metal cutting tools in turning. Int. J. Heat Mass Transf. **41**, 613–623 (1998)
9. Attia, M.H., Cameron, A., Kops, L.: Distortion in Thermal field around inserted thermocouples in experimental interfacial studies – Part 4: End effect. J. Eng. Ind. **124**, 135–145 (2002)
10. Gottwein, K.: Die Messung der Schneidentemperatur beim Abdrehen von FIußeisen. Ma-schinenbau – Der Betrieb. Z. Gestalt. Betr. Wirtsch. **4**(23), 1129–1135 (1925)
11. Shore, H.: Thermoelectric measurement of cutting tool temperatures. J. Wash. Acad. Sci. **15**(5), 85–88 (1925)

12. Herbert, E.G.: The measurement of cutting temperatures. Proc. Inst. Mech. Eng. **110**(1), 289–329 (1926)
13. Hehenkamp, T.: Untersuchungen über den elektrisch kompensierbaren Verschleiß an Drehmeißeln aus Hartmetallen. Archiv Eisenhüttenwes. **29**(4), 249–256 (1958)
14. Stephenson, D.A.: Tool-work thermocouple temperature measurements – theory and implementation issues. J. Eng. Ind. **115**(4), 432–437 (1993)
15. Santos, M.C., Araújo Filho, J.S., Barrozo, M.A.S., et al.: Development and application of a temperature measurement device using the tool-workpiece thermocouple method in turning at high cutting speeds. Int. J. Adv. Manuf. Technol. **89**(5), 2287–2298 (2017)
16. Byrne, G.: Thermoelectric signal characteristics and average interfacial temperatures in the machining of metals under geometrically defined conditions. Int. J. Mach. Tools Manuf. **27**(2), 215–224 (1987)
17. Junge, T., Liborius, H., Mehner, T., Nestler, A., Schubert, A., Lampke, T.: Method for process monitoring of surface layer changes in turning of aluminium alloys using tools with a flank face chamfer. Procedia CIRP **87**, 432–437 (2020)

Modeling of Contact Conditions and Local Load in Bevel Gear Grinding

M. Solf[1(✉)], J. Brimmers[1], and T. Bergs[1,2]

[1] Laboratory for Machine Tools (WZL) of RWTH Aachen University, Campus Boulevard 30, 52074 Aachen, Germany
m.solf@wzl.rwth-aachen.de
[2] Fraunhofer Institute for Production Technology (IPT), Steinbachstraße 17, 52074 Aachen, Germany

Abstract. Grinding is an established process for gear manufacturing, as good geometric and surface quality can be achieved. For bevel gears, grinding is used in case of high demands on accuracy and reproducibility. In industry, design of bevel gear grinding processes is usually based on experience. An efficient design of grinding processes can be performed based on the cutting force. Knowledge of the cutting force is necessary to predict the process influence on the workpiece and the wear of the grinding tools. For bevel gear grinding, no cutting force models exist. To model the cutting force in grinding processes, the contact conditions must be known. In this report, a model of the geometric contact conditions in bevel gear grinding is presented. The model is validated by comparing the simulated bevel gear flank with the ideal flank. Finally, the relation between simulation and measured process loads is analyzed.

Keywords: Bevel gear grinding · Force modeling · Efficient process design

1 Introduction and Motivation

Due to increasingly high demands regarding load-carrying capacity and noise excitation behavior of gear drives, hard finishing of gears has become a necessary process step in many cases [1]. Hard fine machining by means of grinding is an established process for the manufacturing of gears for a wide range of applications, as high geometric and surface quality can be achieved [2]. For bevel gears, grinding is used especially in case of high demands on geometric quality and reproducibility, such as for automotive transmission applications [3].

In the industrial environment, the design of bevel gear grinding processes is usually performed based on experience [4]. Suitable process parameters are determined for each workpiece geometry and tool specification in empirical studies. It is not known whether the derived process parameters are in the range of the maximum productivity. An efficient design of productive bevel gear grinding processes can be based on the cutting force. Knowledge of the cutting force is necessary to predict the

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 246–255, 2021.
https://doi.org/10.1007/978-3-662-62138-7_25

thermal influence on the workpiece as well as the wear of the grinding tool [5]. The cutting force therefore plays a decisive role in a knowledge-based design of grinding processes.

2 State of the Art

In the field of bevel gear manufacturing, the cutting process has been the main focus of research in the past years [6–8]. For plunging bevel gear grinding, an empirical model to predict the risk of grinding burn considering the process parameters was developed [4]. However, the results are only valid for limited boundary conditions. Models for predicting the cutting force in bevel gear grinding do not exist so far.

A model that is frequently used in grinding processes is the cutting force model according to Werner [9]. According to this model, the normal grinding force is proportional to the chip cross-section and the chip thickness [9]. This relation is in agreement with other cutting force models, such as the model according to Kienzle [10]. The same proportionality was later also used in other grinding force models, such as those by Bock or Jankowski [11, 12]. The investigations of Grinko have shown that the cutting force is also proportional to the chip cross section in discontinuous profile gear grinding [13]. Furthermore, Scheffler already successfully transferred the approach of Werner to continuous generating gear grinding [9, 14].

In addition to the proportionality to chip thickness and chip cross section, other characteristic geometric process values are also assumed to be proportional to the cutting force in other grinding force models. Thus, Masslow and Brückner found proportional relations between the material removal rate and the cutting force [15, 16]. The same relation was later shown by Rütti and Kampka for gear honing [17, 18].

Overall, it can be seen that the exact contact conditions between tool and workpiece are required to model the grinding force. In earlier investigations, the applicability of the Werner model for plunging bevel gear grinding was analyzed by approximating the material removal [19]. Due to the complex contact conditions, it has not yet been possible to carry out a precise examination of the relation between the geometric process values and the process forces for plunging or generating bevel gear grinding.

3 Objective and Approach

The prediction of the cutting force is of fundamental relevance for a knowledge based grinding process design [5]. Therefore, the objective of the investigations presented in this report is to check the transferability of existing approaches for the cutting force calculation to bevel gear grinding. The prediction of the cutting force in grinding processes is usually based on the geometric contact conditions between grinding wheel and workpiece. To date, the contact conditions in bevel gear grinding are usually determined in a simplified way by means of analytic equations, averaged over the process duration [20]. In the following, this knowledge deficit shall be eliminated.

In order to derive the contact conditions, the process kinematics of plunging and generating bevel gear grinding are implemented in a numerical model. Out of the

intersection between grinding wheel and bevel gear, geometric process values such as material removal and chip cross section can be determined for every time step and every location on the grinding wheel profile and on the tooth flanks. In this report, the simulation model will be validated based on a comparison of the simulated tooth gap geometry with nominal data. Subsequently, it is checked whether the cutting force in bevel gear grinding can be predicted based on the geometric contact conditions in combination with existing force models. In the future, modeling of the local cutting force out of the local geometric contact conditions in bevel gear grinding shall enable the optimization of process design as well as a prediction of thermal damage and tool wear.

4 Modeling of Bevel Gear Grinding

Many established approaches to predict the cutting force are based on geometric process values. For this reason, the contact conditions in bevel gear grinding are to be modeled in order to enable the determination of the geometric process values. The simulation model is presented in the following.

4.1 Introduction of the Simulation Model

An established method to predict the geometric contact conditions in gear manufacturing is the numerical penetration calculation. Penetration calculations have already been developed at WZL for gear hobbing [21], generating gear grinding [14] and bevel gear cutting [6]. The method was transferred to bevel gear grinding within the scope of the investigations presented. The simulation procedure is shown in Fig. 1.

To model bevel gear grinding, a pre-machined discretized tooth gap geometry and a bevel gear data set in Klingelnberg neutral data format are required. The pre-machined tooth gap discretized into a point cloud arranged in planes can be generated by a simulation of bevel gear cutting or deep grinding. The neutral data include the geometry of the ground bevel gear and the grinding wheel as well as the process kinematics. Out of the neutral data, a discretized grinding wheel is calculated.

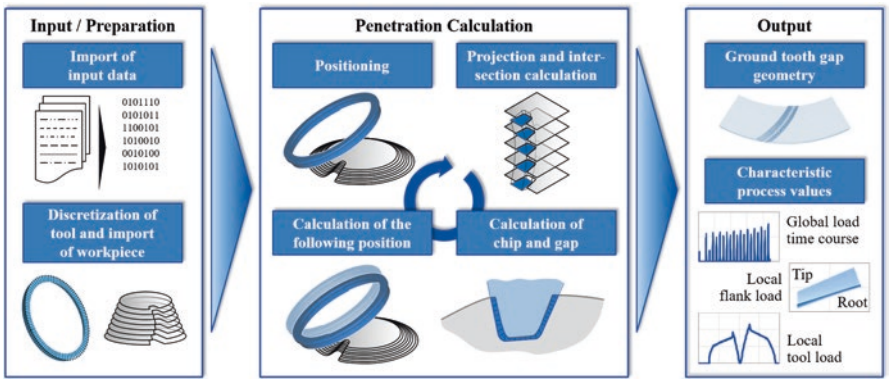


Fig. 1. Modeling of bevel gear grinding process with simulation BevelGrind

The pre-machined bevel gear gap and the grinding wheel are positioned in relation to each other based on the process kinematics, center of Fig. 1. Then, the intersection of the bevel gear and the grinding wheel is determined. By projecting the cutting edge onto the individual discretized planes of the gap, the three-dimensional intersection is reduced to a two-dimensional problem. The intersection points between the planes of the gap and the projected grinding wheel contour are determined. Based on this, the determination of the cut cross-section and the new gap geometry is carried out for the individual tooth gap planes. Subsequently, the tooth gap and grinding wheel are repositioned for the next time increment until the full infeed is achieved.

The penetration calculation is followed by the evaluation of the characteristic geometric values of the bevel gear grinding process such as the chip thickness or the material removal rate. The geometric values are evaluated either globally, as course over the process time, or locally, related to the grinding wheel profile or the tooth flank.

4.2 Validation of the Simulation Model

The manufacturing simulation of bevel gear grinding shall be used to predict the geometric process values. To enable an accurate prediction, it must be ensured that the model represents the geometric relations with sufficient accuracy. The validation of the prediction will be done by comparing the simulated tooth flank geometry with the ideal geometry. For this purpose, a method is applied that is based on a procedure previously used for the validation of the bevel gear cutting simulation BevelCut [22].

As input data for the validation, the simulated tooth gap and nominal measurement data are required. The nominal measurement data are generated in the bevel gear design software and represent the ideal flank geometry. The simulated flanks as well as the nominal measurement data are given as point clouds with different point distribution. First, the points that represent the flanks of the simulated gap are adapted to the point distribution of the nominal measurement data by means of interpolation [22]. Then, the validation is carried out with a pointwise distance calculation between nominal measurement data and simulated flanks. The result of the validation for the simulation of plunging bevel gear grinding is shown in the upper half of Fig. 2.

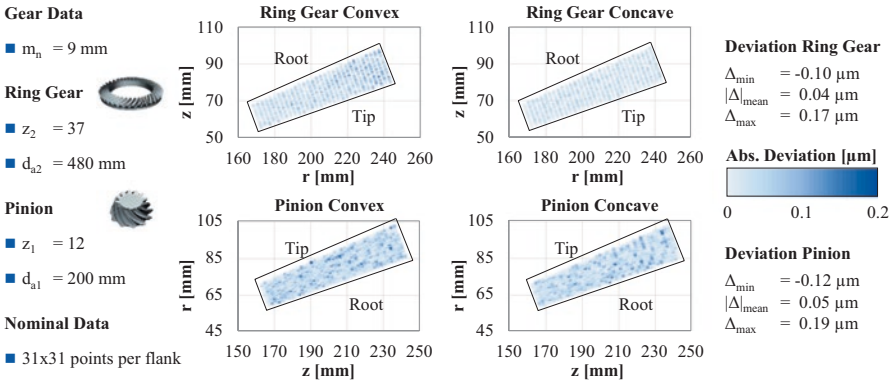


Fig. 2. Validation of bevel gear grinding simulation

On the left, the deviations are shown for the convex flank, on the right for the concave flank. Light colors correspond to a low deviation, dark colors to a high deviation between simulation and nominal measuring data. The comparison shows that the largest deviation occurring is $\Delta_{max} = 0.17$ µm. The mean value of the absolute deviation is $|\Delta|_{mean} = 0.04$ µm. The comparison with the nominal measurement data shows that modeling of plunging bevel gear grinding is performed with high accuracy.

Afterwards, the simulated pinion geometry was compared with the corresponding nominal measurement data. The resulting deviations are shown in the bottom of Fig. 2. For the pinion, the maximum deviation is $\Delta_{max} = 0.19$ µm. The average deviation is $|\Delta|_{mean} = 0.05$ µm. The results show that generating bevel gear grinding is modeled with high accuracy. Since the accuracy of the simulation correlates with the temporal and spatial resolution, a reduction of the deviations could be possible. However, as the proven accuracy is comparable to the accuracy of common gear measuring machines it is assumed to be sufficient for a prediction of the force and the local loads.

5 Comparison of Simulation and Measurements

In the following, the simulated course of the material removal rate will be compared with measurement results. In bevel gear grinding, measurement of the cutting force is only possible after a complex adjustment of the gear clamping, which can also have an effect on the system stiffness. As an alternative to the measurement of the cutting force, process monitoring of grinding processes can be based on the measurement of the tool spindle power [23]. For grinding processes, the tool spindle power is commonly linear proportional to the tangential cutting force and can be determined from the machine tool control or by means of external power sensors [24, 25]. In the following chapters, tool spindle power measurements performed with a Montronix power sensor will be shown. Considering existing force models, the material removal rate and the measured tool spindle power are expected to show a proportional relation.

The spindle power measurements were carried out with a within the project DFG KL 500/187-1 at Scania CV AB in Södertälje, Sweden. The tests were performed on a Gleason PHOENIX®II 600G bevel gear grinding machine with Saint-Gobain grinding wheels that contained rod-shaped sintered corundum (Altos) grains. The test gears were manufactured from conventional case hardening steel. Pinion grinding was performed by generating and ring gear grinding by a plunging process with vector feed and eccentric motion.

5.1 Correlation of Simulation and Measurement for Plunging

The comparison of the simulated material removal rate with the measured course of the tool spindle power over the machining of one gap is first done for plunging bevel gear grinding. The result of the comparison is shown in Fig. 3. As already shown in previous investigations, the spindle power averages the eccentric-related variation [24]. Therefore, the course of the spindle power is compared with the average value

of the material removal rate. In the simulation, the average material removal rate increases rapidly at the beginning of the process, but then remains almost constant. However, the measurement shows that the tool spindle power increases over the entire process duration, see right side of Fig. 3.

The increasing power in plunging bevel gear grinding despite theoretically approximately constant material removal was attributed to a change in the contact conditions due to elasticity in previous investigations [19]. Elastic deformation of grinding wheel and machine tool can lead to a significant change of the effective contact conditions compared to the theoretical contact conditions in an ideally rigid system [26, 27]. From the measurement it is also evident that the spindle power increases throughout a time period that significantly exceeds the theoretical grinding duration Δt. This suggests that the spindle power in the process already increases before contact between grinding wheel and bevel gear occurs. This increase in spindle power before initial contact was attributed to oil friction in the contact zone in previous investigations [28].

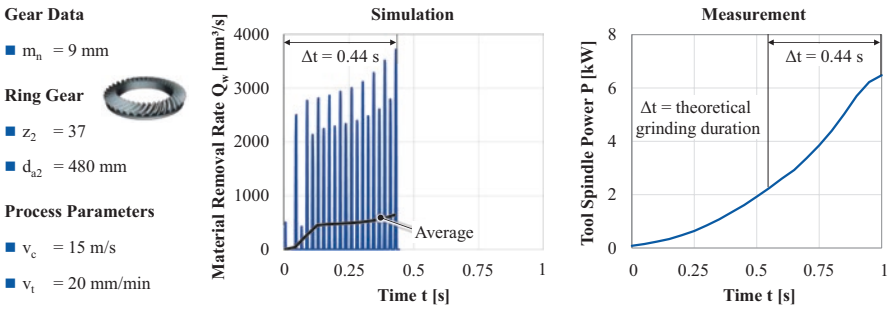


Fig. 3. Comparison of simulation and measurement for plunging bevel gear grinding

Overall, the comparison of the calculated material removal rate and the measured spindle power supports earlier investigations, according to which it was assumed that elasticity-related changes during plunging bevel gear grinding lead to an excessive increase in tool spindle power and cutting force [19, 24, 28]. In the future, it will be necessary to investigate more closely how the contact conditions in plunging bevel gear grinding change due to elasticity during the machining of one tooth gap.

5.2 Correlation of Simulation and Measurement for Generating

In this chapter, the simulated material removal rate and the spindle power are compared for generating bevel gear grinding. The result of the comparison for grinding of one gap is shown in Fig. 4. In contrast to plunging bevel gear grinding, the courses for generating bevel gear grinding show a high similarity. Both the material removal rate and the spindle power increase linearly up to the maximum value and then drop linearly back to zero with a comparable gradient. The simulated process duration is also equal to the measured duration.

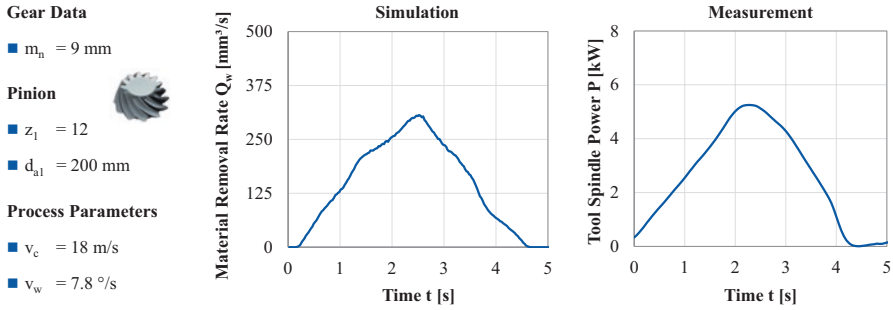


Fig. 4. Comparison of simulation and measurement for generating bevel gear grinding

The results show that for generating bevel gear grinding, the spindle power and thus the tangential cutting force is proportional to the material removal rate. This is in agreement with common force models [15–18]. From the determined correlation, it can be seen that the discrepancy between material removal rate and spindle power in plunging bevel gear grinding is a process phenomenon and cannot be explained by a general behavior of the bevel gear grinding machine.

The examinations reinforce the assumption that in plunging bevel gear grinding, the elasticity in combination with the special geometry lead to a change in contact conditions that is not considered in current force models. In the following, it will be examined in more detail to what extent the force in generating bevel gear grinding can be predicted using the developed numerical simulation model.

6 Modeling of the Cutting Force for Generating

As previously explained, the cutting force in grinding processes is often assumed to be proportional to the material removal rate. The material removal rate is linear proportional to the feed rate, which for generating bevel gear grinding is represented by the rolling velocity v_w. The cutting speed v_c has no direct influence on the material removal rate. Rütti proposed the following equation for calculating the cutting force F(t) from the material removal rate Q_w and the cutting speed v_c, see Equation (1) [17]. The coefficient K represents influence factors such as grinding wheel, coolant and temperature, which are assumed to be constant.

$$F(t) = K \cdot \frac{Q_w(t)}{v_c} \tag{1}$$

As the components of the cutting force are commonly approximately proportional to each other, it is assumed that the tangential component of the cutting force can also be calculated according to the equation. In Fig. 5, the tangential cutting force calculated with Equation (1) from the simulated material removal rate Q_w and the cutting speed v_c is compared to the tangential cutting force calculated from the measured spindle power. The coefficient K was determined based on one measurement and was not

adjusted afterwards. In the future, the influence factors on the coefficient K will be analyzed and determined for different process specifications.

The comparison shows that the equation according to Rütti in combination with the simulation model of generating bevel gear grinding can be used to predict the influence of the rolling velocity v_w and the cutting speed v_c on the course of the cutting force with good agreement. Thus, the objective to predict the cutting force could successfully be realized for generating bevel gear grinding.

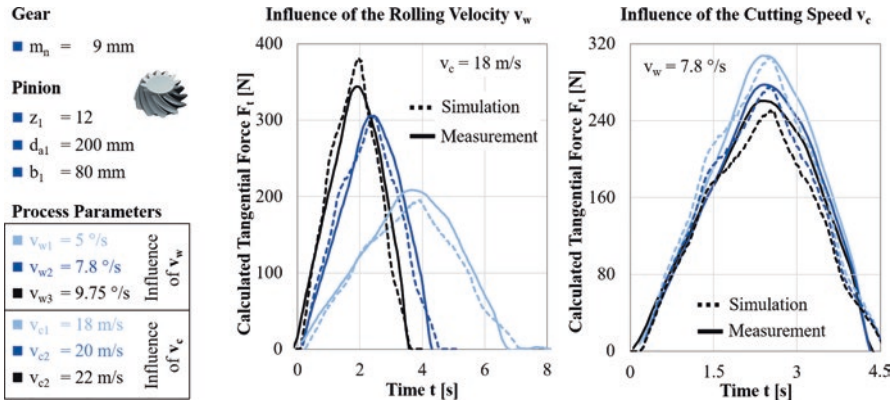


Fig. 5. Comparison of calculated force and measurement for generating bevel gear grinding

Based on the simulated local material removal rate related to the gear flank or the grinding wheel profile, the developed simulation in combination with the force model can be used for a prediction of local loads on bevel gear and grinding wheel. This can enable the prediction of local thermal damage and grinding wheel wear in the future. For plunging bevel gear grinding, it must be investigated in more detail to what extent the effective contact conditions change over the process time due to elasticity.

7 Summary and Outlook

Grinding of bevel gears is an established process in the production of gear drives. Due to the complex kinematics, the predictability of bevel gear grinding processes has been limited. To determine the cutting force, the exact contact conditions between bevel gear and grinding wheel must be known. In the present report, a numerical simulation model was introduced with which the contact conditions in plunging and generating bevel gear grinding can be determined. With the help of the simulation model, the characteristic geometric process values can be predicted. In this report, a geometric validation of the simulated tooth gap was carried out using nominal measurement data. It was found that the maximum deviations between the simulated tooth flanks and the nominal flanks are $\Delta_{max} < 0.2\,\mu m$ for both plunging and generating bevel gear grinding. The proven accuracy is assumed to be sufficient for further modeling purposes.

In the investigations presented, the simulated course of the material removal rate for plunging and generating bevel gear grinding was compared with the measurement of the tool spindle power. For plunging bevel gear grinding, a different course of material removal rate and spindle power was obtained. This fact was already established in earlier investigations and was attributed to changes in contact conditions caused by elasticity [19, 24, 28]. However, a proportional correlation between simulated material removal rate and measured tool spindle power was determined for generating bevel gear grinding.

In addition, it was found that in combination with an equation established by Rütti the influence of the process parameters on the cutting force in generating bevel gear grinding can be predicted with the simulation model [17]. Thus, the objective to develop a model for the cutting force could be achieved for generating bevel gear grinding. For plunging bevel gear grinding, it will be necessary to investigate in more detail how the contact conditions change during the process due to elasticity.

Acknowledgement. The authors gratefully acknowledge the financial support by the German Research Foundation (DFG) for the project [KL 500/187-1] as well as the support of the project partners Scania CV AB and Saint-Gobain Abrasives GmbH.

References

1. Karpuschewski, B., Knoche, H.J., Hipke, M.: Gear finishing by abrasive processes. CIRP Ann. **57**(2), 621–640 (2008). https://doi.org/10.1016/j.cirp.2008.09.002
2. Bausch, T.: Innovative Zahnradfertigung. Verfahren, Maschinen und Werkzeuge zur kostengünstigen Herstellung von Stirnrädern mit hoher Qualität, 3rd edn. Kontakt & Studium, vol. 175. Expert, Renningen (2006)
3. Stadtfeld, H.J.: A split happened on the way to reliable, higher-volume gear grinding. Gear Technol. (5, September/October 2005)
4. Weßels, N.: Flexibles Kegelradschleifen mit Korund in variantenreicher Fertigung. Dissertation, RWTH Aachen University (2009)
5. Tönshoff, H.K., Peters, J., Inasaki, I.: Modelling and simulation of grinding processes. CIRP Ann. Manuf. Technol. **41**(2), 677–688 (1992)
6. Rütjes, U.A.: Entwicklung eines Simulationssystems zur Analyse des Kegelradfräsens. Dissertation, RWTH Aachen University (2010)
7. Herzhoff, S.: Werkzeugverschleiß bei mehrflankiger Spanbildung. Dissertation, RWTH Aachen University (2013)
8. Hardjosuwito, A.F.: Vorhersage des lokalen Werkzeugstandweges und der Werkstückstandmenge beim Kegelradfräsen. Dissertation, RWTH Aachen University (2013)
9. Werner, G.: Kinematik und Mechanik des Schleifprozesses. Dissertation, RWTH Aachen University (1971)
10. Kienzle, O.: Die Bestimmung von Kräften und Leistungen an spanenden Werkzeugen und Werkzeugmaschinen. VDI-Z **94**(11/12), 299–305 (1952)
11. Jankowski, R.: Mathematische Modellierung des Umfangsschleifens auf Basis der Spanungsgrößen an der Einzelschneide. Dissertation, Hochschule Wismar (1991)
12. Bock, R.: Schleifkraftmodell für das Außenrund- und Innenrundschleifen. In: Saljé, E. (ed.) Jahrbuch Schleifen, Honen, Läppen und Polieren. Verfahren und Maschinen. Vulkan, Essen (1987)

13. Grinko, S.: Thermo-mechanisches Schädigungsmodell für das (Zahnflanken-) Profilschleifen. Dissertation, TU Magdeburg (2006)
14. Scheffler, F.: Zerspankräfte beim kontinuierlichen Wälzschleifen von Stirnradverzahnungen. Dissertation, RWTH Aachen University (2019)
15. Masslow, E.N.: Grundlagen der Theorie des Metallschleifens. Technik, Berlin (1952)
16. Brückner, K.: Der Schleifvorgang und seine Bewertung durch die auftretenden Schnittkräfte. Dissertation, RWTH Aachen University (1962)
17. Rütti, R.: Beitrag zur Prozessmodellierung am Beispiel des Zahnradhonens. Dissertation, TH Zürich (2000)
18. Kampka, M.: Lokal aufgelöste Zerspankraft beim Verzahnungshonen. Dissertation, RWTH Aachen University (2017)
19. Solf, M., Löpenhaus, C., Klocke, F.: Contact conditions in bevel gear grinding. In: Schmitt, R., Schuh, G. (eds.) 7. WGP-Jahreskongress. Apprimus, Aachen (2017)
20. Klingelnberg, J.: Kegelräder. Grundlagen, Anwendungen. Springer, Berlin (2008)
21. Sulzer, G.: Leistungssteigerung bei der Zylinderradherstellung durch genaue Erfassung der Zerspankinematik. Dissertation, RWTH Aachen University (1973)
22. Klocke, F., Brecher, C., Löpenhaus, C., Mazak, J.: Describing and evaluating deviations for bevel gear flanks. Procedia. CIRP **62**, 221–226 (2017)
23. Brinksmeier, E., Tönshoff, H.K., Czenkusch, C.: Modelling and optimization of grinding processes. J. Intell. Manuf. **9**, 303–314 (1998)
24. Solf, M., Löpenhaus, C., Bergs, T., Klocke, F.: Analysis of the grinding force in plunging bevel gear grinding. In: Schmitt, R., Schuh, G. (eds.) Advances in Production Research. Proceedings of the 8th Congress of the German Academic Association for Production Technology (WGP). Springer International, Cham (2018)
25. Klocke, F., König, W.: Fertigungsverfahren 2. Schleifen, Honen, Läppen, 5th edn. VDI-Buch. Springer, Berlin (2017)
26. Malkin, S., Guo, C.: Grinding technology. Theory and application of machining with abrasives, 2nd edn. Industrial Press, New York (2008)
27. Dietz, C.: Numerische Simulation des kontinuierlichen Wälzschleifprozesses unter Berücksichtigung des dynamischen Verhaltens des Systems Maschine – Werkzeug – Werkstück. Dissertation, ETH Zürich (2017)
28. Solf, M., Löpenhaus, C., Bergs, T.: Einfluss der elastischen Nachgiebigkeit auf die Kraft beim tauchenden Kegelradschleifen. Forsch. Ingenieurwes. (2019)

Design of Pulsed Electrochemical Machining Processes Based on Data Processing and Multiphysics Simulation

I. Schaarschmidt[1(✉)], S. Loebel[1], P. Steinert[1], M. Zinecker[1], and A. Schubert[1,2]

[1] Professorship Micromanufacturing Technology, Chemnitz University of Technology, 09107 Chemnitz, Germany
ingo.schaarschmidt@mb.tu-chemnitz.de
[2] Fraunhofer Institute for Machine Tools and Forming Technology, 09126 Chemnitz, Germany

Abstract. The process principle of electrochemical machining (ECM) is based on anodic material dissolution at the interface between the workpiece surface and an electrically conductive solution without any mechanical or significant thermal impact on the workpiece surface. As the material removal mechanism is contact- and force-free it is independent of mechanical properties of the workpiece material such as strength or hardness.

In this paper, a methodology to perform the process design of pulsed electrochemical machining (PECM) for manufacturing a three-dimensional geometry based on a sequence of standardized material characterization experiments, data processing and multiphysics simulation is shown. As a part of parameter studies, the machining parameters cathode feed rate and process voltage were varied to determine their influence on the resulting workpiece geometry. Resulting process conditions as for example the working gaps and electric current density distributions were analyzed and compared with experimental results.

Keywords: Pulsed electrochemical machining (PECM) · High-strength materials · Process design

1 Introduction

Pulsed electrochemical machining (PECM) is based on the same principle like conventional electrochemical machining (ECM) but is characterized by a pulsed electric current in combination with an oscillating tool. PECM is suitable for the manufacturing of high strength materials for instance powder metallurgical steels and it is therefore an established manufacturing technology in the fields of medical devices, aerospace components and tools [1].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 256–265, 2021.
https://doi.org/10.1007/978-3-662-62138-7_26

Figure 1 shows the principle of pulsed electrochemical machining with oscillating cathode, which can be described by four consecutive phases. The distance between the cathode and anode decreases from phase I to phase II and an electric current pulse is triggered at the bottom dead center of the oscillation and anodic material dissolution takes place. Afterwards, the cathode moves to the top dead center of the oscillation in phase III and the working gap is flushed to provide fresh electrolyte for the next electric current pulse in phase IV. Main influences on the electrolyte properties during pulsed electrochemical machining are the evolution of process gases such as hydrogen and oxygen, an increase of the electrolyte temperature and the occurrence of reaction products [2].

PECM has several advantages in comparison to conventional electrochemical machining. The implementation of a pulsed electric current, instead a continuous current, leads to higher accuracy and process stability [3]. A further improvement is the combination of a pulsed electric current with an oscillating cathode to increase the distance between the anode and the cathode during the pulse-off time and in consequence renew the electrolyte conditions within the working gap [4].

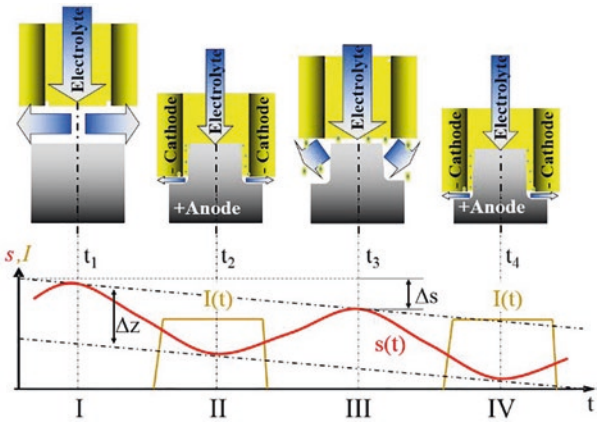


Fig. 1. Principle of pulsed electrochemical machining (PECM) with oscillating cathode (*s* - position of cathode, *I* - electric current); Flushing (I); Dissolution (II); Flushing (III); Dissolution (IV)

Due to the high complexity of the process principle and the variety of adjustable process parameters, the design of the PECM process for manufacturing of three-dimensional geometries requires a cost intensive and time-consuming, iterative procedure to achieve precise results. Therefore, multiphysics simulation is a suitable method to enable the process design for ECM and PECM by calculation of resulting workpiece geometries, analyzing electrolyte properties within the working gap and by optimization of the tool shape [5]. For this purpose, it is necessary to develop simulation models, which describe relevant physical phenomena as well as the formation of the resulting workpiece geometry within practical computational effort limits.

In this study a methodology to perform the process design for a rotationally symmetric external geometry based on a sequence of standardized material characterization experiments, data processing and multiphysics simulation is shown. The workpiece geometry is based on typical dimensions of punches for forming processes. The recorded data of multiple characterization experiments were merged and analyzed in order to calculate relevant electrochemical material properties. The results were integrated into a central PostgreSQL database set up and served as input for the PECM process simulation models. The multiphysics simulations were performed to model the machining process considering relevant physical phenomena and the experimentally obtained material-specific dissolution behavior. As a part of the parameter studies, the machining parameters cathode feed rate and process voltage were varied to verify the validity within a considered parameter range and identify limits of the simulation models. Resulting working gaps and electric current density distributions were analyzed and compared to experimental results.

2 Design of Pulsed Electrochemical Machining Processes

Figure 2 summarizes the simulation-based design of PECM processes schematically. First step is the analysis of the initial workpiece geometry, material, treatment condition as well as the requirements of the part to be machined, such as geometric tolerances or surface quality. Following a rough concept of the material removal process with its relevant input parameters, such as tool geometry, feed rate, working voltage as well as pulse frequencies and pulse durations, is outlined. Based on the order to implement this concept into a process simulation model, the electrochemical properties of the workpiece material have to be determined experimentally.

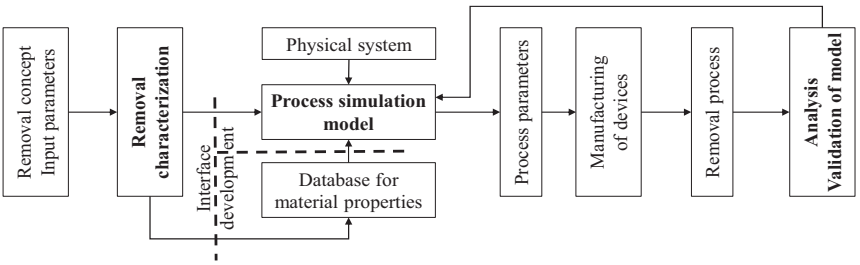


Fig. 2. Methodology for the simulation-based design of pulsed electrochemical machining processes

During this removal characterization after DIN SPEC 91399 [6], the current characteristics are recorded under variation of input parameters in order to map material removal velocities for a sufficient range of current densities, which may occur during the removal process. After that, all relevant electrochemical material properties are calculated from the recorded process data and integrated into both an online

database and process simulation models for subsequent use via developed interfaces, as detailed in Sect. 3.2. With all necessary input parameters at hand, suitable machining parameters are determined within the process simulation model. With the aim to validate the simulated results, these machining parameters are further transferred to the ECM machine. Current characteristics as well as working distances can now be compared between numerically simulated and real machining process. Ultimately, limitations of the simulation models are identified to evaluate the applicability of multiphysics simulation for the design of PECM of external geometries.

3 Standardized Determination of Electrochemical Material Characteristics

3.1 Experimental Setup

The material characterization of the 1.3202 alloy was conducted on the EC manufacturing machine PEMCenter 8000 following DIN SPEC 91399 [6, 7]. A water-based solution with a of $NaNO_3$ content of 8% was used as electrolyte. Both the cathodic tool electrode and the anodic sample workpiece were of cylindrical shape with a diameter of 12 mm, aligned coaxially. Therefore, a constant machining area during the experiments was achieved, allowing for direct calculation of the occurring normal current densities from the recorded current characteristics. Relevant process parameters are summarized in Table 1.

Table 1. Process parameters for determination of removal characteristics of 1.3202

Symbol	Parameter	Value
σ_{el}	Electric conductivity of electrolyte	69 mS/cm
p_{in}	Inlet pressure	3.5 bar
p_{out}	Outlet pressure	1.0 bar
v_f	Cathode feed rate	(0.007 … 0.3) mm/min
U_q	Process voltage	(5.5 … 14.2) V
c	Duty cycle	0.2

3.2 Data Evaluation

In order to describe the electrochemical properties of 1.3202 with sufficient precision, 52 removal experiments under variation of cathode feed rate and process voltage were conducted. These properties include removal velocities, current efficiencies and sums of overpotentials, reducing the effective voltage while machining electrochemically, as functions of the current density present on the workpiece surface. These functions cannot be measured directly. To calculate them from the accumulated process data of

all experiments, an automated data evaluation workflow as an interface between characterization experiments, material database and simulation environment was developed within the data science tool KNIME Analytics Platform [8]. As outlined in the workflow in Fig. 3, the results can be presented tabularly and graphically and is subsequently integrated into a PostgreSQL database structure. The separated section on the right-hand side of this workflow allows accessing all data collected in the past in a format suitable for integration in multiphysics simulation. As this data is available online, it can also be accessed via web interface.

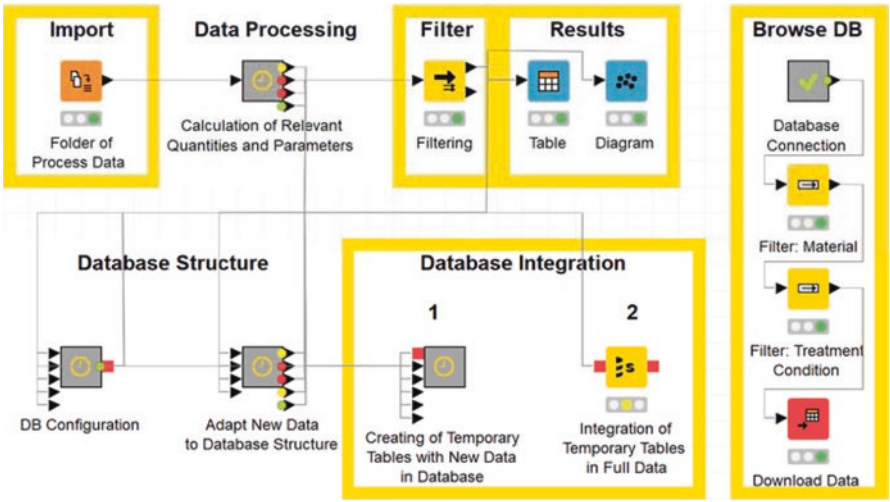


Fig. 3. Data evaluation workflow in KNIME Analytics Platform consisting of consecutive import, processing, visualization and database integration of process data, recorded during the material characterization. Previously uploaded data can be accessed, filtered and downloaded

3.3 Results

Figure 4 shows exemplarily the calculated removal velocity v_a and current efficiency η as functions of the current density J present on the surface of the 1.3202 workpiece. These material-specific functions are a required input for the process simulation model and can be described by continuous segments with no material removal occurring at current densities below 19 A/cm^2. The accuracy of this model will be further increased by determining the sum of overpotentials. These overpotentials fluctuate between 3 and 8 V averaging 5.1 V for the characterized 1.3202 alloy.

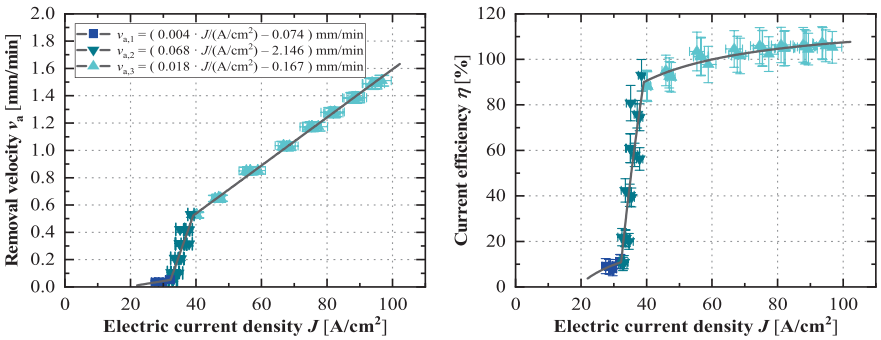


Fig. 4. Removal velocity v_a (left) and current efficiency η (right) as functions of normal current density J

4 Multiphysics Simulation

4.1 Geometry and Methods

The commercial finite element method software COMSOL Multiphysics was used to perform the simulations. Considered physical phenomena are electrodynamics, thermodynamics and fluid dynamics. The initial state of the process simulation model is a cylindrical-shaped sample with an initial outer radius $r_{A,0}$ of 10 mm, which can be seen in Fig. 5. After a certain simulated machining time an equilibrium state is reached. The specific time, when an equilibrium state is reached depends on the cathode feed rate. At this state, the final outer radius of the workpiece is formed and the workpiece geometry stays constant related to the cathode geometry. Figure 5 schematically shows the initial state of the model on the left side and the geometry at the equilibrium state on the right side. The front working area is between the angular cathode face and workpiece.

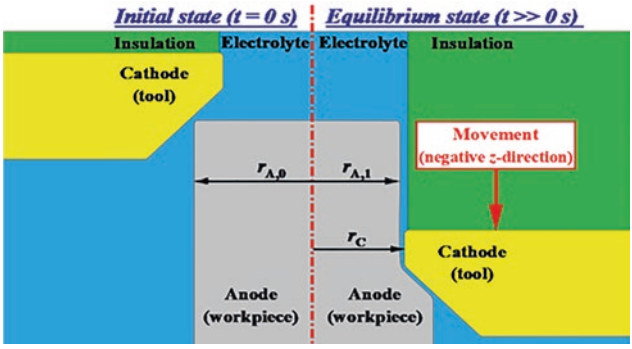


Fig. 5 Initial Geometry (left); Resulting geometry at equilibrium state (right)

The side working area is between the vertical cathode inner surface and workpiece and can be characterized by the resulting side working gap s_s. In this context, the side working gap is an important criterion because the precise adjustment of the final workpiece geometry can be achieved by varying the side working gap due to purposeful definition of the process parameter. In this study, the side working gap s_s is the lateral difference between inner cathode radius r_C and outer radius of the anodic workpiece after the machining process $r_{A,1}$ ($s_s = r_c - r_{A,1}$). Calculating the removal velocity at the anode surface is fundamental to simulate the workpiece formation due to anodic dissolution. Therefore, the current density dependent removal velocity of the workpiece material was implemented and a primary current distribution model was used to calculate the normal current density at the anode surface. In consequence, the charge transfer in the electrolyte follows Ohm's law and electrode kinetic and concentration-dependent effects are neglected. The current density is a function of the electric conductivity σ of the electrolyte and the electric field E. To take the influence of heat generation and hydrogen production into account the electric conductivity of the electrolyte is defined as a function of temperature T and fraction of hydrogen ϕ_{H_2}. A full description of the simulation methods and definition of domain and boundary conditions can be seen in [9–11].

4.2 Results

Selected results of the side working gap s_s and electric current I as functions of the cathode feed rate v_f and process voltage U_q are presented in the following chapter. The resulting geometry was measured after the machining process while the electric current was monitored and stored during the process by the monitoring system of the PEMCenter 8000. A tactile measurement device was used to determine the radii experimentally.

Influence of the process voltage on the side working gap and electric current. In Fig. 6 the resulting side working gap s_s as function of the applied process voltage U_q is shown on the left. It is apparent that the side working gap decreases linearly with decreasing applied voltage within the considered parameter range. Minimum side working gaps in experiment and simulation of about 90 µm and 100 µm, respectively, can be observed. The maximum difference of about 20 µm (13%) can be seen at a process voltage of 10 V. On the right, Fig. 6 shows the electric current I as function of the applied process voltage U_q after reaching the equilibrium state. It can be seen that the electric current increases analogue to the side working gap linearly with increasing process voltage. An offset of the simulation results compared to experimental results is noticeable. However, the implemented multiphysics simulation model predicts the influence of the cathode feed rate on the resulting side working gap very well. In consequence, the functionality of the developed methodology could be confirmed for the considered process and input parameters.

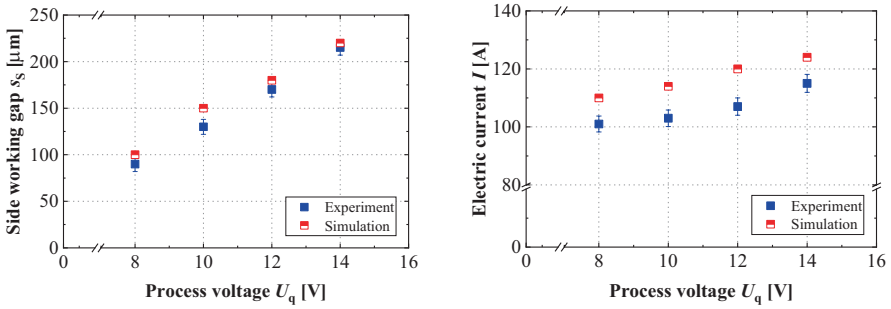


Fig. 6. Side working gap s_s (left) and electric current I (right) as functions of the process voltage U_q; $v_f = 0.3$ mm/min, $c = 50$ Hz · 4 ms = 0.2

Influence of the cathode feed rate on the side working gap and electric current. Figure 7 shows on the left the resulting side working gap s_s as function of cathode feed rate v_f in negative z-direction. It can be seen that experimental results show a decrease of the side working gap with increasing cathode feed rate. Furthermore, from Fig. 7 (left) it can be assumed that there will be no significant reduction of the side working gap with further increasing cathode feed rate as a result of the regressive approximation. Results from simulation confirms the relation. Very slow cathode velocities with very large workpiece formation cannot be predicted accurately by simulation. Differences of side working gap between experiment and simulation at low cathode velocities are a consequence of the summation of inaccuracies due to input parameters, simulation methods, numerical solution methods and meshing. All inaccuracies add up and increase with longer simulated dissolution time and workpiece formation.

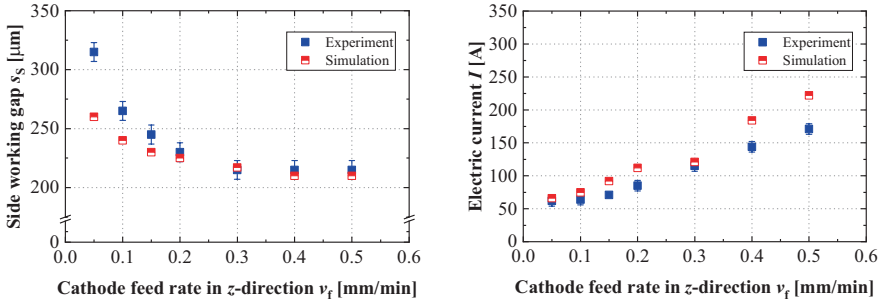


Fig. 7. Side working gap s_s (left) and electric current I (right) as functions of cathode feed rate in z-direction v_f; $U_q = 14$ V, $c = 50$ Hz · 4 ms = 0.2

Figure 7 shows on the right side the resulting electric current I as function of the cathode feed rate v_f in negative z-direction. The electric current increases with increasing cathode feed rate. Again, the simulation overestimates the electric current for all considered values of v_f. A limited validity of the developed methodology for values of

$v_f \geq 0.1$ mm/min can be derived from the results, since the difference between the simulation and experiment exceeds a value of 15% for a cathode feed rate of 0.5 mm/min.

5 Summary and Conclusion

In this study a methodology for the simulation-based process design for pulsed electrochemical machining (PECM) was shown. The entire methodology was presented exemplarily for the manufacturing of a rotationally symmetric external geometry. Simulation models of ECM processes required material specific dissolution characteristics, which are determined in standardized characterization experiments. Relevant electrochemical properties of steel 1.3202, such as removal velocities and current efficiencies, were calculated from arising process data via developed interfaces and subsequently integrated into both, an online database and a process simulation model. Numerical studies on the developed model for PECM of complex external geometries were accompanied by experimental verification on an industrial EC manufacturing machine. Relevant process parameters such as process voltage and cathode feed rate were varied and the resulting workpiece geometries were characterized.

This study confirmed the possibility to adjust the side working gap significantly by changing the process parameters. Analysis of the results showed that changing the process voltage is the preferred method to vary the side working gap due to the wide adjustment range of 125 µm. The process voltage was the process parameter whereby the reduction of U_q results in a decrease of the side working gap and the electric current. Overall the simulation models showed good agreement with the experiments but limitations for simulation models could be derived. Future work should include the analysis of the influence of different pulse frequencies and pulse durations on the side working gap.

References

1. Fransens, J.R., Regt, C., De, C., Zijlstra, H.: Pulsed precision ECM applications in the field of consumer products and medical applications. In: Proceedings of the 10th INSECT, Saarbrücken, pp. 15–23 (2014)
2. Bard, A., György, I., Scholz, F.: Electrochemical Dictionary. Springe, Berlin (2012). https://doi.org/10.1007/978-3-642-29551-5
3. Lauwers, B., Klocke, F., Klink, A., Tekkaya, E., Neugebauer, R., McIntosh, D.: Hybrid processes in manufacturing. In: CIRP Annals – Manufacturing Technology, CIRP, pp. 561–583 (2014). https://doi.org/10.1016/j.cirp.2014.05.003
4. Rajurkar, K.P., Zhu, D., McGeough, J.A., Kozak, J., De Silva, A.: New developments in electro-chemical machining. In: CIRP Annals – Manufacturing Technology, pp. 567–579 (1999). https://doi.org/10.1016/S0007-8506(07)63235-1
5. Hinduja, S., Kunieda, M.: Modelling of ECM and EDM processes. In: CIRP Annals – Manufacturing Technology, CIRP, pp. 775–797 (2013). https://doi.org/10.1016/j.cirp.2013.05.011

6. Deutsches Institut für Normung e. V.: DIN SPEC 91399: Methode zur Bestimmung von Prozesseingangsgrößen für das elektrochemische Präzisionsabtragen – Anforderungen, Kriterien, Festlegungen (2018)
7. Meichsner, G., Hackert-Oschätzchen, M., Krönert, M., Edelmann, J., Schubert, A., Putz, M.: Fast determination of the material removal characteristics in pulsed electrochemical machining. In: CIRP Conference on High Performance Cutting – HPC 2016, pp. 123–126. Elsevier B.V. (2016). https://doi.org/10.1016/j.procir.2016.03.175
8. Loebel, S., Hackert-Oschätzchen, M., Meichsner, G., Schubert, A.: Development of interfaces for material data integration in models of electrochemical machining processes. In: Proceedings of the 13th INSECT, p. 8 (2017)
9. Schaarschmidt, I., Zinecker, M., Hackert-Oschätzchen, M., Meichsner, G., Schubert, A.: Multiscale multiphysics simulation of a pulsed electrochemical machining process with oscillating cathode for microstructuring of impact extrusion punches. Procedia CIRP. **58**, 257–262 (2017). https://doi.org/10.1016/j.procir.2017.04.005
10. Schaarschmidt, I., Meichsner, G., Zinecker, M., Schubert, A.: Multiscale model of the PECM with oscillating cathode for external geometries using a virtual switch. In: COMSOL Conference Rotterdam 2017 (2017).
11. Schaarschmidt, I., Hackert-Oschätzchen, M., Meichsner, G., Zinecker, M., Schubert, A.: Implementation of the machine tool-specific current and voltage control characteristics in multiphysics simulation of electrochemical precision machining. Procedia CIRP. **82**, 237–242 (2019). https://doi.org/10.1016/j.procir.2019.04.142

Functional PVD Hard Coatings for High Temperature Cutting Processes

N. Stachowski(✉), N. C. Kruppe, T. Brögelmann, and K. Bobzin

Surface Engineering Institute, RWTH Aachen University, Kackertstraße 15, 52072 Aachen, Germany
stachowski@iot.rwth-aachen.de

Abstract. To improve the wear resistance of cutting tools, PVD (physical vapor deposition) coatings are well established in manufacturing. Therefore, different PVD hard coating systems have been developed until now. Nevertheless, the demands on the functionality of PVD hard coatings have become more specific. Reasons for this include increased cutting speeds, the machining of high strength materials with low heat conductivity as well as the development towards minimum quantity lubrication and dry machining. In summary, all these leads to increasing temperatures at the cutting edge. As a result, the wear resistance to thermally induced failure mechanisms and self-lubricating PVD hard coatings getting more important. In this work, important requirements for PVD coatings for high temperature cutting processes are presented. Furthermore, an overview of self- lubricating oxidic coating systems based on aluminum oxide (Al_2O_3) and Magnéli-phases as well as their wear behavior in cutting application is given.

Keywords: Tool coatings · PVD · Magnéli-phases · High temperature application

1 Introduction

The manufacturing industry is part of a constant development process and undergoes significant chances due to new demands, especially in the mobility and energy sector. Therefore, new sustainability goals are formulated and corresponding technical solutions are developed. For research, this means the development of new, specially adapted materials as well as efficient and cost effective processing methods for them. An industrial well established variant is metal cutting. To improve the wear resistance of cutting tools, the application of PVD or CVD (chemical vapor deposition) coatings are state of the art. Often CVD is preferred when a very thick coating with high wear volume is required, i.e. in turning. In contrast, thinner PVD coatings are more appropriate for coating sharp cutting edges. Additionally, lower coating temperatures made the coating of heat treated tool steels possible and PVD technology can also be used to synthesize metastable coating systems. An industrial established example is titanium-aluminum-nitride (Ti,Al)N [1]. In general, PVD coatings for cutting tools need to provide

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 266–274, 2021.
https://doi.org/10.1007/978-3-662-62138-7_27

wear protection against abrasion, adhesion and protection against oxidation [1]. PVD coatings have a significant effect on the tool life, since interactions between tool and workpiece always occur at the surfaces. From an economic point of view, more than the increase in service life, an increase of the cutting performance and cost reduction of manufacturing are of considerable interest [2]. As a result, the demands on the functions of PVD hard coatings are becoming more specific. Reasons for this include increasing cutting speeds, machining of materials with higher strengths and lower heat conductivity as well as the development towards minimum quantity lubrication or dry machining. In summary all these lead to increasing temperatures at the cutting edge [3]. In dry turning of unalloyed steels, temperatures of up to T = 1,000 °C can be reached [4]. Simulations for the machining of titanium alloys have also shown that temperatures up to T = 1,150 °C can occur [5, 6]. As a result, the wear resistance to thermally induced failure mechanisms and reducing cutting forces by self-lubricating PVD coatings are getting more important. This can be achieved e.g. by integrating solid lubricants into PVD coatings. This work briefly summarizes important requirements for PVD coatings for cutting tools exposed to peak temperatures up to T = 1,000 °C on the one hand. On the other hand, an overview of oxide coatings based on Al_2O_3 and Magnéli-Phases and their wear behavior in cutting applications is given.

2 Wear on Cutting Tools in High Temperature Applications

Cutting tools in high-performance cutting processes are exposed to unfavorable ambient conditions. These include varying stress conditions, cyclical mechanical and thermal loads and high cutting temperatures [7]. Thermal loads additionally increase if the workpiece material has a low thermal conductivity such as titanium alloys or austenitic, high-alloyed steels. With these material groups, it is not possible that heat occurring during machining dissipates quickly into the chip or the workpiece. Therefore, the most heat is transferred to the cutting tool. Additionally, any contact between the cutting tool and the workpiece surface causes frictional processes that trigger tool wear [4]. According to the current state of research, various individual mechanisms need to be considered for the collective term wear. These mechanisms overlap in tribological contact conditions. Hence, it is difficult to distinguished clearly between their cause and effect on the tool failure. Figure 1 shows schematically the dependence of tool damage on cutting temperature.

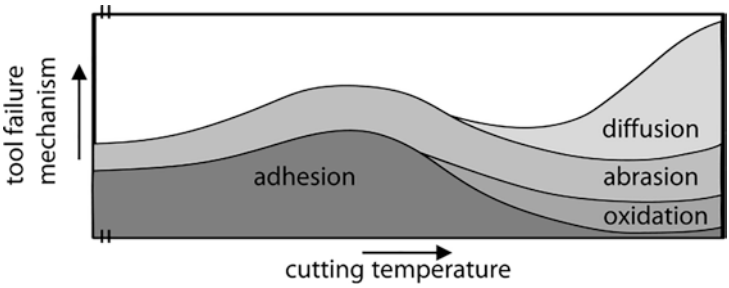


Fig. 1. Temperature dependence of dominant failure mechanisms at cutting tools [4]

It can be seen, that the influence of thermally induced damage mechanisms such as diffusion and oxidation increases with rising cutting temperatures. Additionally, due to high thermal loads at the cutting edge, chipping, transverse and comb cracks or partially plastic deformations occur more frequently. In [8] it is mentioned that while mechanical damage, including abrasion, early fractures and fatigue, occurs nearly independently of the cutting temperature in every machining process, thermally induced damage such as diffusion and chemical reactions increases drastically. They are not the direct cause of tool damage, but they lead to a higher susceptibility for damage from abrasion or adhesion [8]. Therefore, an effective measure to prevent early tool failure is the deposition of a protective PVD hard coating on cutting tools. A self-lubricating effect additionally contributes to minimizing the cutting forces, the acting mechanical loads and the resulting temperatures.

3 Solid Lubricants

In general, lubricants are used to separate contact surfaces in tribological systems. This leads to a reduced coefficient of friction (CoF), reduced heat and noise emission and an increased protection against wear [9]. In cutting applications the used cooling lubricant (KSS) is additionally utilized to dissipate the generated heat from the cutting zone. Nevertheless, often the cooling lubricant cannot reach the tribological contact point between the cutting edge and formed chip. In such contact situations, PVD coating systems can contribute to a friction reduction at the contact point. Therefore, PVD coating systems with a high chemical stability and high chemical inertness at elevated temperatures or PVD coatings with integrated oxidic solid lubricants can be used. The general classification of solid lubricants is not standardized. Nevertheless, they can be categorized for technical applications into five groups: Metallic soaps, soft metals, compounds with layered lattice structure, polymers and oxidic and fluoridic compounds [10]. The most solid lubricants are not applicable at temperatures above T = 400 °C or their mechanical properties are not sufficient. In contrast to that, oxidic solid lubricants offer a high suitability for usage in high temperature applications. Suitable tribologically active coatings are either Al_2O_3, which can be used in the as deposited state, or PVD coating systems based on the in situ formation of self-lubricating oxide phases under tribological stress in machining based on refractory metals such as vanadium (V).

3.1 Al_2O_3 – As Deposited Friction Reduction

Oxidic coatings based on Al_2O_3 are particularly suitable as a coating that offers high wear protection and friction reduction already in the deposited state. In general, Al_2O_3 has a high hot hardness a good chemical stability and contributes to friction reduction in tribological contacts [11]. Al_2O_3 exists in different phase structures. The α-phase (corundum) is the only one that is thermodynamically stable at all temperature and pressure ranges. In the temperature range T = 400 °C and T = 1,000 °C several metastable polymorphic phases are available [12]. Depending on the arrangement of the oxygen ions, they can be divided into the hexagonal closest packed (hcp) and face centered cubic (fcc) lattice structure. The hcp structures include the trigonal α-phase,

the orthorhombic κ-phase and the hexagonal χ-phase. Among the fcc structures the cubic γ-phase, the cubic η-phase, the monoclinic θ-phase and either the tetragonal or the orthorhombic δ-phase can be assigned [12]. In contrast to the phase stability of common γ-Al_2O_3 ceramics, annealing tests in oxygen containing atmosphere have shown that γ-Al_2O_3 PVD coatings can provide a phase stability up to a significantly increased temperature range between T = 900 °C and T = 1,000 °C. However, at higher temperatures, an irreversible phase transformation from the metastable γ-Al_2O_3 to the stable α-Al_2O_3 can occur [13]. Such a phase transformation is rather critical for coating applications since a volume change leads to increased internal stresses and coating delamination [14]. Nevertheless, investigations have shown, that γ-Al_2O_3 offers a great potential as a coating system for cutting applications [11, 13–15], due to its high hot hardness, its high chemical stability and the contribution to a low tendency of adhesion towards workpiece materials. In general, the eligibility of Al_2O_3 as a wear protective PVD coating for cutting tools is already known. Nevertheless, the high power pulse magnetron sputtering (HPPMS) technology offers the possibility to improve the coating morphology and properties of the deposited γ-Al_2O_3 coatings on an industrial scale. In [14] the high temperature and machining behavior of a HPPMS-Al_2O_3 coating system was compared with that of an medium frequency magnetron sputtering (mfMS)- Al_2O_3 coating system. Therefore, in order to achieve the highest possible coating adhesion, an adapted (Ti,Al)N interlayer was deposited by direct current magnetron sputtering (dcMS) or mfMS first. Then the γ-Al_2O_3 top layer was synthesized by HPPMS or mfMs. Scanning electron microscopy (SEM) cross section fractures of both coating systems are shown in Fig. 2.

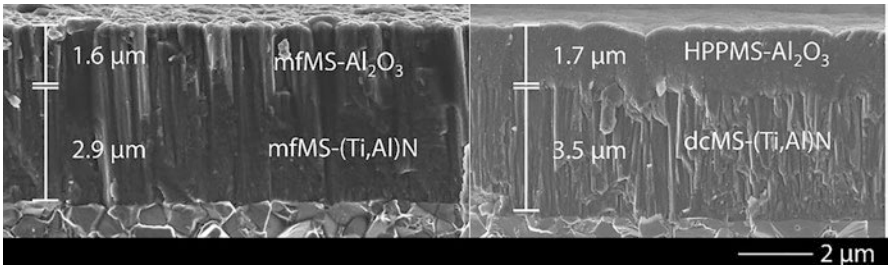


Fig. 2. SEM cross section fracture of the γ-Al_2O_3 multilayer coating system [14]

To analyze the high temperature behavior in an oxygen containing atmosphere, the samples were annealed at T = 800 °C, T = 900 °C and T = 1,000 °C for t = 1 h. The subsequent investigations have shown, that the HPPMS-Al_2O_3 provides a significant higher oxidation resistance compared to mfMs-Al_2O_3. The phase structure of the HPPMS-Al_2O_3 coating was stable up to T = 900 °C. At T = 1,000 °C the first areas of phase transformation to a α-Al_2O_3 were visible between the interlayer and the top layer. In case of the mfMS-Al_2O_3 coating such a phase transformation has already taken place at T = 900 °C. Additionally, the cutting behavior was investigated by milling the high strength cast alloy GGG70. Figure 3 visualized the width of the flank wear marks (VB) as a function of the cutting length (l_f) of both Al_2O_3 coating system compared to a commercial titanium-based reference PVD coating. Furthermore, the flank wear resulting at the end of the milling tests is shown. In both test series the

HPPMS-Al_2O_3 coating system has covered the longer cutting length until the wear criterion of VB < 0.1 mm was reached compared to the other coatings. On average, an increase in cutting length of approximately $l_f = +80\%$ compared to the mfMs-Al_2O_3 and of $l_f = +13\%$ compared to the reference coating was achieved. This is attributed to the higher resistance against abrasive wear, and a better thermal and chemical barrier effect of the γ-Al_2O_3 coating system.

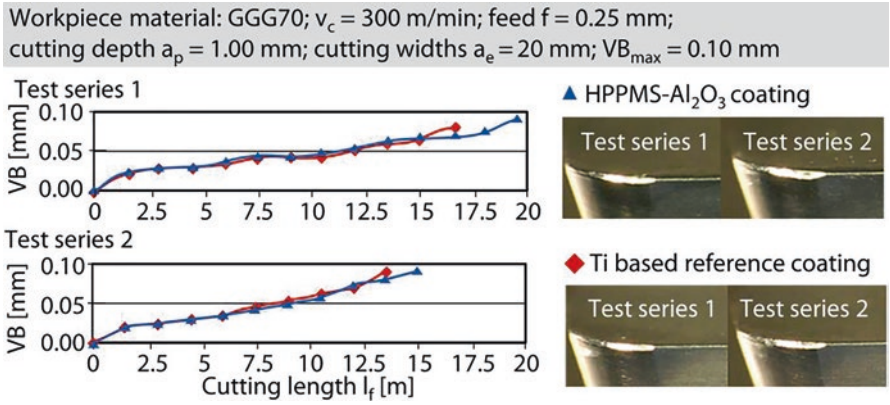


Fig. 3. Width of the wear marks as a function of the cutting length and the resulting wear [14]

3.2 Transition Metal Oxide Hard Coatings – In Situ Formation of Self-Lubricating Effect

Sub-stoichiometric oxide compounds are suitable to generate an in situ formation of self-lubricating oxide phases in PVD coatings. A number of these oxide compounds are classified as Magnéli-phases after their discoverer Magnéli [16]. Magnéli observed that a recurring pattern of oxide formation is present in different material systems. He described this as a homologous series that resembles a rhenium oxide structure (ReO_3-structure) with periodic atomic oxygen defects. The oxygen defects are arranged in weak bonded planes within the lattice structure, wherefore the planes act as slip planes under shear stress. A schematic representation of this is shown in Fig. 4a) using a ReO_3-structure as an example [16, 17].

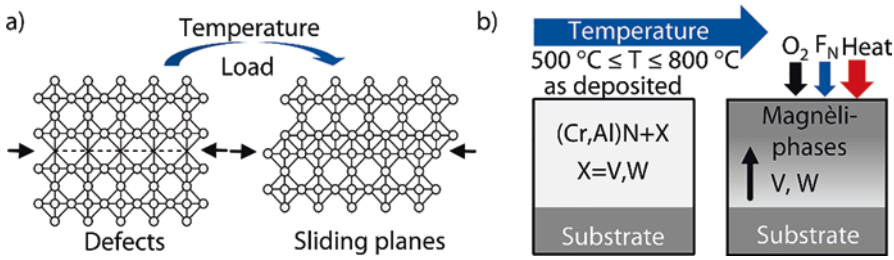


Fig. 4. a) Representation of the ReO_3-structer without (left) and with (right) shear planes, b) Temperature activation of the formation of Magnéli-phases [17]

Theoretically suitable for the formation of Magnéli-phases are the transition metals vanadium, titanium, molybdenum or tungsten [18]. Which oxide phase is formed in a cutting process depends on the number of missing oxygen atoms until a stoichiometric state is reached [17]. Oxides with a high ion potential contribute to a higher friction reduction at elevated temperatures than oxides with a lower ion potential, because the increased ionic potential shields cations more effectively. Therefore, they are not able to form strong covalent bonds and the formed oxides are softer and easier to shear [19]. In addition, these oxide compounds often have a lower melting point. Due to their ion potential in [19] the oxides Re_2O_7, B_2O_3, V_2O_5, MoO_3 and WO_3 are stated as low friction oxides. In particular, the oxides MoO_3, V_2O_5 and WO_3 are highlighted as promising candidates for friction reduction by the formation of Magnéli-phases [17, 20]. In order to integrate self-lubricating Magnéli-phases into PVD coatings, the embedding of Ti, Mo or V in already industrially established PVD hard coatings is a frequently used approach. Triggered by high cutting temperatures self-lubricating Magnéli-phases form due to oxidation processes [3, 17]. Successfully developed PVD coating systems that follow this approach are (Ti,Al)N / (VN) multilayer coatings and coating systems in which V atoms are directly integrated into the crystal lattice of (Ti,Al)N or (Cr,Al)N which is exemplary shown in Fig. 4b). In [21] the tribological and cutting behavior of a (Ti,Al)N/VN multilayer coating was analyzed. At temperatures above T = 700 °C a decrease in CoF due to the formation of Magnéli-phases was determined. Moreover, in machining aluminum alloys for aerospace and automotive components the achieved service life was twice as long as of the uncoated reference tool [21, 22]. In [23] the cutting performance of (Ti,Al,V)N coated cutting tools were analyzed in dry milling 42CrMo4. The achieved service life was increased by approximately $\Delta t = +50\%$ compared to the (Ti,Al)N coated reference tool [3]. In [24], V containing (Cr,Al)N PVD coatings were synthesized by a combination of Magnetron Sputtering (MS) and HPPMS. By means of SEM significant changes in the coating morphology were detected compared to the reference state "as deposited" after annealing for t = 4 h at T = 800 °C. Also a change in chemical composition have shown that at T = 800 °C the vanadium content and the oxygen content at the surface increase. This correlates with the change in coating morphology shown in Fig. 5.

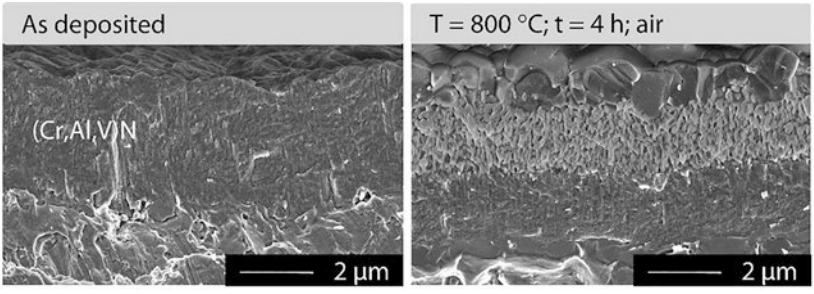


Fig. 5. Coating morphology by SEM before and after annealing [24]

In comparison to the reference state it is noticeable that a three-layer morphology has been developed after annealing which is attributed to diffusion processes of vanadium to the surface. The tribological behavior determined by PoD analysis at temperatures up to T = 800 °C correlates with these results. The comparison of the received CoF's in Fig. 6 indicates a significant decrease of the CoF at elevated temperatures.

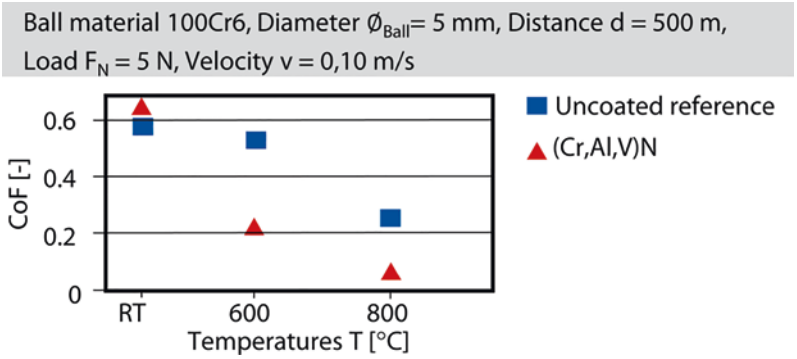


Fig. 6. CoF analyses of (Cr,Al,V)N and uncoated reference sample against 100Cr6 [24]

This behavior is attributed to the formation of Magnéli-phases [24]. While turning tests of TiAl6V4 and X153CrMoV12 at Iscar LTD, [17] revealed that a self-lubricating (Cr,Al,V)N coating has great potential to significantly increase tool life compared to a (Ti,Al)N coated reference tool. The cutting parameters and the achieved tool life are shown in Fig. 7a). The service life of the (Cr,Al,V)N coated tool until reaching the wear criterion $VB_{max} = 0.3$ mm was increased by factor 2.2 compared to the reference. An increase of tool service life was also achieved when turning X153CrMoV12. In terms of the occurring flank and crater wear that occur, a reduction was also achieved in both variants. In Fig. 7b) macroscopic images of occurring flank wear after machining TiAl6V4 for $t_s = 16$ min are shown. In case of both coating systems no break-outs at the cutting tools are visible. Nevertheless, the (Cr,Al,V)N coated cutting tool shows smaller flank wear after the same machining time.

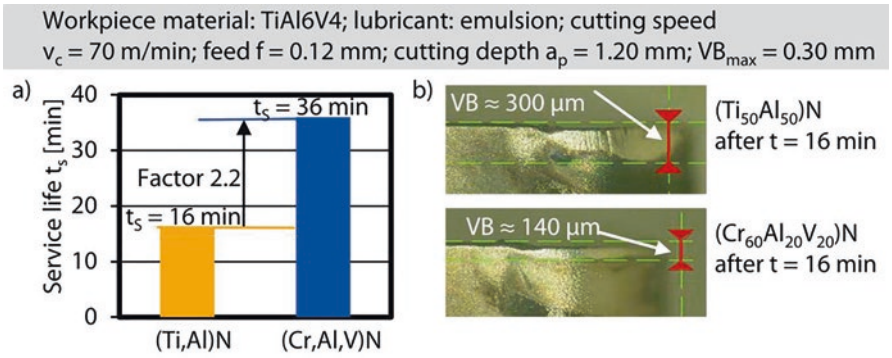


Fig. 7. a) Service life of (Cr,Al,V)N coated cutting tools after cutting TiAl6V4 and b) occurring flank wear after t = 16 min [17]

4 Conclusion

In this work important demands on PVD hard coatings for cutting tools exposed to peak temperatures up to T = 1,000 °C were introduced and an overview of coating systems based on Al_2O_3 and Magnéli-phases is given. In summary, PVD coatings in high temperature cutting applications need a sufficiently high resistance against thermally induced failure mechanisms. In addition, PVD coatings can contribute to the reduction of friction, thus reducing the tribological load on the tools and lowering the heat generation in the tribological contact. Due to the high temperatures in tribological contacts, PVD coating systems with a high chemical stability and high chemical inertness at elevated temperatures or PVD coatings with integrated oxidic solid lubricants are particularly suitable for high temperature cutting processes. At the current state of research there are PVD coatings available that can be used as deposited, such as Al_2O_3 and coating systems were the friction reduction is based on the in situ formation of binary oxide phases of transition metals. In machining tests, an increase in service life was achieved by almost all considered coating systems. Promising fields of application for such PVD coatings are difficult to machine materials such as titanium alloys and challenging cutting operations such as high speed cutting (HSC) or high performance cutting (HPC) with sharp cutting parameters.

Acknowledgements. The author gratefully acknowledge the financial support of the German Research Foundation, Deutsche Forschungsgemeinschaft (DFG) within the research project BO 1979/69-1.

References

1. Bobzin, K.: High-performance coatings for cutting tools. CIRP J. Manuf. Sci. Technol. **18**, 1–9 (2017)
2. Kalss, W., Reiter, A., Derflinger, V., Gey, C., Endrino, J.L.: Modern coatings in high performance cutting applications. Int. J. Refract. Met. Hard Mater. **24**, 399–404 (2006)
3. Franz, R., Mitterer, C.: Vanadium containing self-adaptive low-friction hard coatings for high-temperature applications: a review. Surf. Coat. Technol. **228**, 1–13 (2013)
4. Klocke, F.: Fertigungsverfahren 1 – Zerspanung mit geometrisch bestimmter Schneide, 9th edn. Springer Vieweg, Berlin (2018)
5. Ye, G.G., Xue, S.F., Tong, X.H., Dai, L.H.: Influence of cutting conditions on the cutting performance of TiAl6V4AMR. Adv. Mater. Res. **337**, 387–391 (2011)
6. Müller, B.: Thermische Analyse des Zerspanens metallischer Werkstoffe bei hohen Schnittgeschwindigkeiten. Dissertation, Aachen (2004)
7. Wang, F., Zhao, J., Li, Z., Li, A.: Coated carbide tool failure analysis in high-speed intermittent cutting process based on finite element method. Int. J. Adv. Manuf. Technol. **83**, 805–813 (2016)
8. Childs, T.: Metal Machining: Theory and Applications. Arnold, London (2000)
9. Busch, C.: Solid lubrication. In: Mang, T., Dresel, W. (eds.) Lubricants and Lubrication, pp. 694–714. Wiley, Weinheim (2007)
10. Bobzin, K.: Oberflächentechnik für den Maschinenbau, 1st edn. Wiley, Weinheim (2013)

11. Åstrand, M., Selinder, T.I., Fietzke, F., Klostermann, H.: PVD-Al_2O_3-coated cemented carbide cutting tools. Surf. Coat. Technol. **188–189**, 186–192 (2004)
12. Wefers, K., Misra, C.: Oxides and Hydroxides of Aluminum. Alcoa Laboratories, New Kensington (1987)
13. Bobzin, K., Bagcivan, N., Immich, P., Ewering, M.: Thermal investigation of Al_2O_3 thin films for application in cutting operations. Adv. Eng. Mater. **11**, 590–594 (2009)
14. Bobzin, K., Brögelmann, T., Kruppe, N.C., Basturk, S., Klocke, F., Gerschwiler, K., Kölker, W., Bolz, S., Kohlscheen, J.: Development of HPPMS-Al_2O_3-coatings for the machining of hard-to-machine materials. Materialwiss Werkstofftech. **49**, 1287–1300 (2018)
15. Ewering, M.T.: Synthese, thermische Stabilität und tribologisches Verhalten nanokristalliner γ-Al_2O_3-Schichten für die Zerspanung. Dissertation, Shaker, Aachen (2012)
16. Magnéli, A.: Structures of the ReO_3-type with recurrent dislocations of atoms: 'homologous series' of molybdenum and tungsten oxides. Acta. Cryst. **6**, 495–500 (1953)
17. Brugnara, R.H.: Hochtemperaturaktive HPPMS-Verschleißschutzschichten durch Bildung reibmindernder Magnéli-Phasen im System (Cr,Al,X)N. Dissertation, Shaker (2015)
18. Glasser, L.: Systematic thermodynamics of magneli-phase and other transition metal oxides. Inorgan. Chem. **48**, 10289–10294 (2009)
19. Erdemir, A.: A crystal-chemical approach to lubrication by solid oxides. Tribol. Lett. **8**, 97–102 (2000)
20. Gulbiński, W., Suszko, T., Sienicki, W., Warcholiński, B.: Tribological properties of silver- and copper-doped transition metal oxide coatings. Wear **254**, 129–135 (2003)
21. Luo, Q., Robinson, G., Pittman, M., Howarth, M., Sim, W.M., Stalley, M.R., Leitner, H., Ebner, R., Caliskanoglu, D., Hovsepian, P.E.: Performance of nano-structured multilayer PVD coating TiAlN/VN in dry high speed milling of aerospace aluminium 7010–T7651. Surf. Coat. Technol. **200**, 123–127 (2005)
22. Hovsepian, P.E., Luo, Q., Robinson, G., Pittman, M., Howarth, M., Doerwald, D., Tietema, R., Sim, W.M., Deeming, A., Zeus, T.: TiAlN/VN superlattice structured PVD coatings: a new alternative in machining of aluminium alloys for aerospace and automotive components. Surf. Coat. Technol. **201**, 265–272 (2006)
23. Tillmann, W., Momeni, S., Hoffmann, F.: A study of mechanical and tribological properties of self-lubricating TiAlVN coatings at elevated temperatures. Tribol. Int. **66**, 324–329 (2013)
24. Bobzin, K., Bagcivan, N., Ewering, M., Brugnara, R.H., Theiß, S.: DC-MSIP/HPPMS (Cr,Al,V)N and (Cr,Al,W)N thin films for high-temperature friction reduction. Surf. Coat. Technol. **205**, 2887–2892 (2011)

Wear Behaviour of PCBN, PCD, Binderless PCBN and Cemented Carbide Cutting Inserts When Machining Ti-6Al-4V in an Oxygen-Free Atmosphere

F. Schaper(✉), B. Denkena, M.-A. Dittrich, A. Krödel, J. Matthies, and S. Worpenberg

Institute of Production Engineering and Machine Tools, An der Universität 2, 30823 Garbsen, Germany
schaper@ifw.uni-hannover.de

Abstract. One approach to overcome the difficulties in machining titanium can be the selection of a suitable cutting material. However, influences of the chemical tool wear under machining atmosphere has not been considered jet. In this work the tribochemical wear resistance of uncoated carbide tools, polycrystalline cubic boron nitride (PCBN) tools polycrystalline binderless CBN tools, and polycrystalline diamond (PCD) tools is investigated under different atmospheric conditions cutting Ti-6Al-4V. Air, pure argon and silane-doped argon are used to determine the influence of oxygen contents on tool wear. It was found that oxidation and adhesion behaviour were influenced by oxygen content. Up to 33% increase in tool live were obtained with uncoated carbide- and the PCBN tools with high CBN content. The reduced oxidation also affects PCBN tool behaviour. During machining under oxygen-free conditions increased adhesion of Ti-6Al-4V occurs. The increased wear behaviour for binderless PCBN tools lead to significantly reduced tool life.

Keywords: PCD · Oxygen free · Binderless CBN · Cemented carbide · Titanium

1 Introduction

Nowadays, the demand for components made of titanium alloys (following named titanium), ranging from medical to aerospace applications, rises continuously. During the machining of hard-to-machine materials such as titanium, high temperatures occur at the interface between tool and workpiece due to the low thermal conductivity of titanium [1]. High contact and tooling temperatures then give occasion to associated problems such as diffusion wear and oxidation of the tool [1, 2]. Sun et al. [3] measured rake face temperatures in the range up to 900 °C for titanium machining.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 275–283, 2021.
https://doi.org/10.1007/978-3-662-62138-7_28

For the cutting of titanium typically uncoated cemented carbide tools are used [4]. Due to the short tool life of cemented carbide inserts, the usage of highly hard cutting materials such as polycrystalline cubic boron nitride (PCBN) and polycrystalline diamond (PCD) is another possibility for processing titanium [5]. Usually PCBN tools consist of a hard CBN phase surrounded by a ceramic or metallic binder phase. PCBN specifications are classified according to different criteria. Important parameters are the CBN content, the average grain size of the CBN and the composition of the binder phase [6]. Besides hardness, especially at high temperatures, these factors also influence the wear resistance during machining with PCBN tools [7, 8]. Uhlmann et al. detected an increase in tool life with increasing CBN content and decreasing grain size when machining titanium [8]. However, the cutting material with the highest known hardness is diamond, which is used in the form of PCD tools. The hot hardness of PCD is limited by the graphitisation, which starts at about 900 °C. CBN tools, on the other hand, show higher stability, especially against oxidation [9]. Thus, when the cutting speed is increased, which induces a temperature rise above the critical temperature in titanium machining, the wear resistance of PCD tools decreases significantly. This can result in shorter tool life for diamond tools compared to PCBN inserts [8]. Another approach for the machining of titanium is the use of binderless polycrystallince cubic boron nitride, which can be used to increase tool life compared to conventional PCBN tools due to its higher hardness and strength at elevated temperatures [10].

Mercer et al. have found out the interaction between two tribo-partners in general is influenced by the surrounding atmosphere by investigating the abrasive wear of titanium [11]. During machining, not only the titanium alloys are influenced by chemical reactions, but also the investigated tools can react with oxygen to form oxide layers [12–14].

A novel approach to reduce the oxygen content in production processes is the use of silane-doped inert gas to create atmospheres corresponding to an extremely high vacuum (XHV) regarding the oxygen content. By removing the oxygen from the atmosphere, it is assumed that tribochemical processes between tool and workpiece can be influenced to large extent compared to machining in air. In this work, two PCBN tools with different CBN content and binder materials, a binderless PCBN tool and a PCD tool are used in turning applications of Ti-6Al-4V under different atmospheric conditions. The influence of the atmospheric conditions on tool wear of these four tools were investigated and compared to uncoated cemented carbide inserts.

2 Experimental Setup

The machining has been carried out using a CTV400 lathe of the company Gildemeister under dry cutting conditions. A housing is mounted around the machining point, which is shown in Fig. 1. The space between housing and headstock is additionally closed by a plastic housing. Within this closed chamber the supply of inert gas is realized, which forms the ambient atmosphere during machining. Inert gas atmospheres of commercial quality are used for this purpose. However, conventional

inert gases of high quality (5.0) have an oxygen concentration of more than 5 ppmv despite their high purity. The addition of monosilane (SiH_4) to the argon gas completely removes the residual oxygen and water content (usually <20 ppmv) present in the argon inert gas atmosphere by transformation into SiO_2 and H_2. With regard to the presence of oxygen molecules, atmospheres corresponding to an extremely high vacuum (XHV) can be created.

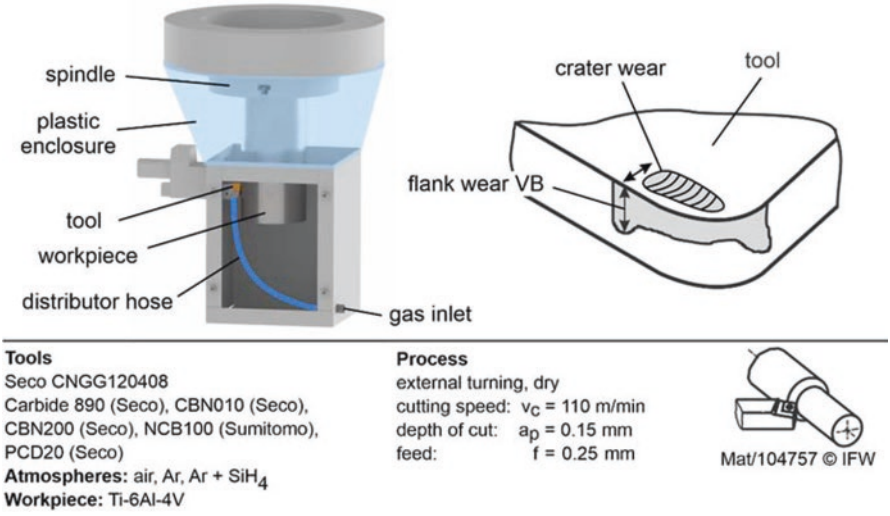


Fig. 1. Experimental setup

For the experiments a cutting speed of $v_c = 110$ m/min, a feed of $f = 0.25$ mm and a depth of cut of $a_p = 0.15$ mm were used. The tool life is assessed for four different tool materials in the cutting of Grade 5 titanium alloy Ti-6Al-4V and compared with the wear behaviour of uncoated cemented carbide inserts. The composition of the workpiece material is given in Table 1. The wear of the tools was measured at intervals of 2.5 min operating time with two tests being carried out for each tool in each atmosphere. The optical wear measurement is carried out by a digital microscope type VHX-600 from Keyence. The investigations are performed in three different atmospheres: air, technically pure argon and silane-doped argon (1.5%) each with three different CBN tools, a PCD tool and an uncoated carbide tool. Cemented carbide inserts of ISO-P designation CNGG 120408 were used for the machining experiments. The carbide tools have a chip-breaker and were uncoated. The carbide insert was supplied by SECO Tools, an EDX analysis on the device EDAX/Ametek octane Elite Super shows a composition of 94 wt.-% tungsten carbide and 6 wt.% of cobalt as binder for the carbide tool. The inserts were classified as fine grain cemented carbide. The PCBN tool with low CBN content, CBN010 from SECO Tools, consists of 50% CBN in a TiC binder phase and has an average grain size of 2 µm. The tool with high CBN content CBN200 is also from SECO Tools and has a CBN content of 85% within a ceramic Co-W-Al binder phase and an average grain size of 2 µm. The binderless PCBN tool NCB100 is from SUMITOMO and consists of 100% CBN with a CBN

grain size of 200 to 500 nm. The specifications of the PCD tool, PCD20 from SECO Tools, show that it has an average grain size of 10 µm and is within a co-binder phase.

In order to assess the influence of the cutting edge micro geometry on the wear behaviour of the tools, the cutting edges were measured with an Alicona Infinite Focus G5 prior to the investigations. The CBN tools and the uncoated carbide tools are used with a tool holder with ISO designation DCLN. This allows a rake angle of $\gamma = -6°$ to be set on the tool body side by the inclination of the tool holder. The uncoated carbide insert has a rake angle of $\gamma = 15°$ due to the chip-breaker geometry. All CBN inserts have a chamfer. In contrast to the previous tools, the PCD tool is used with a tool holder with ISO designation DCLC. The rake angle from the side of the tool holder is therefore 0°. The macro and micro geometries of the tool are shown in Table 2.

Table 1. Composition wt% of Grade 5 Ti-6Al-4V

Al	C	Fe	N	O	V	H	Ti
5.5–6.75	0.1	0.3	0.05	0.2	3.5–4.5	0.015	Balance

Table 2. Tool geometries

	Inclination angle tool-holder [°]	Rake angle of the insert [°]	Clearance angle of the insert [°]	Angle of chamfer [°]	Chamfer length [µm]	Cutting edge rounding [µm]
HM890	−6	15	0	-	-	12
CBN010	−6	0	0	−20	90	20
CBN200	−6	0	0	−20	200	30
NCB100	−6	0	0	−15	145	5
PCD20	0	0	7	-	-	6

3 Results

The results of the wear measurements for the investigated atmospheres and respective cutting materials are presented in Fig. 2. For each atmosphere, the tool life of the different tools is shown in form of a bar chart, which shows the respective mean value of the tool life and the minimum and maximum as a scatter bar. The width of flank wear land VB is measured after an interval of the determined cutting time and thereby, tool life is reached by passing the value VB = 100 µm of the width of flank wear land. Furthermore, the maximum cutting time is determined at 25 min.

Figure 2 shows that reducing the oxygen content of the cutting atmosphere for the carbide tool and the CBN tool with high CBN content (CBN200) has a positive effect on flank wear. An average tool life increase of 2.5 min can be achieved on the cemented carbide tool. For the CBN010 tool, the CBN200 tool and the PCD20 tool, no significant influence of the width of flank wear land in relation to the ambient atmosphere is found. When machining with the PCD tool within all three atmospheres

investigated, the maximum width of flank wear land of 100 µm is not reached. At the time of the maximum cutting time after 25 min all PCD tools have a width of flank wear land $VB = 32 \pm 2$ µm.

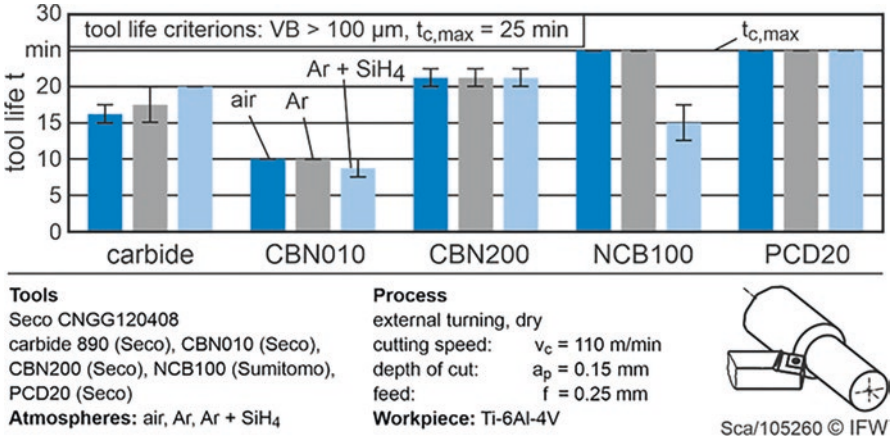


Fig. 2. Tool life of the investigated tools in different machining atmospheres

The most significant environmental influence can be seen in the investigation of the binderless PCBN tools NCB100. Under air and argon this tool reaches the maximum cutting time of 25 min. In contrast, a cutting edge chipping occurs on the binderless PCBN tool during titanium machining within the silane-doped argon atmosphere after 15 min, which leads to a reduction of the tool life by 10 min. Figure 3 shows the initial wear at the flank faces of the investigated inserts after 2.5 min in air, argon and silane-doped argon. It can be seen that scaling takes place on these tools under the influence of oxygen, which can be reduced by exposure to argon and almost suppressed with silane-doped argon. Moreover, during machining in an atmosphere containing silane, the formation of silicon dioxide can be observed, which settles as amorphous dust on the tool, the workpiece, the chips and the enveloping housing.

The scaling formation is additionally superimposed with the effect of adhesion.

Figure 4 shows the rake faces of the cemented carbide tool, the CBN010 tool, the CBN200 tool, the NCB100 tool and the PCD20 tool after a cutting time of 2.5 min and compares them under air and silane-doped argon atmosphere with regards to the observed workpiece adhesion. This comparison demonstrates that the PCBN tools and the PCD tool show an increased adhesion of the workpiece material to the tool in the silane-doped atmosphere, compared to air. Moreover, increased adhesion also occurs on the carbide tool when cutting in a silane-doped argon atmosphere. This phenomenon is additionally accompanied by an increased crater wear and flanking under oxygen free conditions.

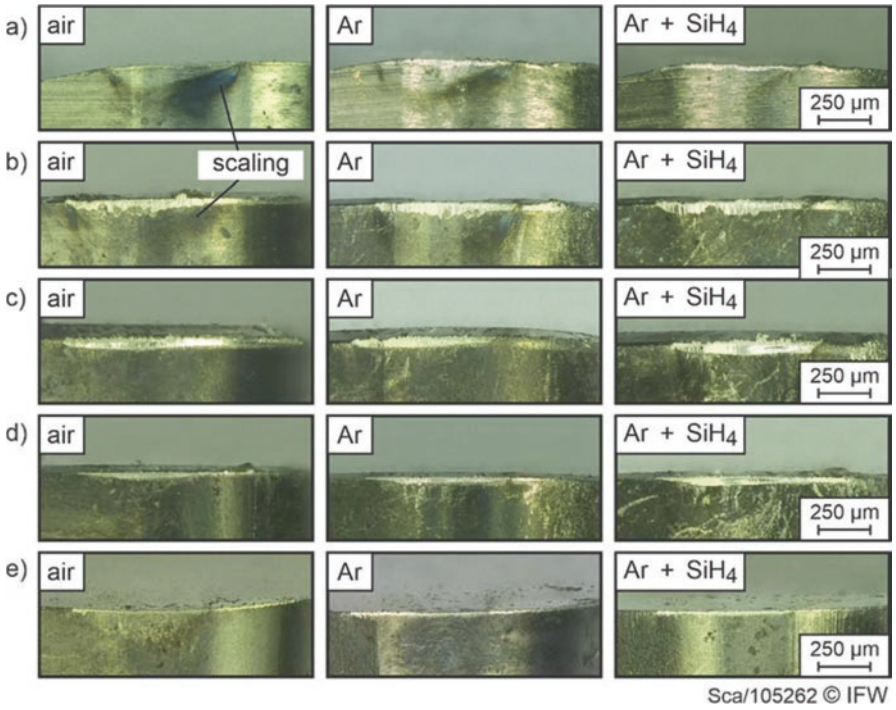


Fig. 3. Flank wear after 2.5 min at a) the carbide tool b) the CBN010 tool c) the CBN200 tool d) the NCB100 tool and e) the PCD tool

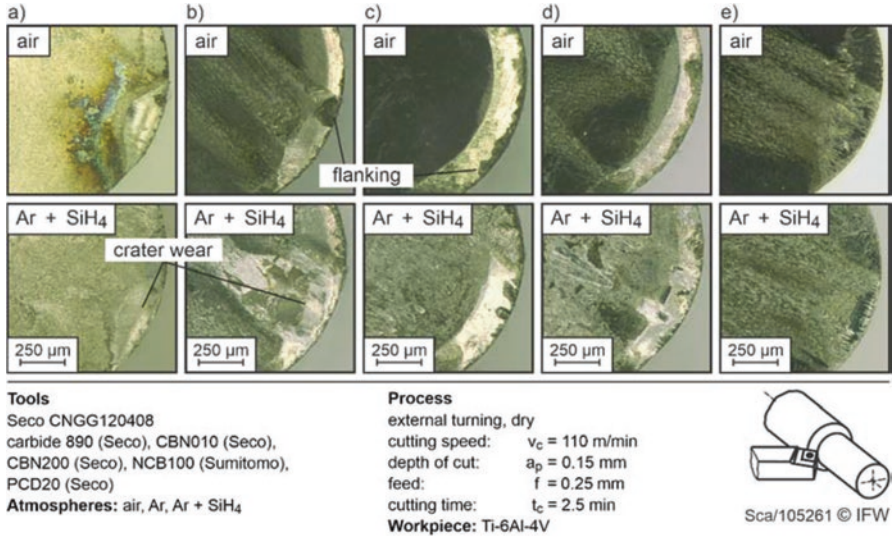


Fig. 4. Rake face after a cutting time of 2.5 min for a) cemented carbide b) CBN010 c) CBN200 d) NCB100 and e) PCD20

In order to investigate the adhesion of the material Ti-6Al-4V with the tool in more detail, the CBN200 tool is examined by scanning electron microscopy and element mapping with respect to titanium elements is used. For the CBN200 tool a cutting time of 2.5 min is considered and the cutting edges are examined under air and silane-doped argon. As can be seen in Fig. 5, more titanium elements adhere to the tool under silane-doped argon, meaning under oxygen-free machining conditions, than under air. This phenomenon can be observed on the flank face but even more significant on the rake face. In addition to a higher density of titanium elements, the area without titanium adhesion under silane-doped argon is much smaller on the rake face compared to air.

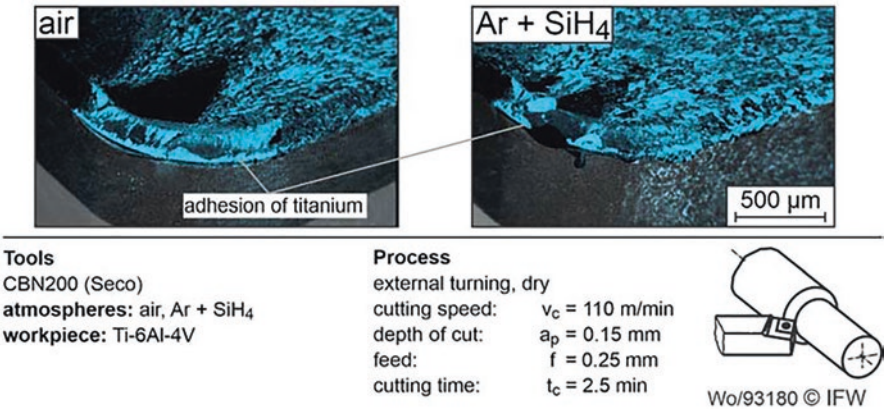


Fig. 5. Element mapping on the rake face of CBN200 tools after 2.5 min cutting time; left: machining under air; right: machining und silane-doped argon

4 Discussion

A possible reason for the increase in tool life of the carbide tool due to the reduction of the oxygen content of the cutting atmosphere can be found in the suppressed scale formation under silane-doped argon. It is possible that the scale layers formed under oxygen-containing atmosphere reduce the wear resistance of the carbide tool. This effect can also be observed on CBN tools containing binder (CBN010 and CBN200) and thus be the reason for the decrease in tool life of the CBN200 tool due to the presence of oxygen.

Furthermore, there is no significant influence of the oxygen content of the ambient atmosphere on the tool life of the PCD tool. Graphitisation has a strongly negative effect on the wear resistance, which cannot be observed under the given process parameters and atmospheric conditions. In future investigations, higher cutting speeds could lead to a more pronounced effect of the atmosphere when machining with PCD.

During machining in a silane-containing atmosphere, the reaction with oxygen produces silicon dioxide, which is deposited as amorphous dust both on the tool and on the workpiece. It is not evident at this point what influence the presence of silicon

dioxide in the effective area between tool and workpiece has on the frictional conditions and wear behaviour of the respective cutting material.

Particulary on PCBN tools increased adhesion of Ti-6Al-4V occurs in the oxygen-free atmosphere, which can reduce the tool life of the binderless PCBN tool. A possible reason for the increased adhesion due to oxygen reduction could be the suppression of oxide formation of Ti-6Al-4V in the presence of oxygen in the tool chip interface. Previous studies have shown that the formation of titanium intermediate oxides and the presence of oxygen during machining of aluminium alloys cause low friction and wear [14, 15]. In the absence of oxygen, increased friction and corresponding temperature increase leads to more adhesion of titanium to the tool. Increased chemical interactions between tool material such as the titanium carbide binder used in the CBN010 tool and titanium in oxygen-free cutting conditions may also be expected due to the absence of the predominant TiO_2 formation in the presence of oxygen, which may also lead to a reduction in tool life [9]. Therefore, titanium-containing binders may be unsuitable for machining titanium alloys such as Ti-6Al-4V, and the prevention of oxide film formation due to an oxygen-free environment may additionally lead to a significantly higher Ti-6Al-4V adhesion to the tool.

As a result, it can be stated that the increased adhesion tendency under oxygen-free conditions leads to increased crater wear and adhesions introduced flanking, which in the case of the binderless PCBN tool can cause cutting edge chipping.

5 Conclusion

Dry turning of titanium alloy Ti-6Al-4V in air, argon and silane-doped argon with an uncoated carbide tool, a PCBN tool with low CBN content, a PCBN tool with high CBN content, a binderless PCBN tool and a PCD tool were investigated. The following conclusions can be drawn from this study:

- A general influence of machining titanium alloy under atmospheres with respect to oxygen level could be verified under all deployed tools.
- A tendency to improve tool life when cutting with uncoated Carbide Tools in low-ogygen atmospheres was observed. The best results are obtained under silane-doped argon,which imitates XHV oxygen levels.
- A reduced scaling formation can be determined as a result of machining under reduced oxygen content on the tools investigated. This has possibly positive effects on the wear behaviour of the tools.
- At the point of action between the tool and the workpiece, higher material adhesion of Ti-6Al-4V on the tool occurs as a result of the reduction of the ambient oxygen content, which is probably caused by changed friction conditions due to the prevention of oxide layer formation, most significant on the PCBN tool white Ti elements in binder.

Acknowledgements. The work presented in this paper has been supported within the framework of the Collaborative Research Centre (Sonderforschungsbereich - SFB 1368) entitled “oxygen

free production" of the German Research Foundation (Deutsche Forschungsgemeinschaft - DFG). The authors gratefully acknowledge the financial support of the German Research Foundation (DFG) within the project.

References

1. Denkena, B., Dege, J.H., Groppe, M., Grove, T.: Zerspanbarkeit von β-Titanlegierungen. VDI-Z Special Werkzeuge, pp. 24–27 (2010)
2. Shaw, M.C.: Metal Cutting Principles. Oxford University Press, Oxford (2005)
3. Sun, S., Brandt, M., Dargusch, M.S.: Machining Ti-6Al-4V alloy with cryogenic compressed air cooling. Int. J. Mach. Tools Manuf. **50**, 933–942 (2010)
4. Chim, Y.C., Ding, X.Z., Zeng, X.T.: Oxidation resistance of TiN, CrN, TiAlN and CrAlN coatings deposited by lateral rotating cathode arc. Thin Solid Films **517**, 4845–4849 (2009)
5. Aspinwall, D.K., Dewes, R.C., Manfle, A.L.: The machining of γ-TiAl intermetallic alloys. CIRP Ann. **54**(1), 99–104 (2005)
6. Denkena, B., Köhler, J., Behrens, L.: Methoden zur Bewertung der Verschleißmechanismen beim Schleifen von PCBN-Wendeschneidplatten, Jahrbuch Schleifen, Honen, Läppen, Polieren: Verfahren und Maschinen, 66th edn. Denkena, B., Hoffmeister, H.-W., Vulkan (2014)
7. Denkena, B., Grove, T., Behrens, L., Müller-Cramm, D.: Wear mechanism model for grinding of PcBN cutting inserts. J. Mater. Process. Tech. **277**, 116474 (2010)
8. Uhlmann, E., Herter, S.: Studies on conventional cutting of intermetallic nickel and titanium aluminides. Proc. Inst. Mech. Eng. [B]: J. Eng. Manuf. **220**, 1391–1398 (2006)
9. Klocke, F.: Fertigungsverfahren 1, 9th edn. Springer, Berlin (2018)
10. Wang, Z.G., Ragman, M., Wong, Y.S.: Tool wear characteristics of binderless CBN tools used in high-speed milling of titanium alloys. Wear **258**, 752–758 (2005)
11. Mercer, A.P., Hutchings, I.M.: The influence of atmospheric composition on the abrasive wear of titanium and Ti-6Al-4V. Wear **124**, 165–176 (1988)
12. Katuku, K., Koursaris, A., Sigals, I.: High-temperature stability of polycrystalline cubic boron nitride cutting tool materials in air. Corros. Sci. **64**, 55–63 (2023)
13. Bagnall, C., Capo, J., Moorhead, W.J.: Oxidation behavior of tungsten carbide-6% cobalt cemented carbide. Metallogr. Microstruct. Anal. **7**, 661–679 (2018)
14. Jaworska, L., Szutkowska, M., Klimczyk, P., Sitarz, M., Bucko, M., Rutkowski, P., Figiel, P., Lojewska, J.: Oxidation, graphitization and thermal resistance of PCD materials with the various bonding phases up to 800 °C. Int. J. Refract. Met. Hard Mater. **45**, 109–116 (2014)
15. Aizawa, T., Mitsuo, A., Yamamoto, S., Sumitomo, T., Muraishi, S.: Self-lubrication mechanism via the in situ formed lubricious oxide tribofilm. Wear **256**, 708–718 (2005)

Influence of Nozzle Position during Cryogenic Milling of Ti-6Al-4V

K. Gutzeit(✉), H. Hotz, B. Kirsch, and J. C. Aurich

TU Kaiserslautern, Gottlieb-Daimler Str., 67663 Kaiserslautern, Germany
kevin.gutzeit@mv.uni-kl.de

Abstract. Ti-6Al-4V was machined by varying the nozzle position during external cryogenic CO_2 cooling. The thermomechanical load was measured and the resulting surface morphology was characterized. The results show a significant influence of the nozzle position on the temperatures in the surface layer. Furthermore, a correlation between the temperatures and the microhardness inside the surface layer of the workpiece was found. This relationship was used to specify an optimal positioning of the nozzle in order to minimize the occurring temperatures.

Keywords: Cryogenic milling · Surface morphology · Ti-6Al-4V

1 Introduction

The surface morphology includes information regarding the topography as well as the metallurgical and mechanical subsurface properties. These properties highly impact the application behavior of a technical component, for example regarding the wear resistance or the fatigue life [1]. Extensive knowledge about the causal correlations between the process parameters, the thermomechanical load and the resulting surface morphology is necessary in order to manufacture components with properties that are suitable to the application [2, 3]. Besides the cutting parameters [4] and the tool properties [5], the cooling strategy has a vast impact on the surface morphology [6, 7], due to the varying effect on heat removal and the resulting temperature in the contact zone.

For a large number of metallic materials, including Ti-6Al-4V, the microhardness of the workpiece surface layer can be increased by machining with cryogenic coolants which are considered as a sustainable and environmentally friendly alternative to conventional cooling strategies [8, 9].

The plastic deformation of the material leads to grain refinement and an increase in dislocation density in the newly created surface layer, resulting in an increase in hardness [10]. When machining with cryogenic cooling, strain hardening is pronounced, since dynamic recrystallization is favored by low temperatures [7]. Furthermore, fewer tensile stresses are introduced into the workpiece surface layer due to the lower thermal load during machining with cryogenic coolants [11].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 284–293, 2021.
https://doi.org/10.1007/978-3-662-62138-7_29

The CO_2 nozzle position can have a significant impact on the flank wear due to a variation in temperature, which was proven by Pittallà et al. [12] during face milling with internal CO_2 cooling through the tool. In previous investigations [13] we observed an influence of the CO_2 nozzle position on the occurring thermal load and the resulting surface morphology during cryogenic turning of stainless steel.

The influence of the nozzle position on the thermomechanical load when milling Ti-6A-4V with external cryogenic CO_2 cooling is still unknown. Since lower temperatures can improve the surface morphology, which is beneficial to the application behavior of the component, the scope of this study is to determine the most efficient nozzle position.

2 Experimental Setup

2.1 Machining Experiments

The experiments are carried out on a 5-axis machining center (Deckel Maho Gildemeister DMU 70 eVolution[1]). The Ti-6Al-4V-workpieces have a cross-section of 10×10 mm^2 and a machining length of $l = 100$ mm. All tests are carried out at a rotational speed of 1592 1/min, resulting in a cutting speed of 50 m/min. Down-milling is used at a feed per tooth of 0.06 mm (feed per revolution = 0.24 mm). The width of cut, as well as the cutting depth are set to 6 mm. CO_2 is extracted from a gas cylinder at room temperature and supplied to the process via two nozzles with an outlet temperature of –78.5 °C [14]. For the applied nozzle outlet diameter of 1 mm the mass flow results in 1.2 kg/min per nozzle. A four-edged end mill (RF 100 Ti - Gühring KG[1]) with a diameter of 10 mm is used as a tool. The manufacturer specifies a cutting-edge radius of 5 µm and a corner radius of 0.5 mm. The rake angle is specified at 5° and the clearance angle at 10°. The tool consists of cemented carbide (DK 460 UF) with a composition of 90% tungsten carbide and 10% cobalt that had a TiAlN coating. The experimental setup is depicted in Fig. 1a.

The position of a single nozzle in relation to the tool is defined by three parameters, which are (see Fig. 1b, c):

- the distance between the nozzle and the tool (d),
- the angle between the nozzle and the end face (α),
- the angle between the nozzle and the feed direction (β).

This adds up to a total of six parameters since there are two individual nozzles. For each of these parameters, three levels are examined (see Table 1). A d-optimal experimental design is created to investigate the influences of the individual parameters (see Table 2), resulting in a total of 39 experiments. These include six experiments around the so-called central point to determine the standard deviation. At the central point, all parameters are set to the second level. The order in which the experiments are carried out is randomized.

Using an evaluation software (Visual-XSel 15[1]), an analysis of variance (ANOVA) is carried out to determine the influence of the examined parameters on the process. The probability-value (p-value) is used as a criterium to indicate whether or not a

parameter has a significant influence. By reaching a p-value of $p<0.05$, the associated parameter can be considered significant with a probability greater than 95%. If the p-value exceeds 0.05, the influence of the parameter is declared non-significant [15].

Table 1. Investigated nozzle positions

Parameter	d_1	α_1	β_1	d_2	α_2	β_2
Level 1	30 mm	15°	0°	30 mm	15°	0°
Level 2	50 mm	30°	30°	50 mm	30°	30°
Level 3	70 mm	45°	60°	70 mm	45°	60°

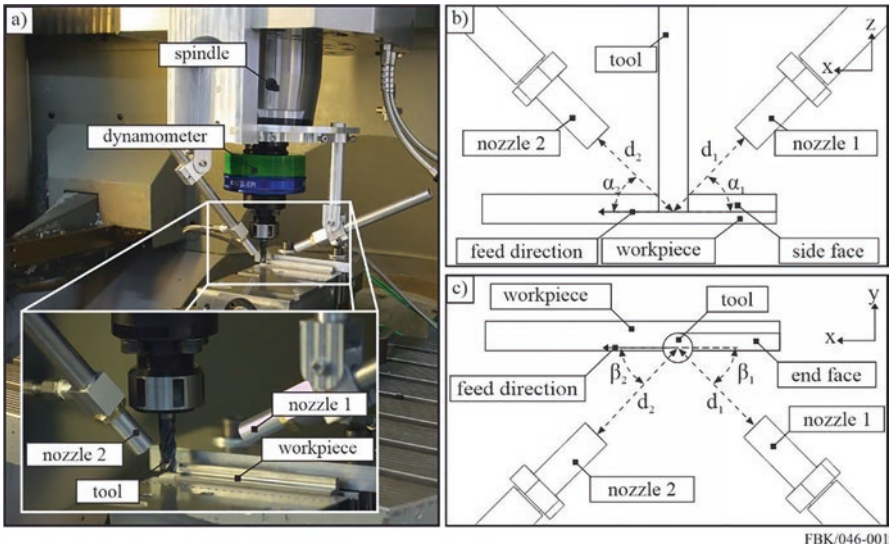


Fig. 1. a) Experimental setup b) Schematic front view of the setup and the parameters examined c) Schematic top view of the setup and the parameters examined

2.2 Measurement Technology

The temperature was measured inside the workpiece using type K thermocouples with a diameter of 1 mm. Six thermocouples were used to measure the temperature inside the end face and the side face. These were arranged at three different positions at each face (see Fig. 2a). The respective distance of the thermocouples along the feed path was 15 mm. The distance between the final surface and the location of the thermocouples inside the workpiece was 1 mm. A representative curve of the temperature as a function of the feed path of the tool is shown in Fig. 2b.

The machining length l indicates when the tool enters and leaves the workpiece. The cryogenic cooling leads to a decrease of the workpiece temperature resulting in a local minimum. Then the temperature curve rises to a local maximum before it begins to fall again. The rise in temperature can be explained by the heat introduced by the tool as a result of cutting. The temperature measured at the local maximum results

from the superposition of the heat introduced by the tool and the heat dissipated by the coolant. The difference ΔT between the starting temperature and the temperature at each local maximum is used to evaluate the cooling performance of the nozzle position.

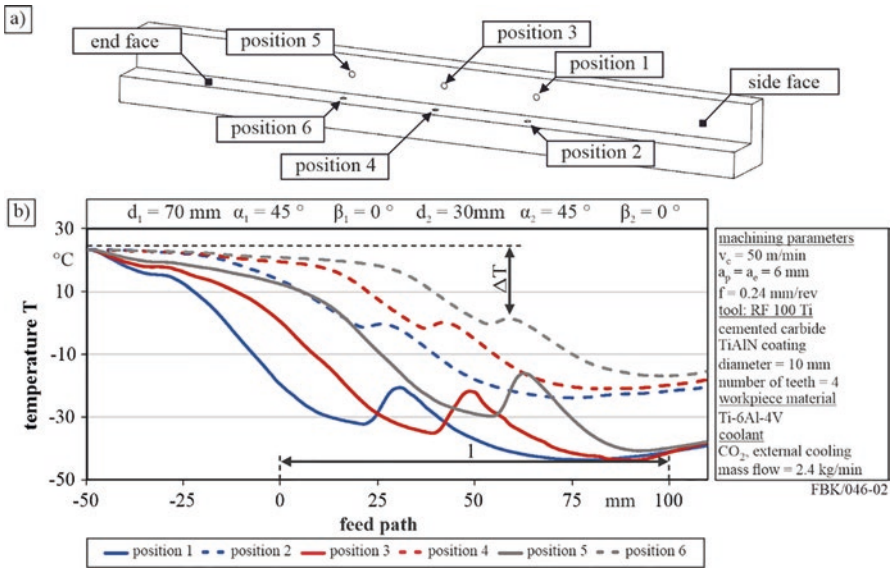


Fig. 2. a) Arrangement of the thermocouples inside the workpiece b) Representative temperature profile

The forces were determined by a rotating dynamometer (Kistler Type 9123C[1]). The passive force F_p results from the cutting force component in the z-direction F_z. The active force F_a is composed of F_x and F_y by vector addition.

The surface roughness was measured with the tactile measurement device MarSurfGD120[1] over a measuring distance of $l_n = 4$ mm ($\lambda_c = 0.8$ mm). On the side face and end face three measurements were carried out each.

The microhardness was measured with a microhardness tester MicroMet 5100[1] with a test load of 0.098 N (HV0.01). For statistical verification, twenty measurements were conducted in the near surface area (distance: 20 µm) of both the side and end face respectively. In order to evaluate a possible increase in microhardness, twenty measurements were also taken inside the core of the workpiece (distance: 400 µm), which is unaffected by the machining process.

3 Results and Discussion

3.1 Thermo-Mechanical Load

The influence of the nozzle position on the determined temperature differences is shown in Fig. 3. The two different nozzles are depicted using different colors, whereas

the faces of the workpiece are separated by different line types. In order to examine a single nozzle parameter, the remaining ones are kept constant in the respective diagram. In addition, the p-values of the examined parameters are depicted for the side and end face.

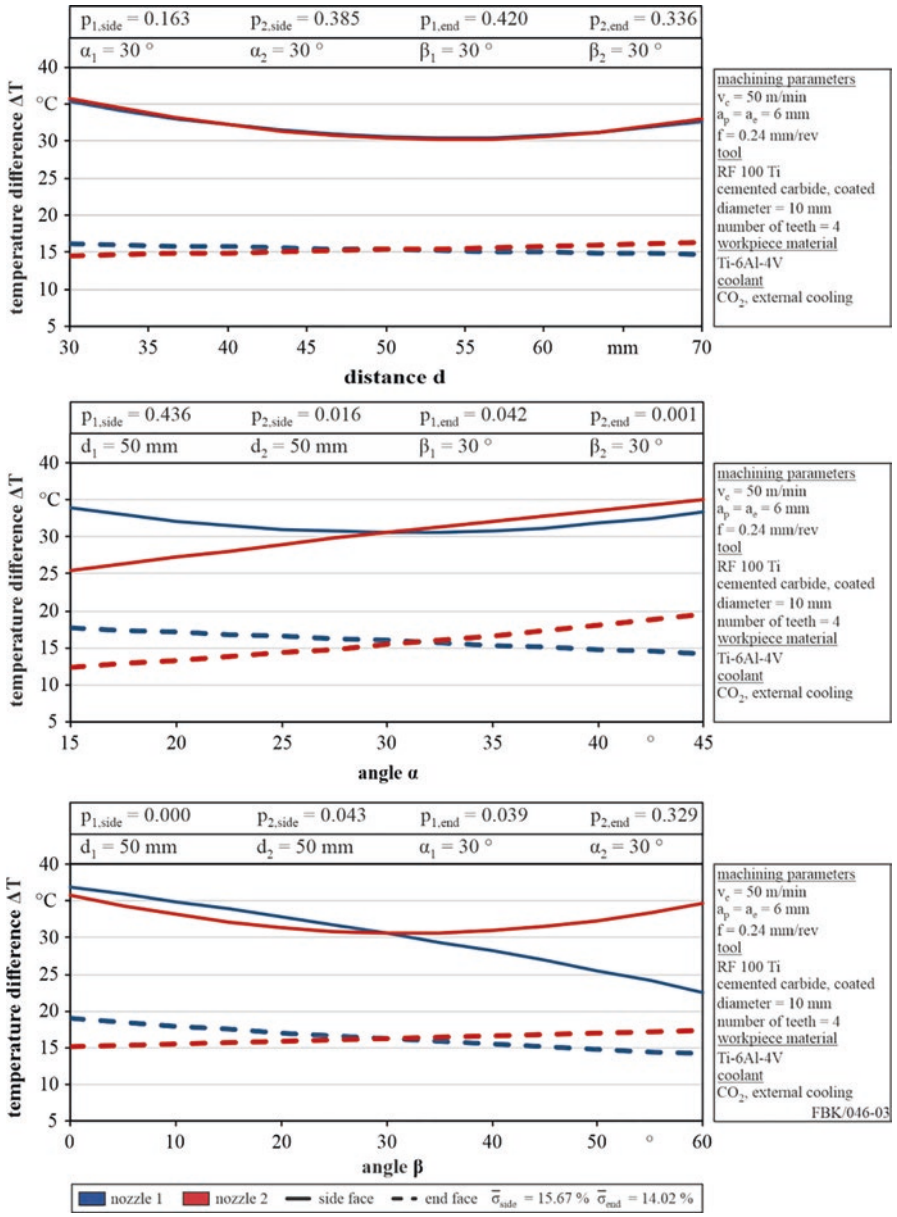


Fig. 3. Influence of the nozzle position on the temperature differences

It can be seen that the nozzle distance has no significant influence on the cooling of the workpiece as every p-value exceeds 0.05. With increasing distance between the workpiece and the nozzle, the time for the coolant to reach the workpiece increases. Hence the time in which the cryogenic medium does heat up as a result of the surrounding room temperature increases as well.

As soon as the CO_2 does heat up, the phase ratio changes. More solid CO_2 particles do sublime. However, the time available to heat up the CO_2 does not seem to be sufficient enough in order to have a significant effect on the temperature due to the small distances investigated. The investigation of the angles shows that the temperature difference at the end face is mainly influenced by α_1 and α_2. The side face is rather influenced by β_1 and β_2.

The temperatures reached at the side face are lower than the ones at the end face, independent of the nozzle position. The lower temperatures result from a better cooling, as more CO_2 is able to reach the side face directly. Less CO_2 reaches the end face directly, as the tool's face side covers the process area. This results in a less effective cooling of the end face and therefore higher temperatures.

For optimum cooling of the end face, α_1 should be set flat, as the CO_2 is more likely to reach the face side of the tool. Contrary α_2 must be set steeply, which results in more CO_2 reaching the contact zone, resulting in better precooling. The improved precooling by a steep α_2 also favors lower temperatures at the side face. However, the most significant influence on the temperature at the side face can be attributed to β_1. At $\beta_1 = 0°$ the lowest temperatures occur, since a large part of the cryogenic medium reaches the contact zone. The angle β_2 should either be set to 0° for better precooling or very steep in order to improve the cooling of the tool. For the values in between, the temperature at the side face increases because a larger amount of CO_2 is wasted to cool material that is removed by the chips and therefore does not reach the side face.

In Fig. 4 the forces are plotted over the associated temperature difference. The active force F_a mainly affects the side face, whereas the passive force F_p affects the end face.

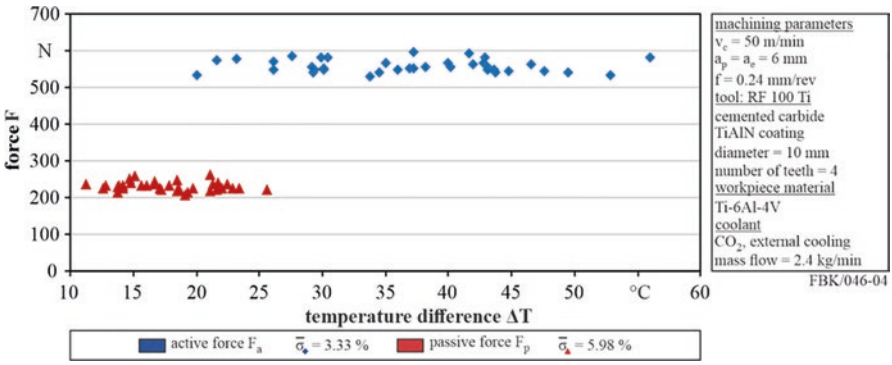


Fig. 4. Force as a function of the temperature difference

It can be stated that there is no significant correlation between the temperature difference and the forces. A possible reason for this is the overlapping of two effects. Lower temperatures can lead to shrinkage of the material. This reduces the chip cross section and thus the forces that occur. However, lower temperatures also cause an increase in the material resistance, which can also result in increasing forces. There is a higher mechanical load at the side face, regardless of the nozzle position.

3.2 Surface Morphology

All p-values regarding the influence of the investigated parameters on the surface roughness exceed $p = 0.05$. The measured roughness at each side lies within the range of $6.59\ \mu m < Rz < 8.14\ \mu m$ at the end face and $1.08\ \mu m < Rz < 2.23\ \mu m$ at the side face respectively. Therefore, it can be deduced that the kinematics during milling have a much higher influence on roughness than the position of the nozzles and the resulting workpiece temperatures.

In Fig. 5 the microhardness is plotted for two selected cases at different distances from the surface. $\Delta T_{Max,\ side}$ represents the case, where the highest temperature difference was measured at the side face ($\Delta T = 55.9 \pm 5.2$ °C). The lowest temperature difference measured at the side face ($\Delta T = 20.0 \pm 2.7$ °C) is represented by $\Delta T_{Min,\ side}$. Compared with the core microhardness at 400 µm, there is an increase in microhardness on both side faces. The increase at the side faces is higher than at the end faces, as higher forces occur. Those favor grain refinement as well as a higher dislocation density, increasing the microhardness. The comparison of the side faces shows that $\Delta T_{Max,\ side}$ has a slightly higher microhardness than $\Delta T_{Min,\ side}$. Since the change in mechanical load can be considered negligible for all cases observed (see Fig. 4), the increase in microhardness can mainly be attributed to the lower temperatures. These counteract thermal softening, which would anneal the dislocations resulting in lower microhardness.

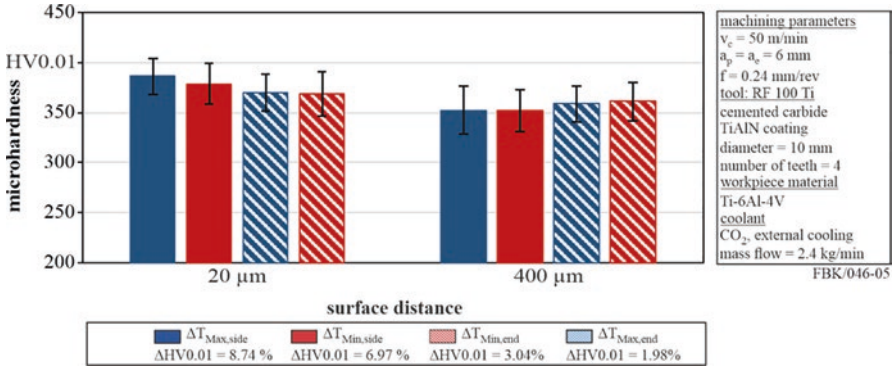


Fig. 5. Increase of the microhardness for the representative cases

4 Conclusion and Outlook

The impact of the nozzle position on the thermomechanical load and the surface morphology was investigated during milling of Ti-6Al-4V with cryogenic CO_2 cooling. It was demonstrated that the examined range of the nozzle distance had no significant influence on the workpiece temperature. The optimal nozzle angles within the investigated range were determined in order to minimize the temperature. The temperature and thus the nozzle position had no significant impact on the forces. A minor increase in microhardness was found in the surface layer of the workpiece side face which increases with decreasing temperatures. The surface roughness at the side face was much lower than at the end face due to the kinematics of the milling process. Both roughness values were not significantly influenced by the nozzle position.

By varying the mass flow and the geometry of the nozzles, the influence of the cooling strategy will further be investigated.

Acknowledgement. This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – project number 172116086 – SFB 926.

[1]Naming of specific manufacturers is done solely for the sake of completeness and does not necessarily imply an endorsement of the named companies nor that the products are necessarily the best for the purpose.

Appendix

See Table 2.

Table 2. Case study

Case	d_1/mm	α_1/°	β_1/°	d_2/mm	α_2/°	β_2/°
1	50	30	30	50	30	30
2	70	15	60	30	30	60
3	70	45	30	70	15	0
4	70	15	60	30	45	0
5	30	15	60	50	15	0
6	50	15	0	70	15	60
7	50	30	30	50	30	30
8	70	15	60	70	15	0
9	30	15	0	30	30	30
10	70	15	0	70	15	0
11	30	45	60	50	30	0
12	30	30	60	30	45	30
13	50	30	30	50	30	30
14	30	45	0	50	45	0
15	70	15	30	70	30	30
16	30	15	0	50	45	60
17	30	45	60	50	15	60
18	70	15	0	70	45	60
19	50	30	30	50	30	30
20	30	30	60	70	45	0
21	30	45	0	70	15	0
22	30	45	60	70	45	60
23	70	45	0	30	15	60
24	50	45	60	30	15	60
25	50	30	30	50	30	30
26	70	45	60	70	30	60
27	50	45	0	70	45	60
28	50	15	60	30	45	60
29	30	30	0	30	15	0
30	50	30	30	50	30	30
31	50	30	30	50	30	30
32	70	45	0	30	45	0
33	30	45	0	30	15	60
34	50	15	30	30	45	0
35	70	45	60	70	45	30
36	70	30	30	30	45	60
37	50	30	30	50	30	30
38	70	30	60	30	15	30
39	30	15	30	70	30	60

References

1. Jawahir, I.S., Brinksmeier, E., M'Saoubi, R., Aspinwall, D.K., Outeiro, J.C., Meyer, D., Umbrello, D., Jayal, A.D.: Surface integrity in material removal processes Recent advances. CIRP Ann. Manuf. Technol. **60**(2), 603–626 (2011)
2. Brinksmeier, E., Meyer, D., Heinzel, C., Lübben, T., Sölter, J., Langenhorst, L., Frerichs, F., Kämmler, J., Kohls, E., Kuschel, S.: Process signatures – the missing link to predict surface integrity in machining. Procedia CIRP 71 – 4th CIRP Conference on Surface Integrity (CSI 2018) **71**, 3–10 (2018)
3. Uebel, J., Ströer, F., Basten, S., Ankener, W., Hotz, H., Heberger, L., Stelzer, G., Kirsch, B., Smaga, M., Seewig, J., Aurich, J.C., Beck, T.: Approach for the observation of surface conditions in-process by soft sensors during cryogenic hard turning. Procedia CIRP 81 – Proceedings of the 52nd CIRP Conference on Manufacturing Systems **81**, 1260–1265 (2019)
4. Rotella, G., Dillon, O.W., Jr., Umbrello, D., Settineri, L., Jawahir, I.S.: The effects of cooling conditions on surface integrity in machining of TiAl6V4 alloy. Int. J. Adv. Manuf. Technol. **71**(1–4), 47–55 (2014)
5. Hotz, H., Kirsch, B.: Influence of tool properties on thermomechanical load and surface morphology when cryogenically turning metastable austenitic steel AISI 347. J. Manuf. Process. **52**, 120–131 (2020)
6. Umbrello, D., Micari, F., Jawahir, I.S.: The effects of cryogenic cooling on surface integrity in hard machining: a comparison with dry machining. CIRP Ann. Manuf. Technol. **61**(1), 103–106 (2012)
7. Pu, Z., Umbrello, D., Dillon, O.W., Jr., Lu, T., Puleo, D.A., Jawahir, I.S.: Finite element modeling of microstructural changes in dry and cryogenic machining of AZ31B magnesium alloy. J. Manuf. Process. **16**(2), 335–343 (2014)
8. Jawahir, I.S., Attia, H., Biermann, D., Duflou, J., Klocke, F., Meyer, D., Newman, S.T., Pusavec, F., Putz, M., Rech, J., Schulze, V., Umbrello, D.: Cryogenic manufacturing processes. CIRP Ann. Manuf. Technol. **65**(2), 713–736 (2016)
9. Pereira, O., Rodríguez, A., Fernández-Abia, A.I., Barreiro, J., López de Lacalle, L.N.: Cryogenic and minimum quantity lubrication for an eco-efficiency turning of AISI 304. J. Cleaner Prod. **139**, 440–449 (2016)
10. Rotella, G., Umbrello, D.: Finite element modeling of microstructural changes in dry and cryogenic cutting of Ti6Al4V alloy. CIRP Ann. Manuf. Technol. **63**(1), 69–72 (2014)
11. Pusavec, F., Hamdi, H., Kopac, J., Jawahir, I.S.: Surface integrity in cryogenic machining of nickel based alloy—inconel 718. J. Mater. Process. Technol. **211**(4), 773–783 (2011)
12. Pittalà, G.M.: A study of the effect of CO_2 cryogenic coolant in end milling of Ti-6Al-4V. Procedia CIRP 77 – 8th CIRP Conference on High Performance Cutting (HPC 2018) **77**, 445–448 (2018)
13. Becker, S., Hotz, H., Kirsch, B., Aurich, J.C., von Harbou, E., Müller, R.: The influence of cooling nozzle positions on the transient temperature field during cryogenic turning of metastable austenitic steel AISI 347. PAMM 18 – Proc. Appl. Math. Mech. **18**(1), e201800447 (2018)
14. Barber, C.R.: The sublimation temperature of carbon dioxide. Br. J. Appl. Phys. **17**(3), 391–397 (1966)
15. Ratnam, C., Arun Vikram, K., Ben, B.S., Murthy, B.S.N.: Process monitoring and effects of process parameters on responses in turn-milling operations based on SN ratio and ANOVA. Measurement **94**, 221–232 (2016)

Lifespan Investigations of Linear Profiled Rail Guides at Pitch and Yaw Moments

S. Ihlenfeldt[1,2(✉)], J. Müller[1], and D. Staroszyk[1]

[1] Institut für Mechatronischen Maschinenbau, TU Dresden, Helmholtzstraße 7a, 01069 Dresden, Germany
steffen.ihlenfeldt@tu-dresden.de
[2] Fraunhofer-Institut für Werkzeugmaschinen und Umformtechnik IWU, Nöthnitzer Str. 44, 01189 Dresden, Germany

Abstract. The "rolling contact related life calculation" (RCRL) of profiled rail guide systems takes the survival probability of every rolling contact into account, that allows a calculation of up to 4 times higher lifespans for pitch and yaw moments on a single guide. New calculation methods allow a user-friendly calculation of the RCRL, which differ from each other by the level of simplification and the computational effort. In this paper, they are summarized and compared to each other.

With the temporary results of the experimental lifespan investigations, the applicability of the RCRL and the mathematical calculation methods can be confirmed.

The mathematical models are based on an iterative calculation of the displacement between the wagon and the guide rail. Based on that, a concept for an indirect measurement of loads on the wagon is presented. Thereby a lifespan calculation at real operating conditions of the profiled rail guide system can be implemented.

Keywords: Profiled rail guides · Lifespan-calculation · Predictive maintenance

1 State of the Art

Profiled rail guides are important components in mechanical and plant engineering and therefore they have to be designed application-specific. With the rolling contact related life calculation (RCRL), lifespans up to 4-times higher than with the conventional, simplified calculation method for pitch and yaw moments on a single profiled rail guide (e. g. [1, 2]) can be calculated [3]. In this case, the single rolling contact forces are not equally over track length. The conventional calculation considers this inadequately and uses the highest contact force of the outer rolling elements for all rolling elements on the track.

The RCRL considers every single rolling contact force by using a single rolling contact related dynamic load rating C_{dyn_sc} and with the life calculation formula for

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 294–303, 2021.
https://doi.org/10.1007/978-3-662-62138-7_30

profiled rail guides the survival probability of every single rolling contact can be calculated. The product of the survival probabilities of the single rolling contacts equates to the survival probability of the profiled rail guide, which is typically defined as 90%. With that, a life span can be calculated iteratively for the considered loading condition [1–3].

The load-dependent single rolling contact forces are input variables for the RCRL and they can be calculated by FE models, analytical or simple numerical calculation models. They are described in Sect. 2.

2 Calculation Models for the Single Rolling Contact Forces

The load-dependent single rolling contact forces can be calculated with the described models. They should consider requirements like:

- calculability of all profiled rail guide sizes,
- consideration of preloads,
- calculability of arbitrary superposed external forces,
- the usage of known or easy to find parameters (e. g. dynamic load rating, preload, rolling element diameters, geometric dimensions of the CAD model, …) and
- the calculability with usually available software.

These requirements cannot be met by all the models completely. At the end of the related sections, the advantages and disadvantages of each calculation model are listed. In Sect. 2.5, the calculation models will be assessed concerning usability.

2.1 FE-Models

Neidhardt introduces a method for modelling a profiled rail guide that can be simulated in a computationally efficient way [3]. On that basis, a profiled rail guide with size 25 and an O-bearing arrangement is modelled (see Fig. 1) [4]. Because there is no applicable measurement method for measuring the single rolling contact forces yet, the FE-model serves as a reference for the developed calculation models explained in Sect. 2.2 et seq.

Instead of modelling the rolling elements as computationally intensive full models, they are modelled as nonlinear springs with a characteristics curve that represents the Hertzian deflection of the rolling contact.

The preload of the profiled rail guide is realised by a thermal stretching layer in the middle of the guide rail.

The static behaviour of the profiled rail guide is very sensitive to the chosen boundary conditions. They have to be chosen in that way, that real pivots and deflections of the guide carriage towards the guide rail are represented.

The connection of the nonlinear springs to the guide rail and the guide carriage is a focus in modelling because of its significant influence on the calculated single rolling contact forces. Therefore, multiple different variants were modelled, whereby the method with rigid inlays for every single rolling contact on the track has shown the best and plausible results.

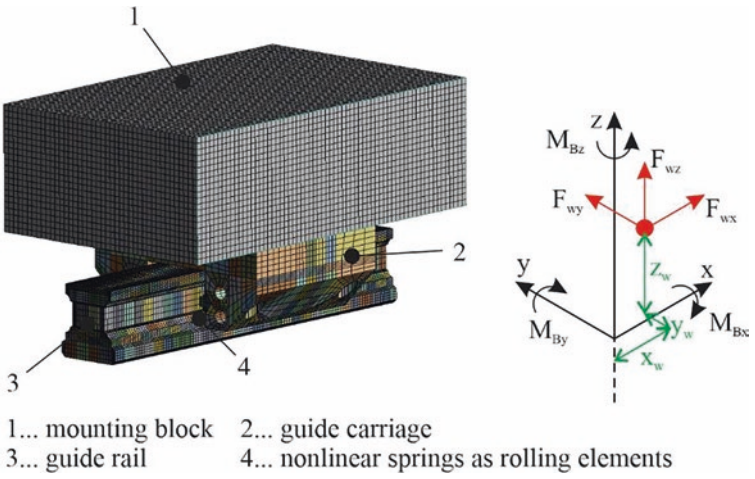


Fig. 1. FE-model of a profiled rail guide with the defined coordinate system [4]

Because the Hertzian deflection in the rolling contact is represented in the characteristics curve of the nonlinear springs, local deflections at the connection-nodes in the FE-model have to be avoided. The transition behaviour between the elastic rolling contact and the elastic guide carriage or elastic guide rail can only be estimated and requires further investigations. Because the deflection of the FE-model is verified with the deflection of a real profiled rail guide, the FE-model is considered sufficiently accurate.

Another point is the degree of simplification of the FE-model towards the real profiled rail guide. Figure 2 shows the resulting single rolling contact forces along a track at a centric load of $F_z = -14300$ N. Thereby the three exemplary details:

– track to track returns at the guide carriage,
– breakouts at the edge of the guide carriage and
– elastic inlet geometries

were added, simulated and compared to each other.

Advantages of FE-models for calculating single rolling contact forces

1. The elastic behaviour of the real profiled rail guide is well approximated.
2. Superposed external forces can be calculated without restrictions.

Disadvantages of FE-models for calculating single rolling contact forces

1. A lot of knowledge in modelling the FE-models is necessary.
2. On an actual PC, the simulation time is about 45 min.
3. Expensive FE-software for the simulation is necessary.

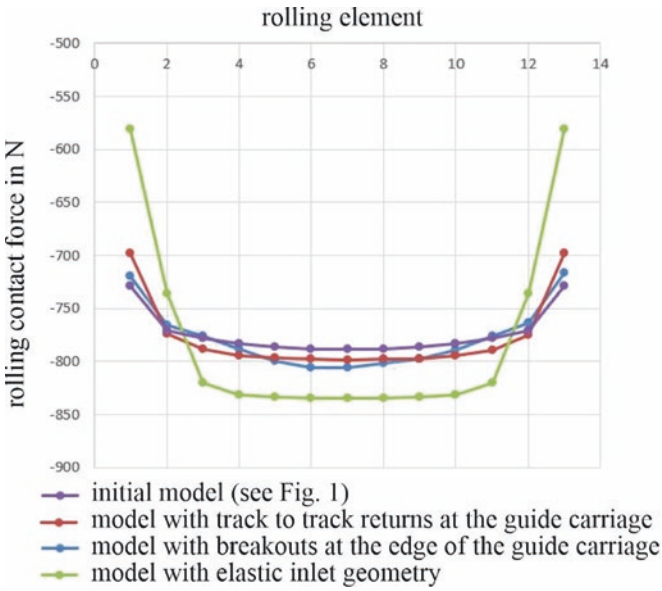


Fig. 2. Single rolling contact forces for different modelled details of the profiled rail guide

2.2 Analytical Model

An analytical and thus fast computing model for calculating the single rolling contact forces is presented and derived in [5].

The approach is based on a linearization of the single rolling contact force curve along the track. The parameter of the linear force curve are calculated by the balance between resulting single rolling contact forces and the acting external load. In the preload range, the calculated single rolling contact forces are adjusted by known analytical formulas for calculating effective forces of profiled rail guides.

Advantages of the analytical method for calculating single rolling contact forces

1. The model can be calculated analytically.
2. The model can be calculated fast and without special software.

Disadvantages of the analytical method for calculating single rolling contact forces

1. The nonlinear Hertzian deflection of the rolling contact is linearized. With low external loads, the single rolling contact forces are approximated good, but with high external loads the deviations, compared to the FE-model, increase.
2. The elastic behaviour of the guide carriage and the guide rail is unconsidered.
3. Superposed external loads cannot be calculated without restrictions.
4. Deflections and torsions of the guide carriage to the guide rail are not calculated and therefore they cannot be used for further investigations.

2.3 Numerical Model with Rigid Guide Carriage and Rigid Guide Rail

A numerical but fast calculable method is described in [6], which considers the Hertzian deflection in the rolling contact area.

The model calculates the deflection in the rolling contact, whereby the deflection of the guide carriage towards the guide rail is defined as linear along the track. The single rolling contact forces can be calculated analytically with the calculated deflections in the belonging rolling contact. To calculate the parameters of the linear deflection of the guide carriage, a numerical solver is used in order to fulfil the force and moment balance between the external load and the single rolling contact forces.

The geometric description of the model is shown in Fig. 3a. It corresponds to the geometric structure of the FE-model.

Advantages of the rigid numerical method for calculating single rolling contact forces

1. The nonlinear deflection of the Hertzian rolling contact and the preload of the profiled rail guide are considered.
2. The calculation needs a simple numerical Solver, which is, for example, included in Excel. Other special software is not needed.
3. Superposed external loads can be calculated.

Disadvantages of the rigid numerical method for calculating single rolling contact forces

1. The model cannot be calculated analytically.
2. The computational effort is higher than with the analytical method (Sect. 2.1).
3. The elastic behaviour of the profiled rail guide is just considered in the rolling contact. The other elasticities of the guide carriage and the guide rail are unconsidered. Therefore the calculated deflections of the guide carriage towards the guide rail are, compared to the real profiled rail guide, too small.

2.4 Numerical Model with Elastic Guide Carriage and Elastic Guide Rail

The model for calculating the single rolling contact forces described in Sect. 2.3 can be extended by the stiffnesses of the guide carriage and the guide rail [6].

Therefore, nonlinear and linear stiffnesses are parameterised only one-time with the FE-model and a centric load on it. The deflection of the profiled rail guide can then be calculated for arbitrary load cases. Figure 3b shows the geometric model with added elastic deflections Δy_{ela_gc}, Δz_{ela_gc}, Δy_{ela_gr} and Δz_{ela_gr}.

The bending of the edges of the guide carriage are considered as nonlinear elastic deflections. Therefore, the differential equation of the bend line for a cantilever beam at every rolling contact area is calculated. The result is a nonlinear deflection $\Delta y_{ela\_gc\_nlin}$ at every rolling contact of the guide carriage [4]. Linear deflections $\Delta y_{ela\_gc\_lin}$ are superposed to the nonlinear deflections and in the calculation model they have to be added in order to get Δy_{ela_gc} (see Fig. 3b).

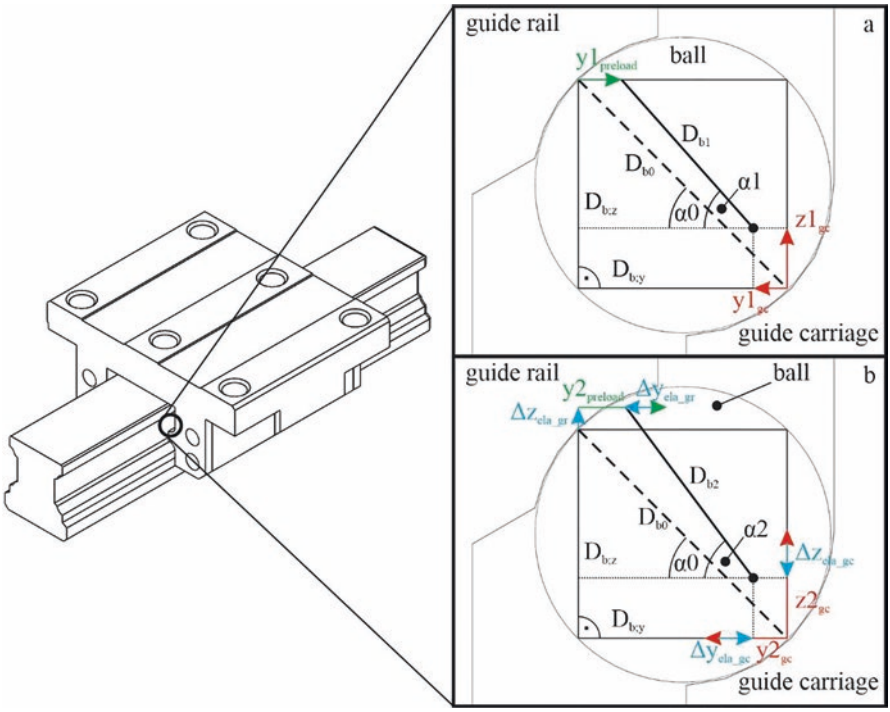


Fig. 3. Geometrical description of the numerical calculation methods (a – rigid guide carriage and rigid guide rail; b – elastic guide carriage and elastic guide rail)

Advantages of the elastic numerical method for calculating single rolling contact forces

1. This calculation model has the same advantages as the rigid model described in Sect. 2.3
2. In addition, the elastic behaviour of the guide carriage and the guide rail can be considered.
3. The calculated deflection of the guide carriage towards the guide rail can be used for further investigations (Sect. 4).

Disadvantages of the elastic numerical method for calculating single rolling contact forces

1. The model cannot be calculated analytically.
2. The computational effort is higher than with the analytical method (Sect. 2.1) and the rigid numerical method (Sect. 2.3)
3. For the one-time parameterisation of the calculation model, an expensive FE-software and knowledge in modelling profiled rail guides are necessary. However, this step can be done by the manufacturer of the profiled rail guides. End users do not need any special software.

2.5 Assessment of the Calculation Models

The models for calculating single rolling contact forces described in Sect. 2.1 to 2.4 can be used for the RCRL. They differ from each other by the computational effort and the calculable results.

From the authors point of view, the below ranking based on the applicability of the models for calculating rolling contact forces is determined.

1. numerical rigid model (Sect. 2.3)
2. numerical elastic model (Sect. 2.4)
3. analytical model (Sect. 2.2)
4. FE-model (Sect. 2.1)

The numerical rigid model for calculating single rolling contact forces fulfils the most of the defined requirements, it is calculable relatively easy and there is no need of expensive FE-software. The differences of the RCRL results between the single rolling contact forces calculated with the numerical rigid model and the FE-model is equal to 4%. Because of the statistical character of the lifespan value for profiled rail guides, this is of no consequence. However, with the elastic numerical method deflections of the profiled rail guide at arbitrary load conditions can be calculated, whereby additional use cases are opened up (see Sect. 4). The differences between the RCRL results with the analytical calculation model and the FE-model are up to 10% and there is no possibility of calculating superposed external loads. For simple load conditions, the analytical model in combination with the RCRL is still a significant improvement compared to the conventional lifespan calculation method for profiled rail guides [6].

The FE-model represents the real profiled rail guide best, but there is a lot of knowledge, much modelling and computational effort and expensive FE-software necessary. Therefore, this method is not applicable for most end users, who want to use the RCRL in their engineering process.

3 Lifespan Investigations

The RCRL is not experimentally validated, so far. Therefore, the "Institut für Mechatronischen Maschinenbau der TU Dresden" conducts lifespan investigations for profiled rail guides of size 25 and balls as rolling elements, according to the standard DIN 631 [7]. The mandatory finger test for damage detection of a test item is assisted by acceleration sensors on every test item. This improves the damage detection and with that the results of the investigations.

Within a first run of the lifespan investigations, the profiled rail guides are centrally loaded, in order to verify the given dynamic load rating $C_{dyn} = 28600$ N. With a second run, the profiled rail guides are loaded with a pitching moment, in order to verify the RCRL. Table 1 gives an overview of the used test parameters and the test results.

The, with standard DIN 631, investigated lifespan L_{B10} with a 10% probability of default is $L_{B10} = 498.12$ km at the first run. This value is considerably smaller

than the calculated lifespan L = 800 km. The dynamic load rating of the profiled rail guide cannot be confirmed. The experimentally determined dynamic load rating is $C_{dyn\_exp} = 24421.91$ N.

Table 1. Test parameter and test results of the lifespan investigations

	Catalogue central	RCRL central	Catalogue eccentric	RCRL eccentric
	Test parameter			
Size	25			
Rolling elements	balls; diameter 4.75 mm			
C_{dyn} (C_{dyn_sc})	28600 N (3232.045 N)			
Longitudinal load	290 Nm			
Load	$F_z = -14300$ N		$F_z = -5250$ N; $x_W = 35$ mm	
Calculated lifespan L	800 km	800 km	183.24 km	793 km
	Test results			
Exp. lifespan L_{B10}	498.12 km		531 km[a]	
C_{dyn_exp} ($C_{dyn\_exp\_sc}$)	24421.91 N (2755.84 N)		/	
Calculated lifespan with C_{dyn_exp}	498.12 km	/	116.65 km	/
FEM	/	498.12 km	/	487.84 km
Numerical rigid	/	500.37 km	/	471.36 km
Numerical elastic	/	498.12 km	/	492.23 km
Analytical	/	506.77 km	/	493.78 km

[a]26 of 30 test items finished (value is changing because of still running exp. investigations)

The pitching moment load on the profiled rail guide at the second run results in a calculated lifespan of L = 793 km with the RCRL and $C_{Dyn\_sc} = 3232.045$ N (catalogue value). With the experimentally determined lifespan of $L_{B10} = 531$ km, the RCRL cannot be verified, also. If the experimentally determined dynamic load rating of $C_{dyn\_exp} = 24421.91$ N ($C_{dyn\_exp\_sc} = 2755.84$ N, first run) and the load parameter of the second run are estimated for the RCRL, a lifespan of L = 487.84 km (FE-model) is calculated. With this approach and referring to DIN 631, L_{B10} is greater than L and therefore the RCRL can be verified.

4 Concept for a Lifespan Calculation at Real Operating Conditions

4.1 Problem Statement

In the conducted lifespan investigations (Sect. 3), the achieved lifespans of the test items ranged from 202.48 km to 2366.78 km (first run) and from 180.23 km to

3773.68 km (second run). This shows the risk for users, if the profiled rail guide fails before the calculated lifespan, but also the high potentials, because most of the profiled rail guides could be used much longer than calculated, in actual operations. In addition to that, the loads on the profiled rail guide cannot always be predicted, because of changing applications of the machine. Thereby the lifespan calculation in the design process could lose validity. Further operating conditions such as constraining forces, caused by manufacturing and mounting tolerances or thermal elongations in the surrounding construction, have an influence of the real lifespans of profiled rail guides. Therefore, the aim of a research project is to capture the real operating conditions of the profiled rail guides and calculate the expected residual life during operation continuously.

4.2 Application Principle

Capacitive distance sensors mounted to the guide carriage measure the distance between the guide carriage and the guide rail continuously and with a high resolution (see Fig. 4). With the approach of the calculation method described in Sect. 2.4, the single rolling contact forces and the RCRL can be calculated control system integrated. With the principle of linear or nonlinear damage accumulation, the residual life of the profiled rail guide can be calculated for changing external (constraining) loads.

The measured distance signals between guide carriage and guide rail are analysed, whereby different influence factors like

- the surface quality of the guide rail, on which the capacitive distance sensors measure (e. g. waviness of the guide rail along travel length, roughness, hole coverings),
- characteristic deflections of the guide carriage due to the run in and run out of the balls into the load zone,
- characteristic damage frequencies and a lot more

have to be considered or compensated.

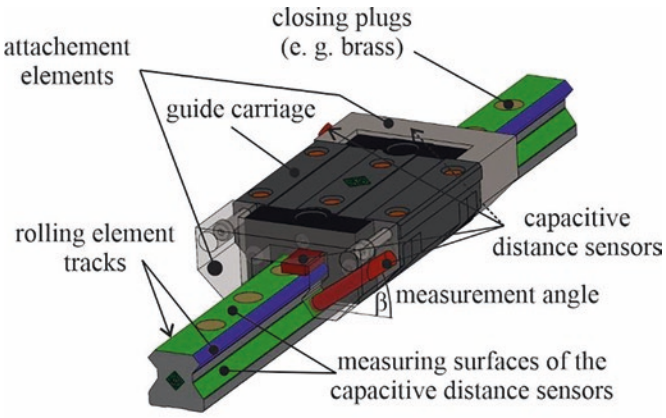


Fig. 4. Capacitive distance sensors mounted to the guide carriage

4.3 Benefit

With the knowledge of the residual life in consideration of the real operating conditions, maintenance measures on machines and plants can be planned more exactly and unplanned downtimes could be avoided. At the same time, the service life potential of the profiled rail guides can be exploited with significant lower risk for the user.

At commissioning of feed axes in machines, the system can detect failures or exceeded tolerances in the surrounding constructions and considers this when calculating the residual life.

5 Summary

In this paper, four different methods for calculating single rolling contact forces for the rolling contact related life calculation are presented. These methods differ from each other in the degree of simplification and thereby in the calculable results.

Lifespan investigations with a pitching moment on a single profiled rail guide have shown that the RCRL can be used in principle. However, there have to be conducted more lifespan investigations in order to confirm the results.

The gained knowledge can be used to determine the influence of real operating conditions on the real lifespan of the profiled rail guide. A scientifical and economical interesting implementation approach is described in the last sections.

References

1. DIN 637:2013-08, Wälzlager- Sicherheitstechnische Festlegungen für Dimensionierung und Betrieb von Profilschienenführungen mit Wälzkörperumlauf, DIN Deutsches Institut für Normung e.V., Beuth, Berlin (2013)
2. Bosch Rexroth, A.G.: Handbuch der Lineartechnik (2006)
3. Neidhardt, L.: Wälzkontaktbezogene Lebensdauer von Profilschienenführungen – Bewertung der experimentellen Ermittlung des Lebensdauerkennwertes (2013)
4. Ihlenfeldt, S., Müller, J., Staroszyk, D.: Lebensdauer von Profilschienenführung unter Momentenbelastung, Manuskript zur 13. VDI Fachtagung „Gleit- und Wälzlagerungen 2019: Gestaltung – Berechnung – Einsatz“, Schweinfurt, 06/2019
5. Ihlenfeldt, S., Müller, J., Staroszyk, D.: Lebensdauer von Profilschienenführungen – Teil 1. wt Werkstattstechnik online **109**(5) (2019)
6. Ihlenfeldt, S., Müller, J., Staroszyk, D.: Lebensdauer von Profilschienenführungen – Teil 2. wt Werkstattstechnik online **110**(3) (2020)
7. DIN 631:2020-03, Wälzlager – Prüfbedingungen zur versuchstechnischen Verifikation der dynamischen Tragzahl von Profilschienenführungen mit kompakten Kugel- oder Rollenumlaufwagen, DIN Deutsches Institut für Normung e.V., Beuth, Berlin (2020)

Towards the Prediction of Compliance Influences on Shape Deviations in Internal Traverse Grinding

N. Schmidt[1(✉)], T. Tsagkir Dereli[1], T. Furlan[2], R. Holtermann[2], D. Biermann[1], and A. Menzel[2,3]

[1] Institute of Machining Technology, TU Dortmund University, Baroper Straße 303, 44227 Dortmund, Germany
nils.schmidt@tu-dortmund.de
[2] Institute of Mechanics, TU Dortmund University, Leonhard-Euler-Straße 5, 44227 Dortmund, Germany
[3] Division of Solid Mechanics, Department of Construction Sciences, Lund University, P.O. Box 118, 22100 Lund, Sweden

Abstract. Internal traverse grinding (ITG) with electroplated cBN tools is a highly efficient process for machining of precision bores. Profiled tools allow for high stock removal and good surface quality in a single axial stroke. However, process control is difficult. Especially shape deviations of machined workpieces are influenced by machining system compliances. Simulation-based solutions can be used to predict and compensate shape errors.

To model the system compliance, deformations have been measured using eddy current sensors in combination with piezoelectric force measurement components. A simplified substitute model has been developed, which represents the deflections of the entire machining system as a function of the process normal force, and grinding investigations with in-process force measurements have been performed. By incorporating the measured forces and the compliance model into an existing simulation system, the influence on the resulting shape of the bore has been predicted in good accordance with the real grinding process.

Keywords: Internal traverse grinding · Compliance · Shape deviations

1 Introduction

The increasing quality requirements for industrial applications demand continuous improvement of tools and machining processes. The grinding of a workpiece usually takes place at the end of a process chain and is often used for the hard and fine machining. Keeping high shape, dimensional and positional tolerances as well as a good surface quality of the workpiece is required [1]. Short machining cycle times also play an important role. For internal grinding of components like bearing and gear

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 304–314, 2021.
https://doi.org/10.1007/978-3-662-62138-7_31

parts, it is difficult to achieve these goals. In industrial applications, these workpieces are most frequently machined using internal plunge grinding (IPG) [2, 3]. In IPG, the full width of the grinding wheel is in contact with the workpiece surface and high surface quality can be achieved. However, this manufacturing process is limited in terms of the material removal rate. Due to the large contact length, high normal forces occur even at low radial feed rates, leading to elastic deformations of the machining system, consisting of machine tool, grinding tool, tool holder, tool spindle, workpiece clamping device and the workpiece itself. In consequence, the relative position of workpiece and tool shifts, which results in shape deviations [4]. High tool wear in internal grinding, resulting from size constraints of the tool and thus a small number of active abrasive grains, further increases the process forces and the resulting shape deviations.

A highly efficient process for machining of bores is internal traverse grinding (ITG) with electroplated cBN tools characterised by a functional split into roughing and finishing zones. This allows a combination of rough grinding and finish grinding in a single axial stroke [4]. Prior to contact with the workpiece, a total radial stock removal $a_{e,tot}$ is set and the feed movement takes place parallel to the inner contour of the workpiece with an axial feed rate v_{fa} [2]. Material removal primarily occurs at the conical roughing zone, while the cylindrical finishing zone smoothens the surface of the workpiece, see Fig. 1.

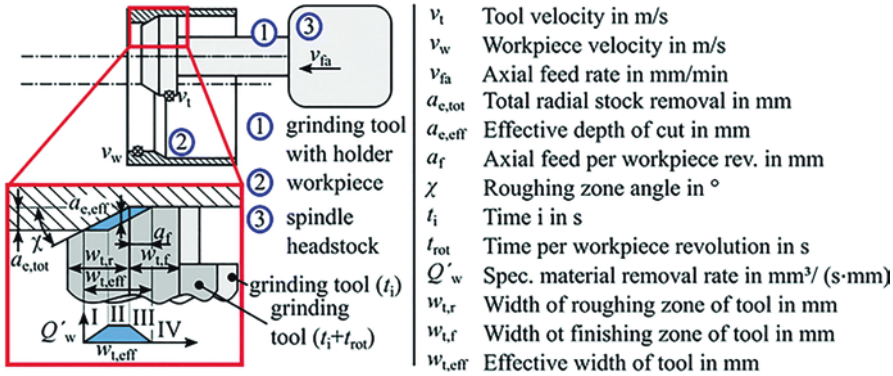


Fig. 1. Fundamental kinematics of ITG.

The distribution of the material removal over the areas I, II and III of the grinding tool depends on the total radial stock removal $a_{e,tot}$, the roughing zone angle χ and the axial feed a_f per workpiece revolution. The effective tool width $w_{t,eff}$ can be calculated as the sum of the contact width in the roughing zone and the width of the material removal in the finishing zone which equals to the axial feed a_f. Because $w_{t,eff}$ in ITG is small compared to IPG, low normal and tangential forces occur in ITG and high material removal rates can be achieved [2, 5]. However, especially in high performance processes, shape errors occur [6]. The heat generated in the process is concentrated on a small area and causes local thermal expansion of the workpiece [6, 7]. In addition, shape deviations are caused by elastic deformations of the machining system [6, 8].

In order to ensure good manufacturing accuracy, simulation-based compensation strategies for shape deviations resulting from workpiece clamping and thermal expansion have been developed and the profile error has been reduced substantially [6, 9]. Machining system deflections have not been considered in these works. However, these represent an important aspect for improving the production quality in internal grinding [8, 10]. A variety of models exists for individual aspects of the machining system, e.g. the dynamics of spindles [11] or contact stiffness [12]. Research on machining system compliance in internal grinding focuses on quasistatic models. The tool and its holder have been modelled using beam theory in, for example, [5] and [8]. A similar modelling approach for ITG that takes the workpiece holder into account has been proposed in [2].

In this article, the compliances of the entire machining system as well as the tool system, consisting of the tool, the tool holder, the tool spindle and the components of the machine tool connecting the spindle with the machine base, are determined experimentally by means of static analogy tests similar to [2]. All parts of the entire system which are not part of the tool system, i.e. workpiece, clamping device and machine tool excluding the tool system, are referred to as the workpiece system.

The influence of the axial tool position is investigated. A simplified model is proposed to represent the deflection of the tool system as a function of the normal forces. Additionally, a simple approach to approximate the deflection of the entire machining system in the contact zone is presented. Subsequently, grinding investigations and geometric-kinematic simulations have been carried out to determine the influence of the compliance on shape deviations.

2 Determination of Machining System Compliances

In ITG, a normal force F_n occurs as soon as the tool starts to interact with the workpiece. With increasing normal force, workpiece and grinding wheel are moved apart due to the compliances of the machining system. As a result, the prescribed total radial stock removal $a_{e,tot}$ is not completely achieved during machining and the resulting inner diameter along the length of the bore is too small. However, entry and exit area of the workpiece contour are exceptions. As the roughing zone of the grinding tool starts to exit the workpiece, the specific material removal rate Q'_w is reduced and the system rebounds due to the decreasing normal force. This results in large shape deviations along the bore length. At the entrance area a similar effect occurs to a minor extent [9].

2.1 Experimental Setup for Compliance Tests

Static analogy tests have been performed in a setup based on the machining of bearing inner rings to investigate the influence of system compliances. All experimental investigations have been performed, in collaboration with *Schaeffler Technologies AG & Co. KG,* on an internal cylindrical grinding machine equipped with the high-speed spindle HV-P 120 - 45.000/18 of the company *GMN Paul Müller Industrie GmbH*

& Co. KG. An electroplated cBN tool with roughing and finishing zone widths of $w_{t,r} = w_{t,f} = 4$ mm and a grain size of B181 in both sections, provided by *August Rüggeberg GmbH & Co. KG*, has been used. The combination of grinding wheel and tool holder is produced from one piece and referred to as the tool in the following. The test setup has been designed to determine the compliance of the tool system and the entire machining system, see Fig. 2.

The tool is positioned at a prescribed axial position z_f relative to the clamped workpiece, while maintaining a radial gap of approximately 5 µm. Subsequently, the tool is moved radially in 30 steps of 0.5 µm in direction f_r towards the workpiece and, thereafter, back to its original position. Neither tool nor workpiece rotate during this static analogy test. After contact is established, the resulting normal force F_n induces a deflection of the entire machining system $\Delta a_e = \Delta a_{e,t} + \Delta a_{e,w}$, corresponding to a reduction of the set total depth of cut $a_{e,tot}$ in ITG. The quantities $\Delta a_{e,t}$ and $\Delta a_{e,w}$ refer to the contributions of tool system and workpiece system, respectively.

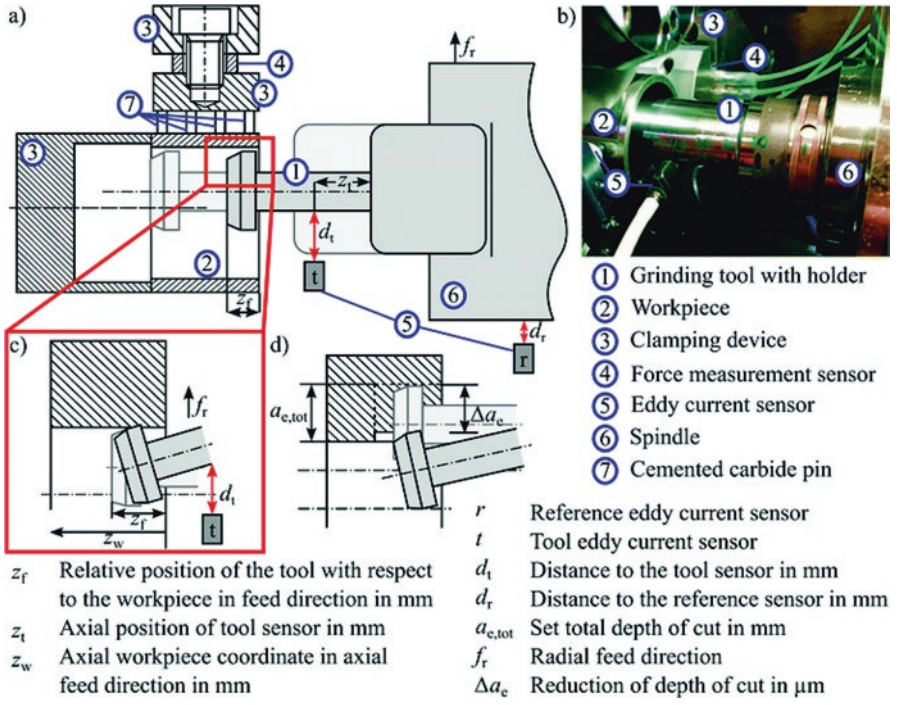


Fig. 2. Experimental setup for compliance tests. a) Schematic representation; b) Photograph: Tool sensor at $z_t = 40$ mm; c) Displacement of the tool system due to movement in radial direction in static compliance tests; d) Machining system deflections reduce the total radial stock removal $a_{e,tot}$ by Δa_e in ITG.

A modified radial clamping device has been constructed, which incorporates a pre-tensioned piezoelectric force sensor (*Kistler 9251A*). The z-axis of the sensor is aligned with the direction f_r in order to measure the normal force F_n. Two

eddy-current sensors (*Micro-Epsilon NCDT 3010-A, sensor U1*) have been used to measure the distances d_t, at a defined position along the tool, and d_r, at a reference position at the base of the tool system. The reference sensor r has been kept at the same position for all measurements, while the other sensor t has been moved to different positions z_t to resolve the deflection along the tool. A limited number of experiments has been performed to investigate the influence of different relative tool positions z_f.

2.2 Experimental Results

Force and distance data (cf. Sect. 2.1) have been analysed using the Python package pandas [13]. An exponential moving mean filter with a half-life of 40 ms has been applied to reduce high-frequency oscillations. The initial distance measured by each eddy current sensor (d_{t0}, d_{r0}) has been approximated by the median of the values corresponding to the first second of the measurement, in which the tool was kept stationary, and has been used to calculate the distance changes as $\Delta d_t = d_t - d_{t0}$ and $\Delta d_r = d_r - d_{r0}$. One representative measurement is illustrated in Fig. 3.

The tool system deflection $u_t = \Delta d_r - \Delta d_t$ at the position of the tool sensor has been approximated for each experiment by the linear fit $u_t(F_n) = b_t + c_t F_n$, shown in Fig. 3c. Based on this approximation, the compliance c_t of the tool system at the tool sensor position has been determined. Since no deflections are expected when no normal force is applied, the value of b_t has been neglected.

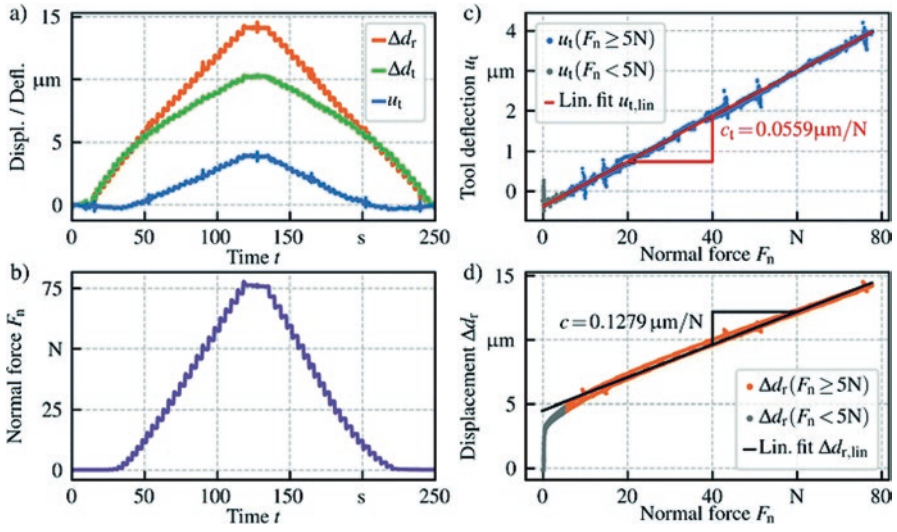


Fig. 3. Plot of filtered data for an exemplary measurement at $z_t = 50$ mm, $z_f = 10$ mm; a) Displacements Δd_r, Δd_t and tool system deflection $u_t = \Delta d_r - \Delta d_t$ at the tool sensor position; b) Normal force F_n over time; c) Tool system deflection u_t and its linear fit $u_{t,lin}$ over normal force F_n; d) Displacement at reference sensor Δd_r and its linear fit $\Delta d_{r,lin}$ over normal force F_n.

2.3 Modelling of Machining System Compliances

Following [2], a model based on classic Bernoulli beam theory has been adopted for the tool system, see Fig. 4a. Unlike the models used in [5] and [8], the compliances of the tool system excluding the tool itself are represented by two springs. A simplified approach assuming a rigid body, has also been evaluated, see Fig. 4b.

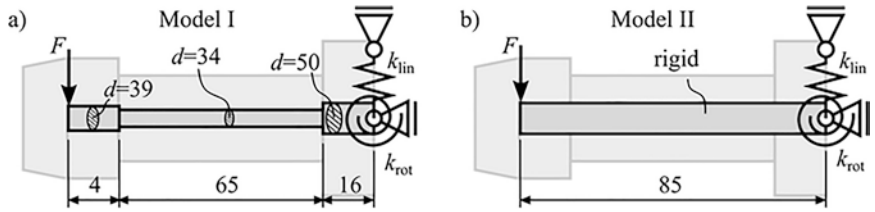


Fig. 4. Modelling of the tool-system deflections. Tool and its holder modelled as a) a beam with varying cross-sections A_i and a Young's modulus of 210 GPa; b) a rigid body.

A least squares fit has been performed for each of the models shown in Fig. 4 to determine the parameters k_{lin} and k_{rot} based on the data obtained for a constant relative axial position $z_f = 10$ mm of the tool with respect to the workpiece. Both models yield similar least squares residuals, and the resulting compliances along the tool axis are shown in Fig. 5a. For the incorporation of the compliance model in simulations, the main interest is the prediction of the deflection of the grinding wheel. While the incorporation of the tool deformation results in more realistic compliances along the tool axis, the predicted compliances at the grinding wheel $\bar{c}_t$ only deviate by approximately 2%.

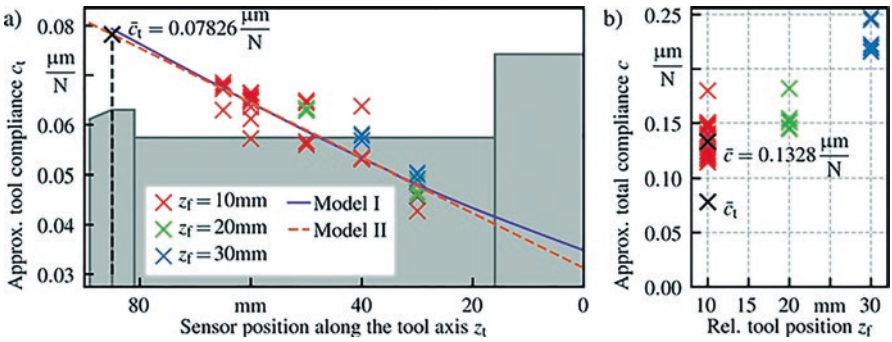


Fig. 5. Approximated compliances; a) Compliance along the tool axis. Measurements at different sensor positions z_t and resulting fit for models I and II. Since good accordance of both models at the point of interest can be observed, the tool compliance at the grinding wheel $\bar{c}_t$ is extrapolated from model II; b) Total compliances for different relative tool positions.

An approximation $\Delta a_{e,t}^{*}$ for the reduction of total depth of cut $\Delta a_{e,t}$ due to elastic deformations of the tool system follows as

$$\Delta a_{e,t}(F_n) \approx \Delta a_{e,t}^{*}(F_n) = \bar{c}_t F_n, \tag{1}$$

taking into account all static compliances of the tool system.

A generalisation of the model including compliances of the workpiece and clamping device is required for the prediction of process results. Once contact is initiated in the static test, the derivative of Δd_r with respect to F_n is assumed to be identical to the compliance c of the entire machining system. If nonlinearities in $\Delta d_r(F_n)$ are neglected, c can be determined in analogy to the tool compliance c_t from the linear fit $\Delta d_{r,lin}$ of $\Delta d_r(F_n)$, see Fig. 3d. Values obtained for c at three different tool positions z_f with respect to the workpiece are shown in Fig. 5b. Assuming that c includes all relevant compliances, the reduction of total depth of cut Δa_e in ITG has been approximated as

$$\Delta a_e(F_n) \approx \Delta a_e^{*}(F_n) = \bar{c} F_n, \tag{2}$$

where the machining system compliance $\bar{c}$ is obtained by averaging all measurements for a given position z_f. A minor influence of z_f on the tool compliance $\bar{c}_t$ can be observed, see Fig. 5a. This effect could be related to the spindle moving out of the headstock for increasing values of z_f and corresponding changes in lever arms. However, an unexpected increase of $\bar{c}$ by about 100% is found for $z_f = 30$ mm compared to $z_f = 10$ mm, see Fig. 5b. This deviation possibly results from small positions changes and elastic deformations of the thin-walled workpiece in the clamping system. Further research regarding the dependency of the compliances on the tool position is necessary to understand and model the influence of the tool position z_f. The following section is based on the values obtained at $z_f = 10$ mm due to the comparatively large data base.

3 Influence of System Compliances on Shape Deviations

ITG experiments have been carried out to investigate the influence of system compliance on shape deviations. The axial feed rate v_{fa}, and thus the process forces and the corresponding deflections, have been varied, resulting in changes of the inner profiles of the machined workpieces. Using a geometric-kinematic grinding simulation (GKS) in combination with the machining system compliance model (cf. Sect. 2.3) and the measured process forces, the resulting workpiece profiles have been calculated.

The test setup used for the grinding experiments is identical to the one described in Sect. 2.1, except no eddy current sensors are utilized. Bearing components have been machined under the experimental conditions listed in Table 1. Each test has been repeated once. After grinding, the inner workpiece profiles along the bore length have been measured at three rotational positions for comparison with the simulated axial profiles, using the form measuring system *Hommel Etamic F455* by *Jenoptik AG*.

Table 1. Experimental conditions and process parameters.

Grinding tool:	Electroplated cBN, B181, $D = 39$ mm, $\chi = 3°$
Workpiece:	100Cr6, hardened to 63 ± 1 HRC Cylindrical: $d = 47$ mm, $D = 53$ mm, $w = 31.4$ mm
Cooling medium:	Grinding oil, $\nu = 5$ mm^2/s
Grinding tool velocity:	$v_t = 80$ m/s
Workpiece velocity:	$v_w = 1.33$ m/s
Axial feed rate:	$v_{fa} = 120, 240, 360, 480, 600$ mm/min
Total radial stock removal:	$a_{e,tot} = 0.125$ mm

3.1 Experimentally Determined Process Normal Forces

The normal force F_n has been measured at the clamping device (cf. Sect. 2.1). Since the measured force is a superposition of clamping and process normal force, separation of these components is required. Additional superposed forces result from the rotational movement of the workpiece and possibly from thermal effects [6, 9]. Also, high-frequency oscillations are induced by the grinding process itself. The normal forces have been separated from the aforementioned factors in two steps: First, a finite-impulse-response filter design based on the Kaiser window, set to a low-pass frequency of $v_{fa}/4$ mm together with a transition width of 2 Hz, has been applied. Subsequently, an offset correction has been performed using constant offsets at the start and at the end of the tool-workpiece engagement with linear interpolation in between. In Fig. 6a, the processed normal force data and the maximum value for one exemplary measurement are shown. Figure 6b displays the maximum force of all measurements plotted against the axial feed rate and exhibits a linear dependency.

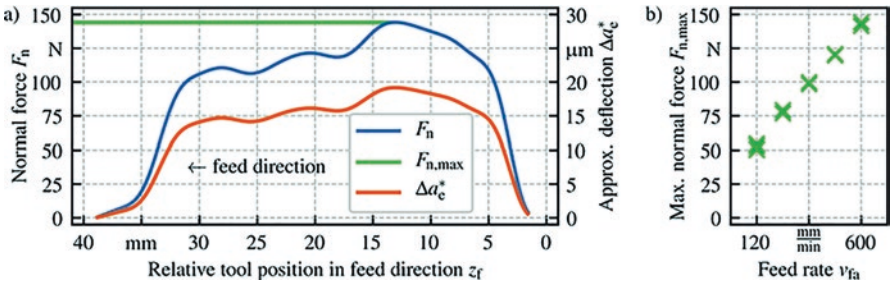


Fig. 6. a) Processed normal forces F_n and approximated machining system deflection Δa_e^* for an exemplary measurement; b) Maximum normal force $F_{n,max}$ for different axial feed rates v_{fa}.

3.2 Simulative Estimation of Shape Deviations

A GKS [14] has been adapted to simulate the resulting workpiece profile based on the compliance model for the entire system, presented in Eq. (2). Since this simulation system does not predict the process forces, the processed normal forces resulting from the experiments $F_n(z_f)$ are incorporated, see Sect. 3.1. The reduction in total depth of cut follows as

$$\Delta a_{\mathrm{e}}^{*}(z_{\mathrm{f}}) = \Delta a_{\mathrm{e}}^{*}(F_{\mathrm{n}}(z_{\mathrm{f}})) = \bar{c} F_{\mathrm{n}}(z_{\mathrm{f}}). \tag{3}$$

Since the compliance model does not take the variable z_{f} into account, see Sect. 2.3, the reduction in depth of cut depends on the position of the tool in relation to the workpiece z_{f} solely through the measured forces. The function $\Delta a_{\mathrm{e}}^{*}(z_{\mathrm{f}})$, is exemplarily shown in Fig. 6a next to the corresponding force measurement. By using $\Delta a_{\mathrm{e}}^{*}(z_{\mathrm{f}})$ as a radial offset of the tool model in the GKS, the achieved radial stock removal has been calculated, referred to as radial stock removed $a_{\mathrm{e,real}}(z_{\mathrm{w}})$. A comparison of the simulation results and the corresponding experiments is depicted for three representative specimens in Fig. 7. The curves are vertically aligned based on the assumption that the total radial stock removal $a_{\mathrm{e,tot}}$ is completely achieved at the end of the exit zone.

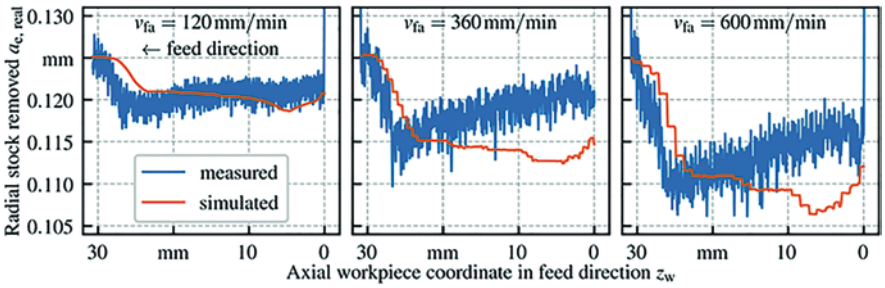


Fig. 7. Achieved radial stock removal along the bore length at three axial feed rates.

Higher axial feed rates v_{fa} and resulting normal forces F_{n} correspond to a reduced diameter along the workpiece profile in both experiments and simulations, and a widening of feed marks visible especially in the simulation results. The material removal is underestimated by the simulation in the entry zone, but overestimated in the exit zone. Due to the gradual decrease of the measured forces and thus the deflections, the errors in diameter exhibited by the main section tend to zero towards the end of the exit zone. This applies to both simulations and measurements. In summary, the influence of the feed rate on the characteristic shape deviations at the exit zone is reproduced using the simulation in good accordance with the measurements. The remaining discrepancies could result from factors not modelled, such as the varying compliance for different tool positions and workpiece deformations due to clamping and thermal expansion.

4 Conclusion and Outlook

Compliance models for the tool system (Eq. (1)) as well as for the entire machining system (Eq. (2)) have been evaluated. Both models have been calibrated using deflection and force data obtained from static analogy tests. For the tool system, two model variations have been compared. The tool has been modelled either as an elastic beam or as rigid body. Both variations result in nearly identical approximations of the

compliance at the grinding wheel for the system under consideration. The model for the entire machining system is based only on the normal force and the radial feed. Major differences in the compliance have been observed at different axial positions of the tool. Workpiece profiles after ITG have been measured and reproduced using a geometric-kinematic grinding simulation based on the machining system compliance model. Good accordance in the resulting profiles has been achieved for different axial feed rates, especially with respect to the characteristic shape deviation in the exit zone.

Additional investigations are planned to verify the underlying assumptions of the machining system compliance model and to include the influence of the axial tool position. In the future, an enhanced compliance model will be incorporated into a thermomechanically coupled simulation framework, similar to [9], to calculate the process forces and deflections monolithically. This will lead to higher quality simulation results, which could be used to develop highly efficient compensation strategies for ITG processes, with the potential to be integrated into CNC systems.

Acknowledgements. Gefördert durch die Deutsche Forschungsgemeinschaft (DFG) - 403857741
Funded by German Research Foundation (DFG) - 403857741

The authors thank the German Research Foundation (DFG) as well as the companies *Schaeffler Technologies AG & Co. KG* and *August Rüggeberg GmbH & Co. KG*.

References

1. Hashimoto, F., Yamaguchi, H., Krajnik, P., Chaudhari, R., Hoffmeister, H.-W.: Abrasive fine-finishing technology. CIRP Ann. **65**(2), 597–620 (2016)
2. Schumann, S.: Mehrskalige Modellierung und Simulation des Hochleistungs-Innenrundschälschleifens. Dissertation, Technische Universität Dortmund. Vulkan, Essen (2019)
3. Klocke, F.: Manufacturing Processes 2: Grinding, Honing, Lapping. Springer, Berlin (2009)
4. Marschalkowski, K., Biermann, D., Weinert, K.: On the characteristics of high-performance internal traverse grinding using electroplated CBN wheels. In: Aoyama, T., Takeuchi, Y. (Eds.), Proceedings of the 4th CIRP International Conference on High Performance Cutting (CIRP HPC 2010), vol. 1, pp. 393–398 (2010)
5. Gao, S., Yang, C., Xu, J., Fu, Y., Su, H., Ding, W.: Optimization for internal traverse grinding of valves based on wheel deflection. Int. J. Adv. Manuf. Technol. **92**, 1105–1112 (2017)
6. Biermann, D., Holtermann, R., Menzel, A., Schumann, S.: Modelling and simulation of thermal effects in internal traverse grinding of hardened bearing steel. CIRP Ann. **65**(1), 321–324 (2016)
7. Okuyama, S., Nishihara, T., Kawamura, S., Hamasaki, S.: Study on the geometrical accuracy in surface grinding – Thermal deformation of the workpiece in traverse grinding. J. Jpn. Soc. Precis. Eng. **4**(28), 13–24 (1994)
8. Pereverzev, P.P., Popova, A.V., Pimenov, D. Yu.: Relation between the cutting force in internal grinding and the elastic deformation of the technological system. Russ. Eng. Res. **35**(3), 215–217 (2015)

9. Holtermann, R., Schumann, S., Menzel, A., Biermann, D.: Modelling and simulation of internal traverse grinding – From micro-thermo-mechanical mechanisms to process models. In: Biermann, D., Hollmann, F. (Eds.) Thermal Effects in Complex Machining Processes. Lecture Notes in Production Engineering, pp. 369–403. Springer, Cham (2018)
10. Tawakoli, T., Rasifard, A., Rabiey, M.: High-efficiency internal cylindrical grinding with a new kinematic. Int. J. Mach. Tools Manuf. **47**, 729–733 (2007)
11. Abele, E., Altintas, Y., Brecher, C.: Machine tool spindle units. CIRP Ann. **59**, 781–802 (2010)
12. Ramos, J.C., Vinolas, J., Nieto, F.J.: A simplified methodology to determine the cutting stiffness and the contact stiffness in the plunge grinding process. Int. J. Mach. Tools Manuf. **41**, 33–49 (2001)
13. McKinney, W.: Data structures for statistical computing in python. In: Proceedings of the 9th Python in Science Conference, vol. 445, pp. 51–56 (2010)
14. Tsagkir Dereli, T., Biermann, D., Menzel, A., Schmidt, N., Furlan, T., Holtermann, R.: Prozessoptimierung für das Innenrundschälschleifen. VDI-Z **3**, 25–27 (2020)

Numerical Modelling of the Aeroacoustic and Flow Behaviour of Chip Fans

C. Menze(✉), C. Zizelmann, M. Schneider, K. Güzel, and H.-C. Möhring

Institute for Machine Tools (IfW), University of Stuttgart, Holzgartenstr. 17, 70174 Stuttgart, Germany
christian.menze@ifw.uni-stuttgart.de

Abstract. During the machining of wood and fibre composites, large quantities of fine dust and chips are produced which are harmful to health. In addition, a chip accumulation can occur at the cutting position, which reduces the production quality and promotes double cutting. In woodworking, chip fans are often used to remove these dusts and chips. These are located directly on the tool holder and generate an air flow directed towards the spindle by the spindle rotation. The additional airflow improves the removal of the dust and chips directly after the chip formation in the direction of the suction hood. As a side product, the chip fans generate high aeroacoustic emissions with increasing rotational speeds. In this paper, the acoustic behaviour of chip fans is characterised by measuring the sound pressure levels at different speeds when running idle. In addition, the aeroacoustic and flow behaviour of chip turbines is numerically modelled with a Computational Fluid Dynamic (CFD) simulation. Thus, in the future, design measures can be analysed for their sound-reducing effectiveness in time-saving simulation studies.

Keywords: Acoustic emission · Simulation · Design optimization

1 Introduction

During the processing of wood and wood-based materials, large quantities of fine chips and dust are produced. These can be harmful to human health and extend the setup times by manual rework to remove the chips. A new type of effective chip extraction can be achieved with chip fans [1]. These are mounted on the tool interface in the immediate vicinity of the Tool Center Point (TCP) and rotate at the spindle speed. This causes a local speed increase at the TCP and the chips are sucked in by the chip fan and ejected in the direction of the suction hood. In Fig. 1 different chip fans are shown. These differ considerably in their design. Chip fan 1 has six flow channels with a diameter a = 95 mm, a height c = 39 mm and the dimension from the lower edge of the chip fan to the surface of the workpiece e is ideally between e = 2–6 mm. Chip fan 2 has eight flow channels with a diameter of a = 98 mm, a

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 315–323, 2021.
https://doi.org/10.1007/978-3-662-62138-7_32

height of c = 36 mm and e = 3–6 mm. The dimensions of chip fan 3 are ten flow channels with a = 99 mm, c = 43 mm and e = 3–6 mm.

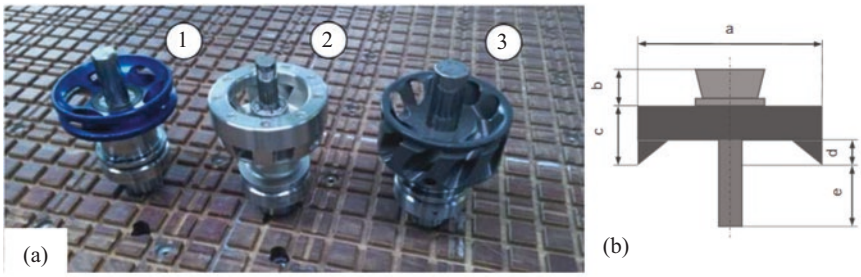


Fig. 1. Different chip fans (a): chip fan 1, chip fan 2, chip fan 3; Geometrical dimensions (b).

As a side effect, chip fans generate considerable acoustic emissions. A preliminary investigation according to the experimental set-up in Fig. 2 shows that the sound pressure level of an idling double-edged end milling tool at 20.000 rpm is 80.57 dB. When machining a wood-based material MDF, with a spindle speed of 20.000 rpm and a feed rate of 6 ms^{-1} the sound pressure level is 96.4 dB. However, chip fan 3 already reaches a sound pressure level of 104.1 dB in idle mode with 20,000 rpm, which already exceeds the legal limits for hearing damage. The comparison in Table 1 illustrates that chip fan 3 is a significant source of noise when idling and exceeds the noise of the tool during idling and machining considerably. For this reason, the following research activities relate to the investigation of the idling chip fan 3.

Table 1. Measured sound pressure levels of an idling tool, a tool in machining process and an idling chip fan 3.

	Idling double-edged milling tool at 20.000 rpm	Machining MDF with a doubled-edged milling tool	Idling chip fan 3
Sound pressure level [dB]	80.57	96.4	104.1

Acoustic emissions that occur at idling speed are often due to fluid mechanical phenomena [2, 3]. For fans and tools, a dominant characteristic blade-passing-frequency is formed as a function of speed and number of blades. Furthermore, turbulences lead to aeropulsive impacts resulting in considerable noise. Groß shows in [4] that turbulent flows in the immediate vicinity of rotating tools cause microscale turbulences and air vibrations. The pressure field generated by these turbulent flows results in a characteristic aeropulsive rotational noise. This effect was also investigated for idle running of circular saw blades [5].

Lighthill [6, 7] has shown how the problem of aerodynamic sound can be posed as an acoustic analogy in which the turbulence provides a quadrupole distribution in an ideal atmosphere at rest. He described the general properties of the induced field and developed the dominant effect of steady low-speed solenoidal source convection.

Ffowcs-Williams and Hawkings [8] extended the analogy to include solids in the calculation volume. This allows the consideration of monopole and dipole sources.

Möhring et al. [9] use the acoustic analogy to calculate the aeroacoustic behavior of circular saw blades in idle mode. Using numerical computational fluid dynamic (CFD)-simulation, fluid mechanical quantities are determined which are transformed into a sound pressure level by the acoustic analogy according to Lighthill and the extension of Ffowcs-Williams/Hawkings. With this computational aeroacoustic (CAA) simulation, the chip space and the tooth shape of the circular saw blades are optimized with regard to low aeroacoustic emissions [10].

This paper presents a transfer of the method presented in [9] to chip fans. At first the problem is presented by experimental investigations. Then a CAA simulation of the chip fans in idling operation is performed. Finally, the results are compared and discussed.

2 Experimental Investigation

To determine the sound pressure level of a chip fan, a commercially available version (chip fan 3, Fig. 1) was clamped into the main spindle of a MAKA PE 170 machine at the Institute for Machine Tools (IfW). To avoid the influence of aeroacoustic effects caused by a tool, a cylindrical shank dummy is clamped in the tool holder of the chip fan instead of an end milling tool. To measure the sound pressure level, a microphone is placed at a measuring distance of one meter from the chip fan (Fig. 2). The sound pressure level is evaluated with a Dactron FOCUS signal analyzer. The blade passing frequency sound is formed as a monopole, turbulences acting on the rigid blade surface can be considered as dipoles and the sound radiation by free turbulences as quadrupoles. These different noise sources are generated at each blade of the chip fan. Since the chip fan is rotationally symmetrical, no directionality of the sound generation is assumed and the sound pressure level is measured in only one position. Future investigations will consider a possible directional characteristic of the sound radiation of the chip fan in an acoustic laboratory.

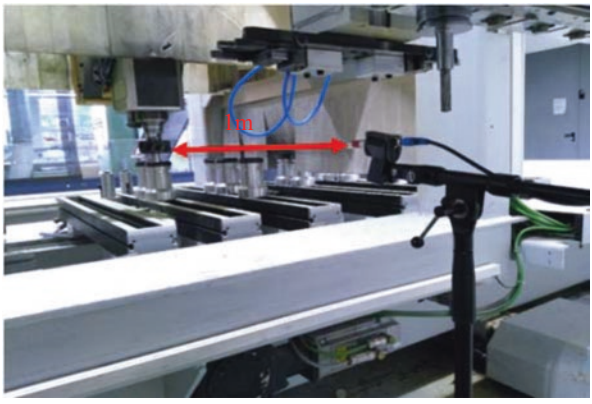


Fig. 2. Measurement of the sound pressure level.

The measurement was performed for characteristic operating rotational speeds of chip fans from 12,000 to 24,000 rpm. Table 2 shows the corresponding sound pressure levels. A spectral analysis of the measurement shows the dominant noise source at the 10th order of the rotational speed. This is the blade passing frequency of the chip fan, which has 10 blades in an equidistant arrangement.

Table 2. Measured sound pressure levels for different rotational speeds.

Rotational speed [rpm]	12,000	14,000	16,000	18,000	20,000	22,000	24,000
Sound pressure level [dB]	89.6	94.4	98.0	101.4	104.1	106.4	109.0

Prior to a simulation of the chip fan, the boundary conditions and basic sound generation mechanisms must be investigated. For this purpose, it is necessary to identify the dominant sound source. Subsequently, the sound generation mechanisms can be further investigated. In the literature, theoretical and experimental research [11] shows that in subsonic flow velocities, fluid displacement and turbulence in free flow contribute comparatively less to the total sound. The forces on the surfaces of the rotor and other stationary parts around which the flow passes are much more decisive.

It is therefore of interest whether the noise is generated by the chip fan or by fluidic effects which occur when the air transported by the chip fan hits or flows around surrounding components. For this purpose, a sound localisation is carried out with an acoustic camera. This limits the location of the sound generation and allows an analysis and identification of possible sound generation mechanisms. Figure 3 shows a recording with the acoustic camera of the chip fan at a rotational speed of 18,000 rpm in idle mode. The analysed frequency range is from 0 to 48 kHz. The experiment shows that the chip fan is the dominant sound source, more precisely the flow outlet at the end of the flow channel. This indicates the existence of blade passing frequency sound as well as vortex noise caused by the passing blades. It also shows that there is no other noise source, which is less than 15 dB quieter than the chip fan. Hence, there are no other dominant sources or reflections in the system, which contribute significantly to the overall sound pressure level. The air abruptly stops after leaving the flow channel and creates turbulent eddies in the vicinity of the chip fan. A flow-mechanical simulation (Fig. 3) confirms this assumption. Therefore, it can be excluded that the noise is caused by effects in the environment, such as turbulence caused by the impact of the flow on surrounding components. A successful noise reduction must be implemented at the chip turbine by constructive design modifications.

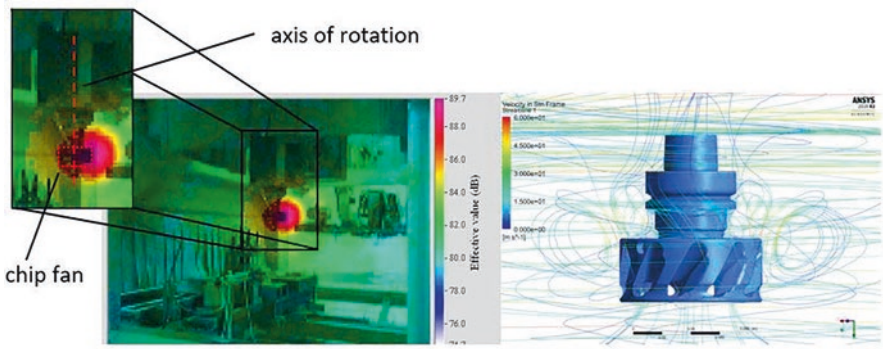


Fig. 3. Experimental sound localisation with the acoustic camera (left); CFD simulation of the chip fan 3 with turbulences at the outlet of the flow channels (right).

3 CFD Simulation

For the simulation of the flow characteristics of the chip fans, a CFD simulation is carried out in ANSYS-CFX. Here, not the chip fan itself is simulated, but the surrounding air in which the chip fan normally rotates. The simulation therefore consists of a cylindrical solid body which represents the air in the environment of the chip fan. The chip fan is cut out of this cylinder. This leaves the entire air space around the chip fan. The size of the calculation room is iteratively selected so that the formation of vortices is not disturbed, but the calculation time can be kept within limits. The diameter of the calculation room is 250 mm and the height is 350 mm. For the mesh tetrahedron elements with a maximum skewness of 0.9 and a minimum orthogonal quality of 0.3 are used. The inner boundary of the air to the chip fan and the outer boundary of the simulation space are defined as frictional walls. Finally, the air is rotated around the rotation axis at the rotational speeds from the experiment (Fig. 4).

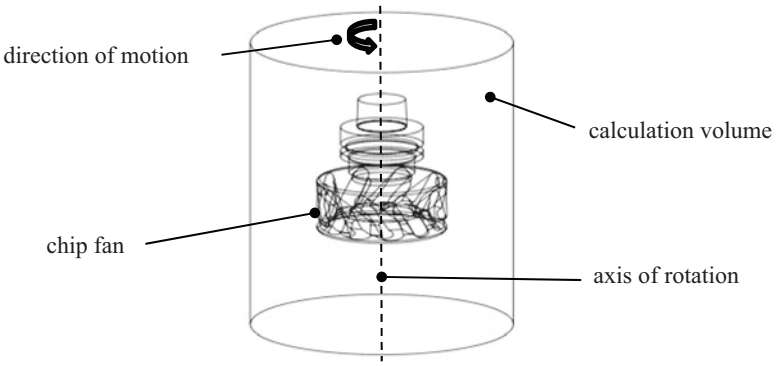


Fig. 4. CAD model of the calculation volume.

Additionally, an analysis of the different turbulence models was carried out to determine their suitability for the presented issue. For this purpose, the common models k-ε, Shear Stress Transport (SST) as well as the Reynold Stress models BSL and SSG were used and simulated for all rotational speeds.

The experimental investigations showed that the dominant noise source is at the outlet of the flow channels of the chip fan. This indicates that turbulent aerodynamic phenomena must be present at this location. The analysis of the CFD simulation confirms this hypothesis. Figure 5 shows the flow with velocity vectors in the stationary frame. It becomes clear that the air at the outlet of the flow channel is highly compressed. This exiting air jet pulls further air by shear friction effects and decays into microscale turbulences. These aerodynamic effects result in considerable noise emissions.

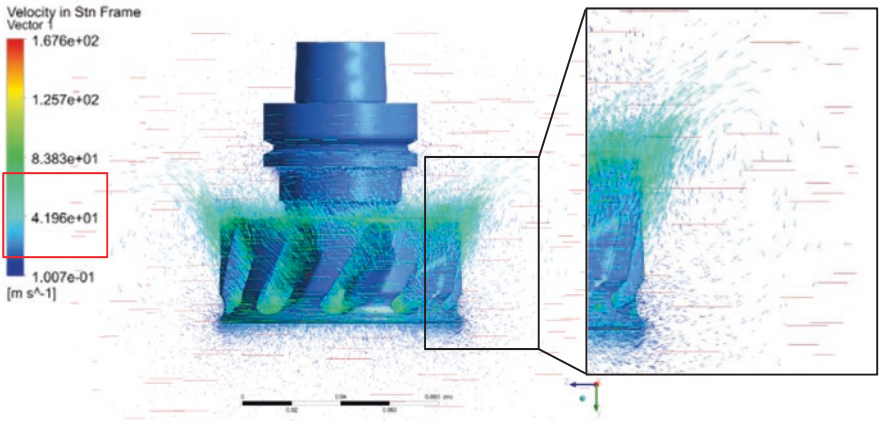


Fig. 5. Velocity in stationary frame (Vector) with turbulent eddies (BSL-Model)

4 Acoustic Analogy

To convert the results of the flow simulation, the acoustic analogy of Lighthill and Ffowcs-Williams/Hawkings is used [6–8].

Lighthill has shown, that the basic equations of fluid mechanics can be converted into a wave equation, which describes the propagation of sound. Therefore, it is possible to calculate sound parameters from results, which are obtained from a flow simulation.

The acoustic analogy can be seen in Eq. (1):

$$\begin{aligned}\frac{\partial^2\rho}{\partial t^2} - c_0^2\frac{\partial^2\rho}{\partial x_i^2} &= \frac{\partial\dot{m}}{\partial t} - \frac{\partial}{\partial x_i}(\rho f_i + \dot{m}v_i) + \frac{\partial^2}{\partial x_i\partial x_j}\left(\rho v_i v_j + p_{ij} - c_0^2\rho\delta_{ij}\right) \\ &= \frac{\partial\dot{m}}{\partial t} - \frac{\partial}{\partial x_i}(\rho f_i + \dot{m}v_i) + \frac{\partial^2}{\partial x_i\partial x_j}T_{ij}\end{aligned} \tag{1}$$

To solve the above equation, the generalized Kirchhoff equation is used. Equation (2) shows, how the sound pressure in a receptor point *p(x,t)* is calculated. The integral *dV* stands for the air volume of the flow simulation, *dS* is the surface of the chip fan.

$$p(x,t) = \int_V \frac{1}{4\pi r}\left(\frac{\partial \dot{m}}{\partial t}\right)_\tau dV - \int_S \frac{1}{4\pi r}\left(\frac{\partial(\rho v_i)}{\partial t}\right)_\tau n_i dS - \frac{\partial}{\partial x_i}\int_V \frac{1}{4\pi r}(f_i + \dot{m}v_i)_\tau dV$$
$$+ \frac{\partial}{\partial x_i}\int_S \frac{1}{4\pi r}\left(\rho v_i v_j + p_{ij}\right)_\tau n_j dS + \frac{\partial^2}{\partial x_i \partial x_j}\int_V \frac{1}{4\pi r}\left(T_{ij}\right)_\tau dV \tag{2}$$

In order to obtain the sound pressure level from the sound pressure, Eq. (3) has to be applied.

$$L_p = 20 \cdot \log_{10}\left(\frac{p(x,t)}{2 \cdot 10^{-5} Pa}\right) dB \tag{3}$$

A more detailed explanation of the calculation steps is displayed in [9]. There, it is shown, how the results of the flow simulation are used to solve the integrals.

5 Results

Figure 6 shows the results of the calculated sound pressure levels gained from the flow simulations. A comparison between the used turbulence models and the measurement is displayed.

There are deviations between the measurement and the calculated results. Almost all results are below the experimental data. Only the SSG-model shows a higher result at 12.000 rpm. The diagram demonstrates that the results of the calculations are highly dependent on the used turbulence model.

The fact, that almost all results are lower in sound pressure level than the measured data is due to the measurement setup. Even though there are no other significant sound sources in the machining center, the setup can still cause an error. Therefore, future measurements will be conducted in an acoustic laboratory to eliminate any potential errors. For the acoustic optimisation, only the relative change of the sound pressure level depending on the rotational speed is relevant. Therefore, the slopes of the balance lines are looked at for comparison.

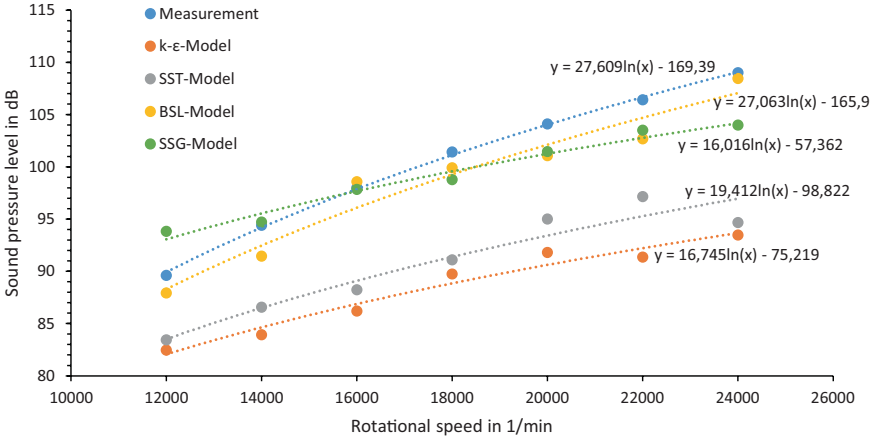


Fig. 6. Comparison between the used turbulence models and the measurement regarding the sound pressure level.

The BSL-model shows almost the same slope than the measurement. The displayed regression functions in Fig. 6 clearly illustrate that fact. The differences between the sound pressure levels are very small. The BSL-model is also optimized for rotating systems, which results in a good convergence behavior of the simulation. Hence, this turbulence model will be used for ongoing investigations and the optimisation of the chip fan.

6 Summery and Future Work

With the performed simulations, it is now possible to predict the acoustic behavior of the chip fan with a sufficient accuracy. The trend of the dependence of the sound pressure level on the rotational speed was shown for the used chip fan. Thus, the foundation for the optimisation of the chip fan geometry in regard to the noise emission is built.

As part of ongoing work, an investigation is being carried out, where the influence of the geometrical properties of chip fans on the acoustic behavior is being regarded. These parameters have yet to be determined before the study can be carried out. The results will be the basis for any future optimisation.

In addition, the flow conditions must be taken into account, since the chip fan must still be able to transport chips and dust to the suction system.

As well as the acoustic behaviour of the chip fan at idle, future simulations should also consider the chips in a two-phase flow.

Furthermore, flow simulations with other chip fans will be done to show the general applicability of the method and to develop a guidance line for designing low noise chip fans.

References

1. Möhring, H.C., Eschelbacher, S., Güzel, K., Kimmelmann, M., Schneider, M., Zizelmann, C., Häusler, A., Menze, C.: En route to intelligent wood machining – current situation and future perspectives. J. Mach. Eng. **19**(4), 5–26 (2019)
2. Kuolt, H.: Geräuschentstehung und Maßnahmen zur Lärmminderung an Holzbearbeitungsmaschinen. Dissertation, Universität Stuttgart (2006)
3. Stehle, T., Groß, L.: International market analysis on the noise emission of woodworking machines. In: First European Forum on Effective Solutions for Managing Occupational Noise Risks, Noiseatwork 2007, Lille (2007)
4. Groß, L., Heisel, U.: Analysis of the flow field of rotating milling cutters in idle mode. In: NAG/DAGA, International Conference on Acoustics, Rotterdam, vol. II, pp. 729–730, Deutsche Gesellschaft für Akustik e. V. (DEGA), Berlin (2009)
5. Stehle, T., Heisel, U., Birenbaum, C.: Simulative Untersuchung von Kreissägeblättern. Wt Werkstatsttech. Online **103**(1/2), 69–75 (2013)
6. Lighthill, M.J.: On sound generated aerodynamically. I. General theory. R. Soc. Pub. Lond. (A) **222**, 564–587 (1952a)
7. Lighthill, M.J.: On sound generated aerodynamically. II. Turbulences as a source of sound. R. Soc. Pub. Lond. (A) **222**, 1–32 (1952b)
8. Ffowcs-Williams, J.E., Hawkings, D.L.: Sound generated by turbulence and surfaces in arbitrary motion. R. Soc. Pub. Lond. (A) **264**, 321–342 (1969)
9. Möhring, H.C., Stehle, T., Güzel, K., Zizelmann, C.: Numerical flow simulation of rotating circular saw blades for the investigation of sound generation mechanisms. J. Mach. Eng. **18**(1), 61–71 (2018)
10. Güzel, K., Stehle, T., Möhring, H.C.: Simulationsgestützte Optimierung des aeroakustischen Verhaltens von Kreissägeblättern. Werkstattstech. Online **1**(2), 18–23 (2020)
11. Carolus, T.: Ventilatoren – Aerodynamischer Entwurf, Schallvorhersage, Konstruktion, 3., überarbeitete und erweiterte edn. Springer Vieweg, Wiesbaden (2013)

Estimation of the Influence of Volumetric Correction Approaches on the Thermo-Elastic Correction Accuracy

X. Thiem[1(✉)], B. Kauschinger[1], J. Müller[1], and S. Ihlenfeldt[1,2]

[1] Professur für Werkzeugmaschinenentwicklung und adaptive Steuerungen, Institut für Mechatronischen Maschinenbau, TU Dresden, Kutzbach-Bau Helmholtzstraße 7a, 01069 Dresden, Germany
xaver.thiem1@tu-dresden.de

[2] Fraunhofer-Institut für Werkzeugmaschinen und Umformtechnik IWU, Nöthnitzer Straße 44, 01187 Dresden, Germany

Abstract. Thermo-elastic errors in machine tools have a significant influence on the machining accuracy. The error at the tool center point can be determined with the structure model based correction, which uses physical based models like finite element models. The output of the structure model is the volumetric thermo-elastic error in the workspace of the machine tool. The correction values of the machine axes are determined based on the volumetric error with the help of a kinematic model. Therefore, three variants are presented in this article and their influence on the correction accuracy is estimated. The evaluation is based on typical thermo-elastic errors in the workspace of machine tools. The influence of the different variants on the correction accuracy is estimated with the Monte Carlo method based on randomly generated errors in the workspace of the machine.

Keywords: Thermo-elastic error · Machine tool · Correction · Compensation · Volumetric error · Structure model · Accuracy · Evaluation

1 Introduction

Thermo-elastic errors dominate the machining errors of machine tools [1]. Up to 75% of the overall geometrical errors of the workpiece are caused by thermal effects [2]. According to [1] new advanced approaches for the reduction of thermo-elastic errors have to improve the accuracy without additional energy consumption and utilize models for the thermal behavior. Model based correction approaches can be divided into structure model based (e.g. [3]), transfer function based (e.g. [4, 5]) and correlative model based correction (e.g. [6]). Structure models are physical based models like finite element models and can be seen as “white box” models. The characteristic thermal behavior of machine tools can be described with transfer functions (“grey box”), for example with first-order lag elements. Correlative models treat the machine as “black box” and usually describe the relation between measured temperatures at

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 324–333, 2021.
https://doi.org/10.1007/978-3-662-62138-7_33

different points in the machine and measured displacements at the tool center point (TCP) or deformations of the axes. This article focuses on the correction of the thermo-elastic volumetric error at the tool center point (TCP) calculated by a structure model. However, all model based correction approaches have to use control-integrated functions for the volumetric correction of the thermo-elastic errors. Therefore, the presented approach is also relevant for other model types. In this article, three implementation variants and the number of necessary grid points calculated by the modeol are evaluated with regard to their influence on the correction accuracy.

The article has the following structure. The geometric correction approach on which the thermo-elastic correction is based is briefly explained in Sect. 2. In Sect. 3, three possible variants for a combined thermo-elastic and geometric error correction are described. The typical thermo-elastic errors at the TCP in the workspace is analyzed and an approach for the calculation of random typical errors are presented in Sect. 4. Based on these randomly generated errors an approach is described to estimate the residual error at the TCP for the three different variants in Sect. 5. The approach was implemented for the kinematic of a demonstrator machine (Sect. 6) and the results are evaluated in Sect. 7. Finally, the article closes with a summary and an outlook (Sect. 8).

2 Geometric Volumetric Correction

The *thermo-elastic* volumetric correction is based on a *geometric* volumetric correction approach. The *geometric* volumetric correction uses axis error parameters according to ISO230-1 [7] and a kinematic model. The following notation is used for the error parameter, e.g. EYX (E-error, Y-error direction; X-examined machine axis). The kinematic model utilizes homogenous transformation matrices (HTM). It is assumed that the error of one axis is independent from the other axes positions and therefore that the components of the machine behave like rigid bodies [8].

The approach used for the correction of rotatory and translational errors is called decoupling method [9]. The flow chart in Fig. 1 (blocks on the lower half below dashed line) depict the principle of the correction for the six degrees of freedom. It starts with the measurement of the error parameters according to ISO 230-1 [7]. The error at the TCP (ΔX, ΔY, ΔZ, $\Delta\Theta_X$, $\Delta\Theta_Y$, $\Delta\Theta_Z$) is calculated based on this error parameters, the kinematic model and the current axes positions in the machine control. At first, the rotatory errors at the TCP are corrected for example by tilting the y- and the z-slide. This correction leads to additional translational errors at the TCP (ΔX_{rot}, ΔY_{rot}, ΔZ_{rot}). Therefore, the entire translational error is corrected in a second step by the axes correction values (e.g. X_k, $Y1_k$, $Y2_k$, $Z1_k$, $Z2_k$, $Z3_k$). Finally, the correction values are added to the current axes positioning values (e.g. X, Y1, Y2, Z1, Z2, Z3). On the example of a demonstrator machine (see Sect. 6) the geometric correction approach was verified and the deviations were reduced approximately by 90% [8].

3 Combined Volumetric Correction

The base for the thermo-elastic volumetric correction is a grid with the thermo-elastic errors at the TCP (ΔTCP) in the workspace of the machine. The grid can be provided by a structure model but also by other thermal models as described in the introduction. The three variants in combination with the geometric correction are depicted in Fig. 1. There are two different real time requirements for the functions (marked with different colored frames). Thermal real times (green) means that the results have to be calculated before relevant thermo-elastic deformations occur. The functions with real-time requirements in the control are marked with a red frame. These functions depend on the current axes positions. Therefore, they have to be recalculated ideally with every interpolation cycle of the control.

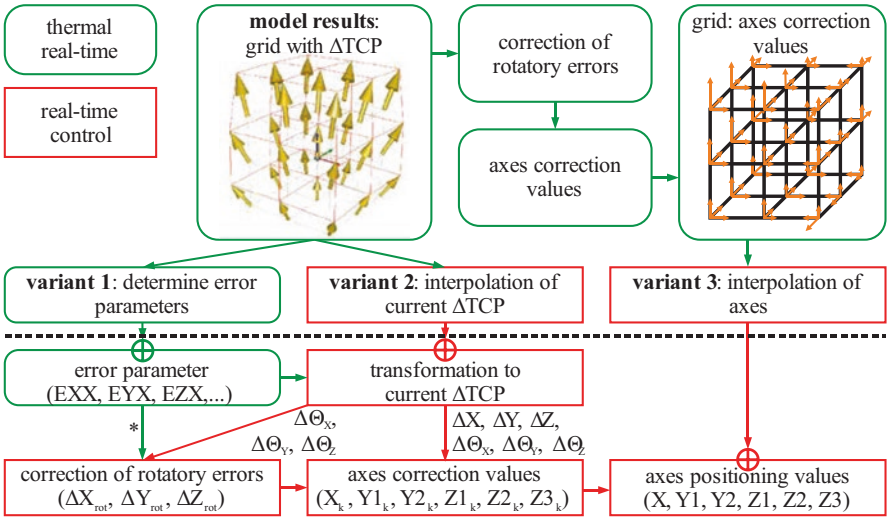


Fig. 1. Variants for the combination of thermo-elastic and geometric volumetric error correction

In the **first variant**, the thermo-elastic error parameters (ISO230-1 [7]) of the axes are determined and added to the geometric error parameters. The determination of the thermo-elastic error parameters is similar to the measurement of the geometric error parameters. The error at the TCP at equally distributed points along the axis direction are used as error parameters of the axis. This is done for each axis separately while the other axes are in zero position. Therefore, three lines with errors at the TCP in the workspace are taken into account. The reference point is the center of the workspace. The error at the TCP is distributed between the error parameter of the three axes at this point. The distribution is determined with the help of the kinematic model by minimizing the residual errors at the TCP at the given grid points. Therefore, least error squares and the interior-point method is used.

The **second variant** uses directly the given grid with the errors at the TCP in the workspace. Depending on the current axes positions the current error at the TCP is linearly interpolated within the grid. The resulting thermo-elastic error at the TCP is added to the geometric error. The difference to the first variant is that thermo-elastic error parameters (see * in Fig. 1) are not considered in the calculation of the translational errors (ΔX_{rot}, ΔY_{rot}, ΔZ_{rot}) caused by the *correction of the rotatory errors* at the TCP. The influence on the correction accuracy is considered to be very small since the thermo-elastic error parameters are typical in the range of micrometers. Therefore, they only have a very small influence on the kinematic transformation, which is used to determine the translational errors ΔX_{rot}, ΔY_{rot}, ΔZ_{rot}.

In the **third variant**, the axes correction values are determined at the given grid with errors at the TCP. Therefore, the functions "correction of rotatory errors" and "axes correction values" (see Fig. 1) are moved out of the control. The axes correction values are linearly interpolated depending on the current axes positions within the grid. Finally, the correction values are added to the axes positioning values. As in variant 2 the thermo-elastic error parameters can't be considered for the correction of the rotatory errors.

4 Typical Thermo-Elastic Error at TCP in Workspace

The base for the evaluation of the three different variants is a grid with errors at the TCP in the workspace. Therefore, machine tool typical thermo-elastic errors are generated. In this way, a wide variety of thermo-elastic errors at the TCP can be evaluated. The kinematic model of the machine and generated thermo-elastic error parameters according to ISO230-1 [7] are used for the calculation of the error at the TCP. The characteristics of the error parameters were analyzed with the help of published measurements of machine tools with a workspace size from 0.5 m × 0.4 m × 0.35 m up to 1 m × 1 m ×1 m [4, 5, 10–15]. The thermo-elastic error parameters depend on the axis position and the time. The typical thermo-elastic error parameters are approximately in the range of ±80 µm respectively ±80 µm/m. An assumption is that the error parameters change linearly over the length of the axis [5, 11, 13, 15]. The nonlinearities of positioning and angular errors are in the range of ±4 µm respectively ±4 µm/m in these publications. Since the linear part of the straightness errors is considered in the squareness errors, the remaining straightness error are nonlinear in the range up to ±10 µm [5, 10, 14, 15]. In the first step (only for positioning and angular error) of the parameter generation, a random straight line over the axis position is determined with a uniform distributed rise and offset. In a second step (for all position dependent errors), at five random points along the axis a uniform distributed random offset in the range of the typical nonlinearities are added to the straight line. Between these five points, the curve is interpolated by a shape preserving piecewise cubic Hermite interpolating polynomial based on [16]. This interpolation prevents an overshoot of the resulting curve and generates a realistic curve of the error parameter over the axis position ($E_{(x)}$). The squareness errors are determined by uniform distributed random values in the range of ±80 µm/m.

The time behavior of the thermo-elastic error parameters after a load step is often similar to a lag element of first order. This property of the error parameter is used for correction approaches based on transfer functions [4, 5, 17]. The time behavior is characterized by the thermal time constant (τ) of the machine or of machine components. The typical time constant is in the range of 0.3 to 6.8 h. In this range, a uniform distributed random time constant is generated. The position dependent error parameters $E_{(x)}$ are scaled by the step response of a lag element of first order with this time constant (Eq. 1). A load step is chosen because it leads to the largest possible rise of the error parameters.

$$E_{total(x,t)} = E_{(x)} \cdot \left(1 - exp\left(-\frac{t}{\tau}\right)\right) \tag{1}$$

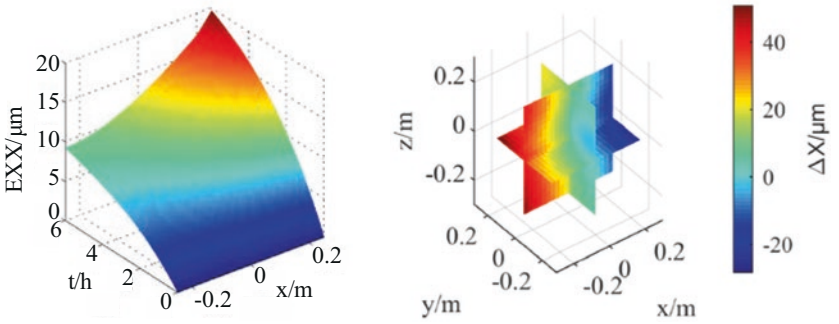


Fig. 2. Example for generated EXX (left) and error ΔX at TCP in workspace (right)

Figure 2 shows an example for a randomly generated positioning error of the x-axis (EXX) over six hours on the left side. The error at the TCP at a reference grid in the workspace is calculated based on these randomly generated error parameters. A kinematic model is used for this calculation. An example for the calculated error of the TCP in X direction (ΔX) is depicted in Fig. 2 on the right. The figure is for one point in time after 6 h. The errors are plotted at the X-Y-plane, the X-Z-plane and the Y-Z-plane to get an overview of the distribution in the workspace.

5 Approach for the Estimation

The following section describes the estimation of the residual error at the TCP, which is used to evaluate the influence of the three variants and the number of given grid points on the correction accuracy. The translational errors at the TCP are summarized with the Euclidean norm (2-norm, $TCP_{E,trans}$), which can be interpreted as length of the displacement vector. The rotational errors at the TCP are summarized with 1-norm ($TCP_{E,rot}$). The characteristic value k_{max} is the maximum residual error with respect to the maximum uncorrected error. The characteristic values are determined for different randomly generated typical thermo-elastic errors at the TCP and implementation variants. A

distribution for the characteristic values is calculated with the help of the Monte Carlo method. All *geometric* error parameters are set to zero for this investigation. Figure 3 shows the flow chart for the determination of the distribution. At first, a reference grid with typical errors at the TCP is generated for 6 h. The reference grid has eleven points in each direction, altogether 1331 points. A subset of points is selected from these points. This subset represent the output of the structure model. In practice, the number of grid points that can be calculated by the structure model is limited by real-time requirements. The subset is used for correction with the three variants (Fig. 1). In the next step, the residual errors at the TCP at the reference grid are determined and the characteristic values are calculated. At the points of the subset, the results of the variants are nearly identical to the reference. The relevant residual errors result from the variant specific interpolation at the reference points between the subset points (structure model output). The calculation of the characteristic values is done every 10 min of the 6 h. After this, the process starts over at the first step until 500 repetitions are reached.

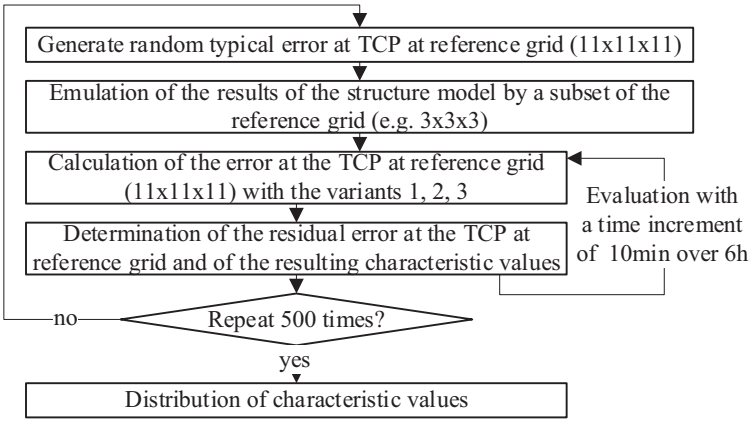


Fig. 3. Flow chart for the determination of the distribution of the characteristic values

6 Demonstrator Machine

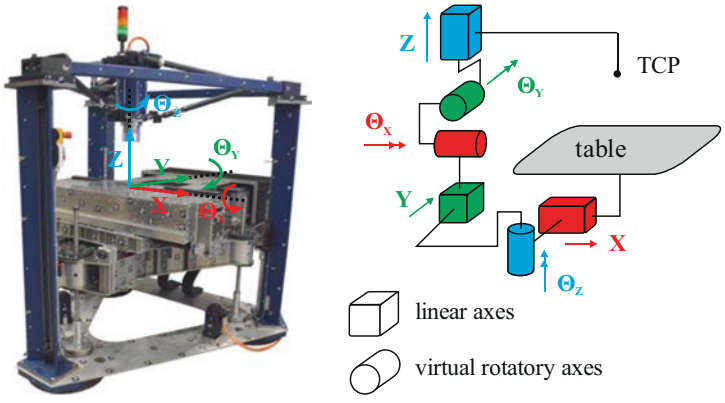


Fig. 4. Demonstrator machine and kinematic chain, based on [8]

As an example a machine with a 3-axis Cartesian kinematic is used. This machine has the special ability to compensate errors in six degrees of freedom. Figure 4 pictures the demonstrator machine on the left. Three ball screw axes drive the z-slide. Flexure bearings connect the ball screw nuts and the guide carriages to the z-slide. Therefore, the z-slide can be tilted around X- and Y-direction in small angles up to 10 mrad. Two parallel linear direct drives actuate the y-slide. Flexure bearings connect the y-slide to the guide carriages of the linear guideways at the z-slide. In this way, the y-slide can be tilted around the z-direction up to 1.6 mrad. The kinematic chain of the machine from the table to the tool center point is sketched in Fig. 4 on the right. Cuboids mark the translational movements. Cylinders mark the rotatory movements. The rotatory axes are named as virtual since they are no rotatory axes in the usual sense [8].

Eq. 2 shows the transformation with HTMs [18] (T) from the machine *table* coordinate system to the *TCP* coordinate system taking the axes positions (*Pos*), pivot points (*PP*), angles of virtual rotatory axes (Θ), squareness errors (*SE*) and axes errors (*E*) into account. The top left index denote the reference coordinate system and the bottom right index denote the target coordinate system. For the calculation of the error at the TCP, it is assumed that the virtual rotatory axes themselves have no error parameters.

$$
\begin{aligned}
{}^{table}T_{TCP} = {}& {}^{table}T_{XSE}\,{}^{XSE}T_{XPos}\,{}^{XPos}T_{XE}\,{}^{XE}T_{YPP}\,{}^{YPP}T_{\Theta Z}\,{}^{\Theta Z}T_{Y}\,{}^{Y}T_{YSE}\,{}^{YSE}T_{YPos}\,{}^{YPos}T_{YE}\\
& {}^{YE}T_{ZPP}\,{}^{ZPP}T_{\Theta X,\Theta Y}\,{}^{\Theta X,\Theta Y}T_{Z}\,{}^{Z}T_{ZSE}\,{}^{ZSE}T_{ZPos}\,{}^{ZPos}T_{ZE}\,{}^{ZE}T_{TCP}
\end{aligned}
\tag{2}
$$

The workspace of the demonstrator machine has the size 0.5 m × 0.4 m × 0.44 m. This is within the range of the workspace sizes of the machine tools analyzed in Sect. 4.

7 Results

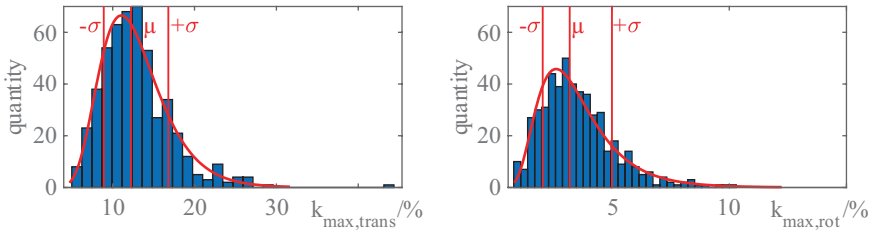


Fig. 5. Example for relative translational (left) and rotatory (right) residual error at TCP

In this section, the results of the Monte Carlo simulation are discussed for the demonstrator machine. In Fig. 5 are typical distributions for the relative residual translational ($k_{max,trans}$) and rotatory error ($k_{max,rot}$) at the TCP depicted. The results in Fig. 5 are from variant 2 using a subset of 125 points (5 points in each axis direction) of the reference grid. The other variants and different numbers of grid points lead to similar distributions. The histograms show distributions, which have a positive skew (right-skewed). They are similar to a logarithmic normal distribution. Logarithmic

normal distributions are suitable to describe multiplicative effects [19]. In this case, the random error parameters were multiplied along the kinematic chain. Therefore, a logarithmic normal distribution is assumed for $k_{max,trans}$ and $k_{max,rot}$. In the figures, the expected value (μ) and 1-sigma-limits (σ) are marked with red lines. Within the limits are 68.27% of the values.

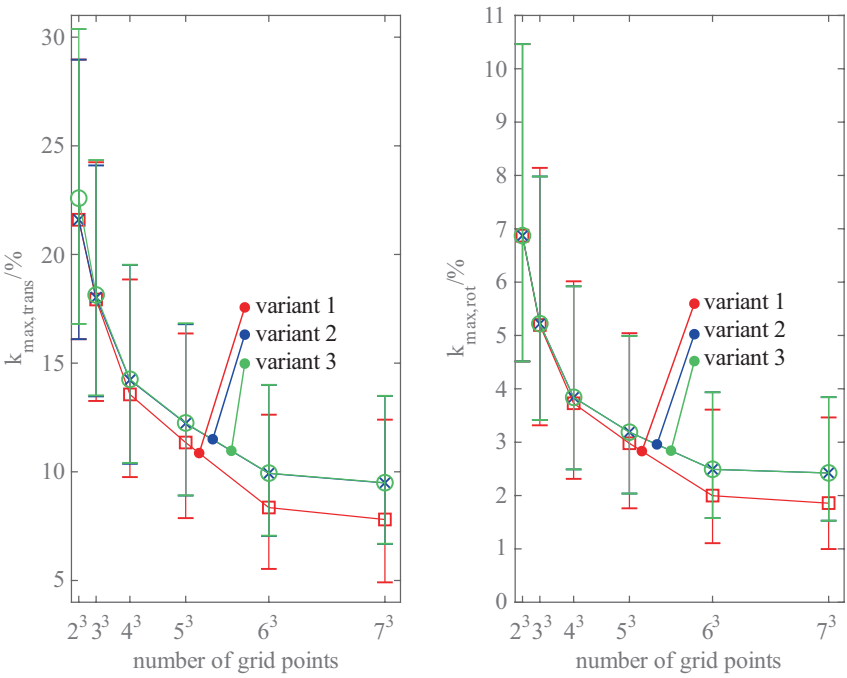


Fig. 6. Relative translational (left side) and rotatory (right side) residual error at the TCP

Figure 6 shows the expected values and the 1-sigma-limits (error bars) of k_{max} depending on the number of points in the subset of the reference grid. In each axis direction always the same number of grid points is used. The three variants only differ slightly in their influence on the correction accuracy, even for a very small number of grid points. For a high number of grid points variant 1 has a lower expected value but also a similar variance compared to variant 2 and 3. The relative translational residual error ($k_{max,trans}$) of variant 2 starts close to variant 1 for low numbers of gird points and approaches variant 3 for higher numbers. The relative rotational residual error ($k_{max,rot}$) of variant 2 and 3 are nearly equal.

The diagrams show, that the number of used grid points has a higher influence on the correction accuracy then the chosen variant. The expected values and the 1-sigma-limits of $k_{max,trans}$ and $k_{max,rot}$ decrease asymptotically with increasing number of grid points. There is still a relevant residual error, due to the nonlinearities of the error parameters and simplifications (e.g. small-angle approximation). In conclusion, the three variants are similar in terms of correction accuracy. The choice of the variant

for implementation depends on the specific control with its capabilities to integrate additional real-time functions. In this example, the structure model should calculate the error at the TCP at least at 27 grid points in the workspace of the machine in order to keep the 1-sigma-limit of the relative residual error of the volumetric correction below 25%.

8 Summary and Outlook

In this article, three variants were introduced for the combined geometrical and thermo-elastic volumetric error correction at machine tools. The variants base on a grid with errors at the TCP in the workspace. The grid represents the output of the structure model or another kind of model for thermo-elastic errors. In the next step, the typical thermo-elastic error parameters of machine tools were analyzed. Based on this analyzation typical thermo-elastic errors at the TCP were randomly generated as reference for the evaluation of the variants. The Monte Carlo method was used to estimate the influence of the different variants and of the number of the given grid points on the correction accuracy for a demonstrator machine. The variants only slightly differ in their influence. The number of given grid points has a higher impact on the correction accuracy than the variants. At least 27 grid points in the workspace should be used in order to keep the 1-sigma-limit of the relative residual error below 25%. In future works, the thermo-elastic displacement of the TCP in the workspace of the demonstrator machine will be measured with the help of a photogrammetric measurement system [20]. Based on the measurement the accuracy of the variants with different numbers of grid points could be evaluated and compared to the results presented in this article.

Acknowledgements. Funded by the German Research Foundation – Project-ID 174223256 – TRR 96

References

1. Wegener, K., Weikert, S., Mayr, J.: Age of compensation – challenge and chance for machine tool industry. Int. J. Autom. Technol. **10**, 609–623 (2016)
2. Mayr, J., Jedrzejewski, J., Uhlmann, E., Alkan, D.M., Knapp, W., Härtig, F., Wendt, K., Moriwaki, T., Shore, P., Schmitt, R., Brecher, C., Würz, T., Wegener, K.: Thermal issues in machine tools. CIRP Ann. Manuf. Technol. **61**, 771–791 (2012)
3. Liu, K., Li, T., Wang, Y., Sun, M., Wu, Y., Zhu, T.: Physically based modeling method for comprehensive thermally induced errors of CNC machining centers. Int. J. Adv. Manuf. Technol. **94**, 463–474 (2018)
4. Bitar-Nehme, E., Mayer, J.R.R.: Modelling and compensation of dominant thermally induced geometric errors using rotary axes' power consumption. CIRP Ann. (2018)
5. Wennemer, M.: Methode zur messtechnischen Analyse und Charakterisierung volumetrischer thermo-elastischer Verlagerungen von Werkzeugmaschinen. Apprimus Wissenschaftsverlag, Aachen (2018)

6. Naumann, C., Priber, U., Ihlenfeldt, S., Riedel, I.: Characteristic diagram based correction algorithms for the thermo-elastic deformation of machine tools. In: Proceedings 48th CIRP Conference on Manufacturing Systems, Naples (2015)
7. ISO 230-1: Test code for machine tools - Part 1: geometric accuracy of machines operating under no-load or quasi-static conditions (2012)
8. Ihlenfeldt S., Müller J., Merx M., Kraft M., Peukert C.: Simplified manufacturing of machine tools utilising mechatronic solutions on the example of the experimental machine MAX. In: Yan X.T., Bradley D., Russell D., Moore P. (Eds) Reinventing Mechatronics: Developing Future Directions for Mechatronics. Springer International Publishing (2020)
9. Hsu, Y.Y., Wang, S.S.: A new compensation method for geometry errors of five-axis machine tools. Int. J. Mach. Tools Manuf. **47**, 352–360 (2007)
10. Brecher, C., Fey, M., Wennemer, M.: Volumetric thermo-elastic machine tool behavior. Prod. Eng. **9**, 119–124 (2015)
11. Scharschmidt, K.H.: Modellbasiertes Verfahren zur Kompensation thermoelastischer Verlagerungen von Werkzeugmaschinen. PZH, Produktionstechn. Zentrum, Garbsen (2011)
12. Weng, L., Gao, W., Lv, Z., Zhang, D., Liu, T., Wang, Y., Qi, X., Tian, Y.: Influence of external heat sources on volumetric thermal errors of precision machine tools. Int. J. Adv. Manuf. Technol. **99**, 475–495 (2018)
13. Yang, J., Zhang, D., Mei, X., Zhao, L., Ma, C., Shi, H.: Thermal error simulation and compensation in a jig-boring machine equipped with a dual-drive servo feed system. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. **229**, 43–63 (2015)
14. Zhang, H., Yang, J., Zhang, Y., Shen, J., Wang, C.: Measurement and compensation for volumetric positioning errors of CNC machine tools considering thermal effect. Int. J. Adv. Manuf. Technol. **55**, 275–283 (2011)
15. Zhang, Y., Yang, J., Xiang, S., Xiao, H.: Volumetric error modeling and compensation considering thermal effect on five-axis machine tools. Proc. Inst. Mech. Eng. Part C J. Mech. Eng. Sci. **227**, 1102–1115 (2013)
16. Fritsch, F.N., Carlson, R.E.: Monotone Piecewise Cubic Interpolation. SIAM J. Numer. Anal. **17**, 238–246 (1980)
17. Gebhardt M.: Thermal behaviour and compensation of rotary axes in 5-axis machine tools. ETH Zürich (2014)
18. Siciliano, B., Khatib, O.: Springer Handbook of Robotics, 1st edn., pp. 13–14. Springer, Heidelberg (2008)
19. Limpert, E., Stahel, W.A., Abbt, M.: Log-normal distributions across the sciences: keys and clues. Bioscience **51**, 341 (2001)
20. Riedel, M., Müller, J., Ihlenfeldt, S., Deutsch, J.: Design of a photogrammetric measurement of displacement and deformation on machine tools. In: Conference on Thermal Issues in Machine Tools, pp. 421–432. Verlag Wissenschaftliche Scripten, Dresden (2018)

Inline Measurement of Process Forces and Development of a Friction Model in Abrasive Flow Machining

S. Roßkamp[1(✉)] and E. Uhlmann[1,2]

[1] Institut Werkzeugmaschinen Und Fabrikbetrieb, Technische Universität Berlin, Pascalstraße 8–9, 10587 Berlin, Germany
simon.rosskamp@iwf.tu-berlin.de
[2] Fraunhofer Institut Produktionsanlagen und Konstruktionstechnik, Pascalstraße 8–9, 10587 Berlin, Germany

Abstract. Abrasive flow machining is widely used for finishing applications. For every workpiece material, chips are formed by the relative movement between abrasive grains and workpiece surface. Thus, the friction in between is an important parameter to describe these mechanisms. A method for inline measuring of process forces in abrasive flow machining is presented. A sealed, technical device for measuring tangential forces and a device for measuring the static pressure were successfully developed. The results show a strong dependence on processing parameters. The data is used for developing a friction model that takes processing parameters into account. Moreover, the friction model is utilized to implement the local slip of the abrasive media on the workpiece surface in flow simulations. Consequently, process designs of machining tasks will be improved and will contribute to an enhanced quality and process stability. By process simulations, time consuming experiments are reduced in order to reduce costs.

Keywords: AFM · Process forces · Friction model · Flow and process simulation

1 Introduction

Innovations in manufacturing lead to developments of new machining technologies like additive manufacturing in 1981. Especially additively manufactured metallic workpieces lack surface quality so that they must be finished just like some conventionally manufactured workpieces. Even inner surfaces of these expensive complex shaped workpieces cannot be reached by stiff tools. Here, abrasive flow machining provides an opportunity for reducing surface roughness, removing burrs and rounding edges.

The mechanisms of chip formation in abrasive flow machining have rarely been investigated yet, as the process takes place in a closed system without insight.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 334–344, 2021.
https://doi.org/10.1007/978-3-662-62138-7_34

Independent from the theory of chip formation, the interaction between fluid, grain, and workpiece surface characterize this process. As a result, the knowledge of the occurring friction is important for process understanding. Against this backdrop, knowledge about friction in abrasive flow machining should be gained for implementing the findings in flow simulations. This way, enhanced two-phase flow simulations will reduce time required for process designs. By conducting fewer pilot tests, the costs will be reduced.

2 State of the Art

2.1 Abrasive Flow Machining

The machining technique of abrasive flow machining was developed in the middle of the last century [1, 2]. Two coaxial cylinders are closed by a piston in each cylinder. By clamping a heat exchanger and a workpiece holder between these cylinders, a closed system is built [3, 4]. The machining technique is shown in Fig. 1 and includes the red cross-hatched workpieces in the workpiece holder. Moreover, the abrasive media is illustrated for a state in the middle of a cycle. After the lower piston reached a set pressure, the movements of both pistons start until they reach the upper dead center. A cycle is completed with return of the pistons to the lower dead center [5, 6].

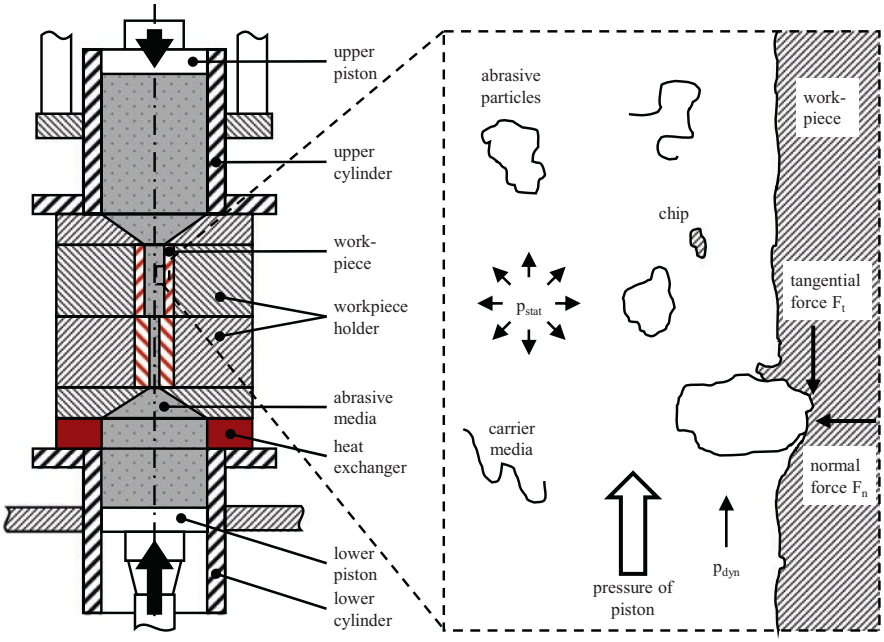


Fig. 1. Machine technique and chip formation in abrasive flow machining [7]

The mentioned abrasive media consists of a mixture of abrasive grains, a carrier fluid, which is mostly a long chain polymer, and additives for adjusting the rheological behavior [4, 8]. Due to the movement of the pistons, the abrasive media is moved along the workpiece surface. The resulting relative velocity between abrasive grains and workpiece surface at the processing location as well as additional physical flow parameters lead to local processing parameters [7, 9]. Besides the local velocity v_l, these are the local static pressure $p_{l,stat}$, the local dynamic pressure $p_{l,dyn}$, the local dynamic viscosity η_l and the local temperature T_l. The local static pressure $p_{l,stat}$ causes the normal force F_n, while the local dynamic pressure $p_{l,dyn}$ results in the tangential force F_t [10, 11]. Depending on the ratios between normal and tangential forces, different movement patterns such as sliding-rubbing, grooving or rolling arise [12, 13]. As a result, the chip formation is similar to that of conventional grinding like micro cutting and micro ploughing and is valid for spherical and ellipsoidal grains [14, 15]. Due to the closed system, it is not possible to prove this theoretical analysis in real processes yet [4]. Despite this, it is shown that the movement patterns as well as the chip formation depend on material, its hardness, and friction at the processing location [16, 17].

2.2 Friction and Fluid Mechanics

Generally, friction is divided into dynamic friction for effects in motion and static friction in an unmoved case [18]. Under both circumstances, the friction coefficient can be described as the quotient of tangential and normal forces as shown in Fig. 1. Here, the friction according to Coulomb [19] is used. Other formulations of friction laws or special cases such as rolling friction are not considered. Regarding abrasive flow machining, there is no publication that deals with friction in a significant scope.

The description of fluids is made by rheological characteristic values. A common one is the viscosity η that can be measured for homogeneous fluids [20]. In case of suspensions with two phases like abrasive media, the measurement of rheological characteristic values is not possible because the abrasive particles' diameters are bigger than the permitted gap between the plates of the rheometer [21]. In this case, Eq. (1) by Krieger Dougherty [22] can be used.

$$\eta = \eta_f \cdot \left(1 - \frac{\alpha_p}{\alpha_{p,max}}\right)^{-\chi \cdot \alpha_{p,max}} \tag{1}$$

In this equation, the viscosity of the abrasive media η depends on the viscosity of the carrier fluid η_f. In a sense of a correction term, a factor is multiplied with information about the packing density α_p and the maximum possible packing density $\alpha_{p,max}$ as well as the morphology factor χ, which is a numerical measure for the grain shape.

For describing the flow properties, the Reynolds number Re is used, which contains the fluid's density ρ, the mean flow velocity $\bar{v}$, a characteristic length, which is the pipe diameter d in case of a pipe flow, and the fluid's dynamic viscosity η [23]. For the devices and processing parameters used, the Reynolds number Re shows the low value of $3.88 \cdot 10^{-4}$, which is mainly influenced by the high dynamic viscosity η of 11880 Pa·s. Thus, laminar flow can be assumed [23].

In flow mechanics, the requirement of no-slip is framed for simplification. This cannot be fulfilled in abrasive flow machining because the relative velocity v_{rel} between an abrasive grain and the workpiece surface would show a value of zero. But without relative movement, chip formation is not possible. On the other hand, full slip would imply flow mechanics without any kind of friction. This cannot be fulfilled either. Against this backdrop, the investigations should be used for clarification of several issues. Besides questions concerning normal and tangential process forces, the friction coefficient and the degree of arising slip are to be calculated. With these findings, models of friction and slip can be developed and implemented in flow simulation. With these developments, the process simulation will show higher quality, which saves time in process design and leads from an increased flexibility to a higher agility.

3 Experimental Set-up

A mandatory requirement of the device for measuring tangential forces is a combination of a leakproof device and the certainty to measure only tangential forces. After some concepts and a few tests, a solution was found and is displayed in Fig. 2. The workpieces with an inner diameter d_i of 9 mm are clamped in a metallic cylinder by two plates with force sensors in between, which is pictured in Fig. 2a. The force sensors are both Typ 9135B from Kistler Instrumente GmbH, Winterthur, Switzerland; a technical drawing is shown in Fig. 2b. These force sensors can be used for measuring axial forces, which correspond to the process tangential forces F_t. The ring design is used for channeling the abrasive media through the measuring tier. For avoiding any friction

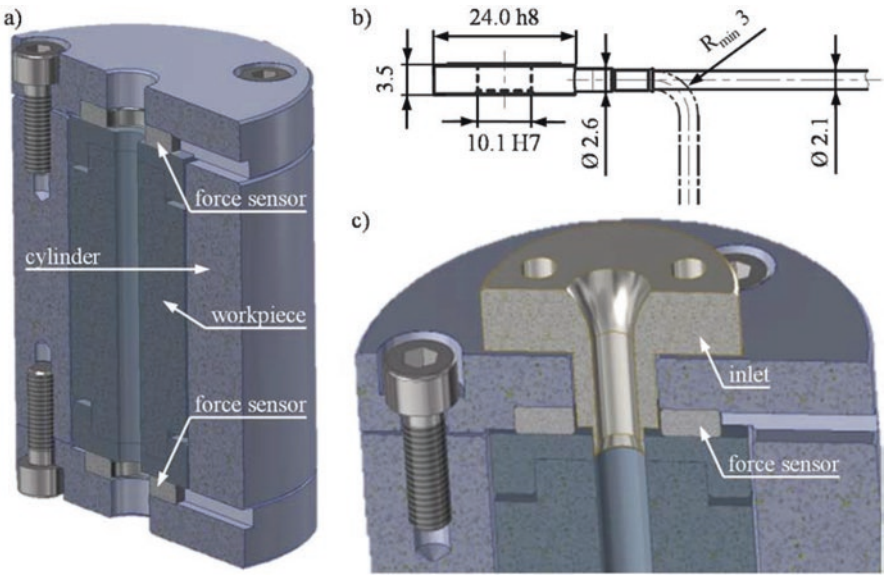


Fig. 2. Device for measuring tangential forces; a) assembly with workpiece; b) technical drawing of force sensor [24]; c) inlet for laminar flow without flow constrictions

on the outer surface of the workpieces, they are clamped with radial allowance. Rough centering is achieved by the inlet illustrated in Fig. 2c. It includes a radial allowance for the inner diameter of the workpiece. The design of the inlet in the shape of a continuous function is chosen to enable a laminar flow of the abrasive media. The sensors must be preloaded by the three screws on each plate. The calibration was conducted reliably and reproducibly with a rod and a polymeric tube for traction.

The surface roughness of the workpieces is measured at nine positions in depths of 10 mm, 25 mm and 40 mm and angular positions of 0°, 120° and 240° by a stylus instrument SJ-410 from Mitutoyo Corporation, Kawasaki, Japan. The processing parameters are chosen according to previous examinations to ensure the comparability of the results. They were published by Uhlmann et al. for the first time in 2016 [3].

4 Results

4.1 Process Forces

After performing technological investigations, every data series received were analyzed in accordance with the piston movement. As a result, for every half of a cycle, one value for the tangential force F_t was calculated. Variations in measurement data are considered by determining the mean value for all measuring points, which are in the reliable range of 95% of all values. By this procedure, statistical outliers are taken in account, which can be explained technologically. The correlation between technological values is pictured in Fig. 3. The diagram should be read in three ways. Firstly, the assumed course of surface roughness Rz depending on the number of cycles is

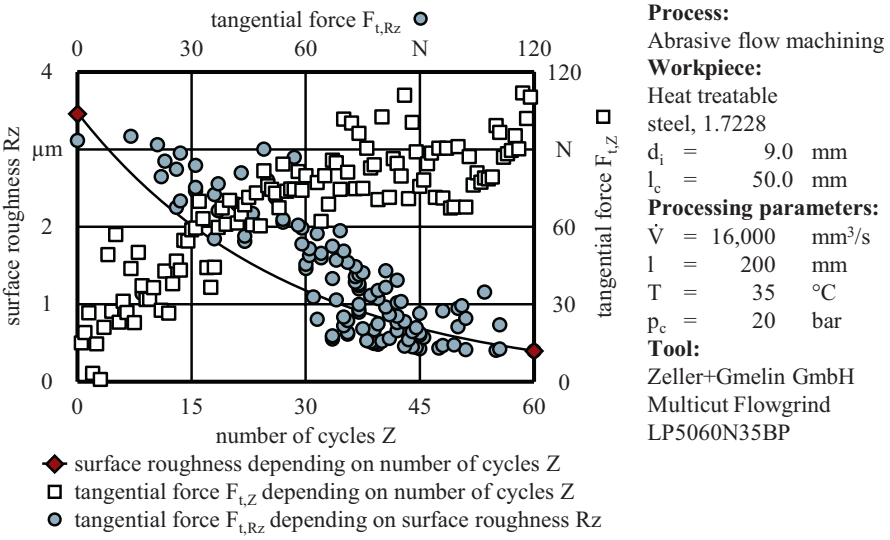


Fig. 3. Correlations between surface roughness Rz, tangential force $F_{t,Rz}$, number of cycles Z and tangential force $F_{t,Z}$

indicated by red rhombi. The surface roughness could not be measured after every cycle, as process interruption would cause variations in media temperature. The pictured regressive trend is based on measurements at cycle numbers zero and 60 and a model that was derived from numerous technological investigations with identical parameters beforehand.

Secondly, the white squares show a slightly degressive correlation between the tangential force $F_{t,Z}$ and the number of cycles Z. This time-dependent correlation could be explained by the changing values of surface roughness Rz. Therefore, the third part of the diagram is given. The blue circles show the tangential force $F_{t,Rz}$ depending on the surface roughness Rz on the vertical axis. This regressive curve shape confirms the theory that the tangential force F_t only seems to be time-dependent, which is an indirect connection. Primarily, the tangential force F_t behaves reciprocally to surface roughness.

In addition to the tangential force F_t, the normal force F_n will be important for the physical value friction. Due to the dependence between static pressure p, the loaded surface A and the normal force F_n, the processing parameter static pressure p is used in Fig. 4. The data points of the static pressure p are displayed in correlation to the number of cycles Z, indicated by the brown circles. The connections between these measuring points are drawn considering the ascending number of cycles Z. A strong dependence cannot be identified, but a slightly degressive behavior of the static pressure p in relation to an ascending number of cycles Z can be assumed. This theory is illustrated by the black line, which displays the regression between both parameters. An explanation for this trend is the increasing temperature of abrasive media with processing time. This behavior is influenced by parameters such as machining technique, processing parameters, geometry and material of the workpiece. In the considered technological investigations, this is plausible and in accordance with the results in literature [7, 25, 26].

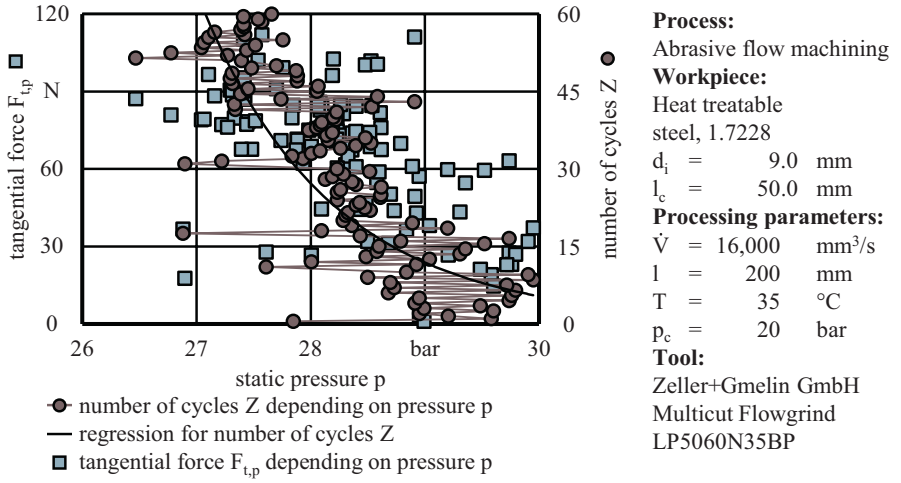


Fig. 4. Tangential force $F_{t,p}$ depending on static pressure p and association between static pressure p and ascending number of cycles Z with its regression

The main purpose of Fig. 4 is to show the weak connection between the tangential force $F_{t,p}$ and the static pressure p. An extremely cautious interpretation could suggest a reciprocal behavior of the tangential force $F_{t,p}$ in relation to the static pressure p. Contrary to this hypothesis, the mean value of about 68 N and its standard deviation of approximately 25 N, which is more than 20% of the range, should be mentioned. Due to the large variance of the tangential force $F_{t,p}$ and the static pressure p, which is in linear relation with the normal force F_n, a high variance of the friction coefficient μ can be expected. Meanwhile, some correlations with a power around two were recognized.

These findings are used to develop the regression depicted in Fig. 5. The slightly degressive behavior of the static pressure p in relation to an ascending number of cycles Z can be seen in the shape of the regression surface. As the friction coefficient μ is the quotient of tangential force F_t and normal force F_n, the spread width in the correlation between static pressure p and friction coefficient μ is plausible. The proportional correlation between the number of cycles Z and the friction coefficient μ can be seen very clearly in the blue framed diagram. In this context, the regressive behavior of the surface roughness Rz against the number of cycles as well as the reciprocal correlation between the tangential forces F_t and the surface roughness Rz should be kept in mind.

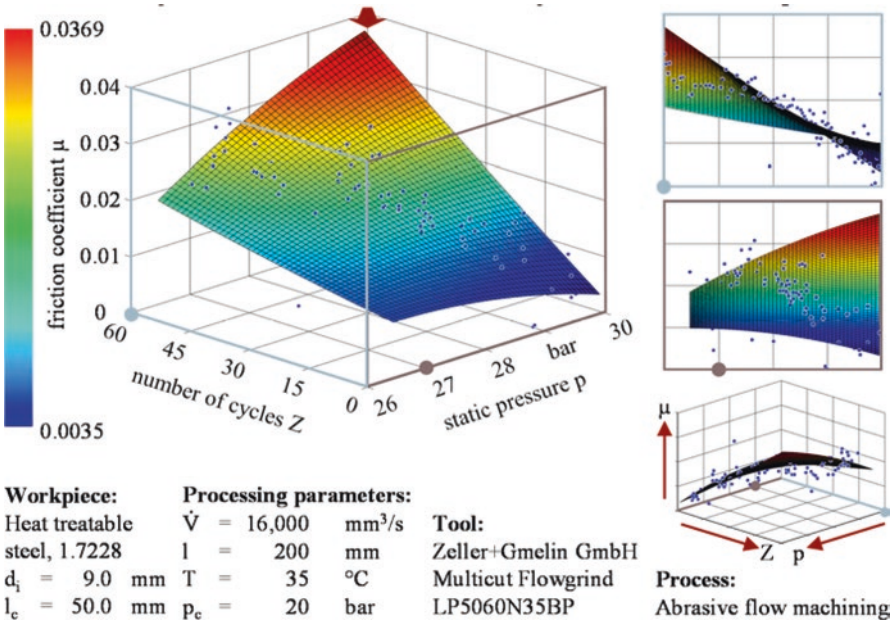


Fig. 5. Friction coefficient in dependence of number of cycles Z and static pressure p

The regression surface as a quadratic approach is derived on the basis of the observations made above. Despite the scattering of the measured values of some technological parameters, the coefficient of determination R^2 is at about 0.76. It is higher for the correlation of other technological parameters and for more input parameters in

the model. The distance between measuring points and regression surface can be seen in the three diagrams on the right side of Fig. 5. The blue framed diagram shows the surface in the plane of the number of cycles Z and the friction coefficient μ, which is displayed for the correlation with the static pressure p in the brown framed diagram. A view of the regression in direction of the red arrow is shown below. Brown and blue circles support the assignment of the axes and their ascending axis values.

4.2 Friction Model

A friction model is expected to improve the existing flow simulation as no-slip and full-slip conditions (Sect. 2.2) and occurring friction on the surface have an impact on local processing parameters. Therefore, the values of local friction are described in dependence of local processing parameters. This can be done by calculating local slip, which is implemented in ANSYS CFX in the form of a superimposed wall velocity u_w. The formula is shown in Eq. (2), with a slip velocity u_s, a wall shear stress τ_w, a critical shear stress τ_c and a normalizing shear stress τ_n. Moreover, the equation contains the slip exponent m, a pressure coefficient B, and depends on the pressure p.

$$u_w = u_s \cdot \left(\frac{\tau_w - \tau_c}{\tau_n} \right)^m \cdot e^{-\frac{B \cdot p}{\tau_n}} \tag{2}$$

The main objective is the development of a process simulation in dependence of local processing parameters, so that the slip velocity u_s and the pressure p are defined as such. For this, the elaborated data can be applied. Initial point of the derivation is the formula for partial slip of Ostwald deWaele-fluids in tube flows, which is shown in Eq. (3).

$$u(r) = \frac{3n+1}{n+1} \cdot (\bar{u} - u_w) \cdot \left(1 - \left(\frac{r}{R} \right)^{\frac{1}{n}+1} \right) + u_w \tag{3}$$

Transformations and applying basic equations of fluid mechanics lead to an analytical equation for the wall velocity u_w, which is displayed in Eq. (4) [20, 23, 27].

$$u_w = \bar{u} - \left(\frac{n}{3n+1} \right) \cdot R \cdot \left(\frac{\tau_w}{K} \right)^{\frac{1}{n}} \tag{4}$$

The wall velocities u_w from Eqs. (2) and (4) must be identical. From rheological qualifications of the carrier fluid and the technological investigations, the parameters are known or can be calculated numerically and yield the values given in Table 1.

Table 1. Numerically calculated solutions

Parameter	B	m	τ_c	τ_n	$u_{s,c}$	$u_{s,p}$
Unit	1	1	N/m^2	N/m^2	m/s	m/s
Value	$3.85 \cdot 10^{-6}$	-0.2715	$2.36 \cdot 10^{-14}$	133.3	-3.793	$5.59 \cdot 10^{-7}$

Special attention should be given to the two righthand columns. Evaluations showed a strong dependence between the total pressure p_{tot}, as the sum of static pressure p and the dynamic pressure p_{dyn}, and slip velocity u_s of the abrasive media on the workpiece surface. With a share of a constant slip velocity $u_{s,c}$ and a share of a pressure dependent slip velocity $u_{s,p}$, the quality of the flow simulation is highly improved. An example for a flow simulation with implemented wall slip is given in Fig. 6.

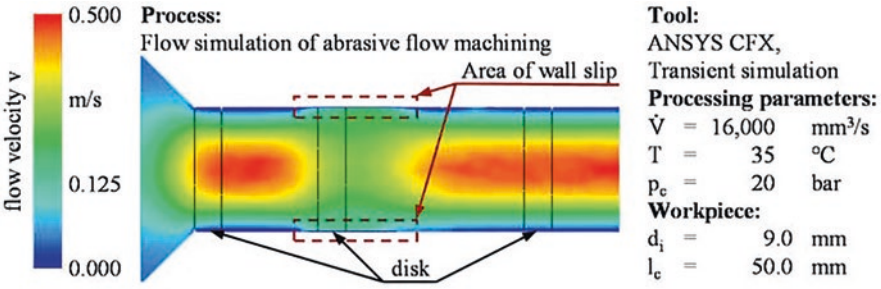


Fig. 6. Flow simulation in abrasive flow machining with implemented wall slip

In this flow simulation, the local processing parameters are calculated within the disks. With these results, wall slip is simulated and shows a strong influence on the flow pattern of the abrasive media. The wall velocity u_w increases due to wall slip in the disks, which results in a decrease of the flow velocity profile because of the continuity equation. After a short length, the wall slip ends and the flow velocity profile returns to the known pattern.

5 Summary and Outlook

Knowing that friction is an important physical value in abrasive flow machining, the need for detailed investigations was recognized. A device for inline measuring of tangential forces F_t is developed and is supplemented by the sensor for measuring static pressure p, which was described previously [7]. The results of technological investigations showed correlations between the tangential force F_t, the static pressure p, the number of cycles Z and the surface roughness Rz. The data was used for describing the friction coefficient μ and the connection to local processing parameters. With these findings, it was possible to derive a wall slip model in ANSYS CFX. This model includes the dependence from local processing parameters. As a result, the flow simulations are improved for higher simulation quality. In the future, further investigations are to be conducted. Correlations between the local velocity v_l and the tangential force F_t are still being observed. Technological investigations with a pure carrier fluid could reveal new insights into the distribution of the friction across carrier fluid and abrasive grains. The presented results can be a basis for this research.

In combination with the process model, the flow simulation can be used for the prediction of processing results in improved process design enhanced quality of

process stability. Time consuming experiments can be reduced in order to reduce costs. Both components are valuable for products with a high variant diversity. Thus, the described results find promising applications as finishing of additively manufactured workpieces.

References

1. Balman, C.: Honing apparatus. U.S. Patent 3,039,234 (1962)
2. Luckey, G.: Method of oliving jewels. U.S. Patent 2,419,687 (1945)
3. Uhlmann, E., Mihotović, V., Roßkamp, S., Dethlefs, A.: A pragmatic modeling approach in abrasive flow machining for complex-shaped automotive components. Procedia CIRP **5**(46), 51–54 (2016)
4. Dong, Z., Ya, G., Liu, J.: Study on machining mechanism of high viscoelastic abrasive flow machining for surface finishing. Proc. Inst. Mech. Eng. Part B J. Eng. IMechE **231**(4), 608–617 (2017)
5. Mahalik, N.: Micromanufacturing and Nanotechnology. Springer, Berlin (2006)
6. Jain, V., Kumar, R., Dixit, P., Sidpara, A.: Investigations into abrasive flow finishing of complex workpieces using FEM. Wear **267**(1–4), 71–80 (2009)
7. Roßkamp, S., Uhlmann, E., Hofmann, R., Gröger, S.: Shape alterations and their holistic geometrical representation in abrasive flow machining. In: Wulfsberg, J., Hintze, W., Behrens, B. (eds.) Production at the Leading Edge of Technology, pp. 249–258. Springer, Berlin (2019)
8. Tzeng, H., Yan, B., Hsu, R., Chow, H.: Finishing effect of abrasive flow machining on micro slit fabricated by wire-EDM. Int. J. Adv. Manuf. Technol. **34**(7–8), 649–656 (2007)
9. Hashimoto, F., Yamaguchi, H., Krajnik, P., Wegener, K., Chaudhari, R., Hoffmeister, H., Kuster, F.: Abrasive fine-finishing technology. CIRP Ann. **65**(2), 597–620 (2016)
10. Gorana, V., Jain, V., Lal, G.: Forces prediction during material deformation in abrasive flow machining. Wear **260**(1–2), 128–139 (2006)
11. Li, J., Su, N., Wei, L., Zhang, X., Yin, Y., Zhao, W.: Study on the surface forming mechanism of the solid-liquid two-phase grinding fluid polishing pipe based on large eddy simulation. Proc. Inst. Mech. Eng. Part B J. Eng. Manuf. **233**(14), 2505–2514 (2019)
12. Fang, L., Sun, K., Cen, Q.: Particle movement patterns and their prediction in abrasive flow machining. Tribotest **13**(4), 195–206 (2007)
13. Murali, K., Raju, H.: Extrusion honed surface characteristics of inconel 625 fabricated by EDM for square shape. Int. J. Eng. Res. Appl. **4**(6), 68–72 (2014)
14. Fang, L., Zhao, J., Li, B., Sun, K.: Movement patterns of ellipsoidal particle in abrasive flow machining. J. Mater. Process. Technol. **209**(20), 6048–6056 (2009)
15. Sankar, M., Jain, V., Ramkumar, J.: Experimental investigations into rotating workpiece abrasive flow finishing. Wear **267**(1–4), 43–51 (2009)
16. M'Saoubi, R., Outeiro, J., Chandrasekaran, H., Dillon, O., Jawahir, I.: A review of surface integrity in machining and its impact on functional performance and life of machined products. Int. J. Sustain. Manuf. **1**(1/2), 203–236 (2008)
17. Gorana, V., Jain, V., Lal, G.: Experimental investigation into cutting forces and active grain density during abrasive flow machining. Int. J. Mach. Tools Manuf. **44**(2–3), 201–211 (2004)
18. Cozza, R., Tanaka, D., Souza, R.: Friction coefficient and abrasive wear modes in ball-cratering tests conducted at constant normal force and constant pressure—preliminary results. Wear **267**(1–4), 61–70 (2009)

19. Coulomb, C.: Théorie des machines simples. Memoire de Mathematique et de Physics de l'Academie Royal, 161–342 (1785)
20. Neto, C., Evans, D., Bonaccurso, E., Butt, H., Craig, V.: Boundary slip in Newtonian liquids: a review of experimental studies. Rep. Prog. Phys. **68**(12), 2859–2897 (2005)
21. DIN 53019-1: Viskosimetrie: Messung von Viskositäten und Fließkurven mit Rotationsviskosimetern. Grundlagen und Messgeometrie. Beuth, Berlin (2008)
22. Krieger, I., Dougherty, T.: A mechanism for non-Newtonian flow in suspensions of rigid spheres. Trans. Soc. Rheo. **3**(1), 137–152 (1959)
23. Metzner, A., Reed, J.: Flow of non-Newtonian fluids—correlation of the laminar, transition, and turbulent-flow regions. Am. Inst. Chem. Eng. J. **1**(4), 434–440 (1955)
24. Kistler Gruppe: SlimLine Sensoren (SLS). Messen von dynamischen und quasistatischen Druckkräften, 0 … 3 kN bis 0 … 80 kN. Winterthur, Schweiz (2020)
25. Jain, R., Jain, V.: Specific energy and temperature determination in abrasive flow machining process. Int. J. Mach. Tools Manuf. **12**(41), 1689–1704 (2001)
26. Fang, L., Zhao, J., Sun, K., Zheng, D., Ma, D.: Temperature as sensitive monitor for efficiency of work in abrasive flow machining. Wear **266**(7–8), 678–687 (2009)
27. Böhme, G.: Strömungsmechanik nicht-newtonischer Fluide, 2nd edn. Vieweg, Wiesbaden (2000)

Fast Evaluation of Volumetric Motion Accuracy of Multi-axis Kinematics Using a Double Ballbar

R. Zhou(✉), B. Kauschinger, C. Friedrich, and S. Ihlenfeldt

Institute of Mechatronic Engineering, TU Dresden, Helmholtzstraße 7a, 01069 Dresden, Germany
ruiqing.zhou@tu-dresden.de

Abstract. In this paper, we present a measuring approach to assessing the volumetric accuracy of machine tools or kinematics with more than 3 interpolating feed axes. It is based on an inexpensive Double Ballbar and suitable for fast performance-tests in an operational state of the machine. The novelty is the recording of large amounts of data in a very short period while continuously moving along measuring paths with up to 6DoF. In order to facilitate this measuring approach, we developed a Double Ballbar with extended measuring range based on an optical sensor and a series of systematic methods, like kinematic error modeling, optimal design of 6D measuring paths as well as post-processing and evaluation of the captured data. We demonstrate that our approach is universally applicable to different types of kinematic structures such as 5-axis machine tools, industrial robots or even parallel kinematics. By experiments, we will also show that the measured data are essentially more informative than those of standardized circular DBB tests are. Moreover, the measuring procedure is automatable and to be carried out under production conditions in the workshop, what makes it practically applicable in various evaluation and decision scenarios.

Keywords: Double ballbar · Fast performance test · Volumetric accuracy · Multi-axis kinematics

1 Introduction

The accuracy of a machine tool largely determines the manufacturing precision of the workpiece. For the verification and validation as well as for the proof of the manufacturing accuracy, the collection and evaluation of information about the positioning accuracy of machine tool is of central importance.

One kind of approaches is acquiring machine accuracy through test workpieces, which is widely utilized for machine acceptance and repeatability test. For diverse machine types and manufacturing processes (3-axis-, 5-axis-, HSC-machine, see Fig. 1), various independent acceptance and test workpieces are defined in standards

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 345–353, 2021.
https://doi.org/10.1007/978-3-662-62138-7_35

and guidelines [1–3]. In addition, there are customer-specific test workpieces as well as test workpieces from control and machine manufacturers. However, those approaches have following essential shortcomings:

- The assessment of the machine accuracy is associated with a relatively high expenditure of material and time.
- The accuracy statement applies only to a very limited working area.

Fig. 1. Left: NCG-HSC test workpiece 2004, middle: NCG-5-axis test workpiece 2005, right: NCG-5-axis test workpiece for microchip removal 2007

In order to improve the efficiency, numerous measuring systems and their associated approaches for rapid performance test were invented, developed and some of them even standardized, such as tests on spatial diagonal [4], circular tests with Double Ballbar (DBB), a cross grid measuring device [4] or the AxiSet™ Check-Up [5] (Fig. 2). However, the assessment from above approaches is restricted to one or two moving axes and only partially instructive. Among those measuring systems, the DBB is advantageous in many ways, especially:

- it is accurate. Typical commercial Double Ballbar could hold a measurement accuracy of 1 μm within 1 mm measuring range [6].
- it is cost efficient, both from the perspective of material and labor.
- it is also efficient in collecting information. Double Ballbar is capable of uninterruptedly capturing errors during 6-D machine movement.

Fig. 2. Rapid test for accuracy: left: comparison scale, middle: circular test with DBB, right: rapid rotatory axis test with AxiSet™ CheckUp

But it was pointed out in work [7, 8], a circular test with DBB could only provide partial information about the accuracy for a 5- or 6-axis machine. Even the assessment

obtained from circular tests with DBB in three main planes remains very limited [9]. In order to make DBB an universally applicable and cost-efficient measuring equipment for rapid performance test, following essential deficiencies in current Double Ballbar measuring system and approach must be remedied:

- It has to be capable of capturing larger errors (*inψmagnitudeψofψmillimeter*) in a much wider measuring range, so as to improve the measuring coverage in working space. There are only Double Ballbar with very restrictive measuring range (for instance, ±1 mm by Reinishaw QC20 [6]) available on the market, which does not suffice for this purpose.
- The measuring path (also the poses on it) must be delicately designed or generated, so that the DBB measurement could deliver a more reliable assessment about the accuracy of machine tools, including 5- or 6-axis machines with rotatory axes. This aims to make the measurement more representative in the whole working space. Researches have revealed the importance of poses for measurement [10].

Hence, in the following sections, we will first briefly introduce the IMD DBB with extensive measuring range and path generation algorithm, which aim to tackle the above problems. Subsequently, we present the continuous DBB measurement and validate our concept with experimental results on machines of both serial and parallel kinematic structure.

2 Measurement Approach and IMD DBB

The continuous DBB measurement could be roughly divided into four steps:

1. The Ballbar has to be calibrated on an extern calibration tool (Fig. 3) and the absolute bar length refers to the calibration value.
2. The center pivot is positioned on the machine table, the spindle is moved to a reference point and the test'zero' coordinates set (Fig. 4 left).
3. The machine is moved to the test start pose and the Ballbar is mounted between two kinematic magic joints (Fig. 4 right).
4. Ballbar measures uninterruptedly while the spindle moves along the measuring path. The measuring values would be either collected with the machine tool itself(intern) or with a dedicated device(extern).

Fig. 3. Zerodur calibrator

Thereafter it could be followed by further analysis and application. A Double Ballbar with extended measuring range was designed and manufactured in Institute of Mechatronic Engineering, TU Dresden (IMD) by considering that:

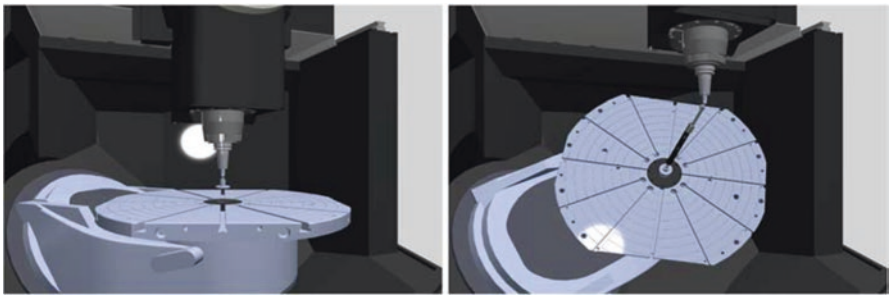


Fig. 4. Setup of Ballbar, left: spindle at reference('zero') position, right: spindle at start position and Ballbar mounted

- it can make a better coverage of the whole working space other than conventional measuring path such as circular test and
- it is capable of capturing the thereby arising larger error.

When compared to the conventional telescopic Ballbar, it utilizes the optical measuring principle instead of inductive (Fig. 5). The optical sensor head scans the scale tape and builds an incremental measuring system with a resolution of 0.1 μm. It is experimentally validated in our former work [11] that IMD DBB could reach an accuracy of 1 μm in 5 mm measuring range, which is comparable to commercial DBBs (≤1 μm in 100 mm measuring range by QC20 [6]).

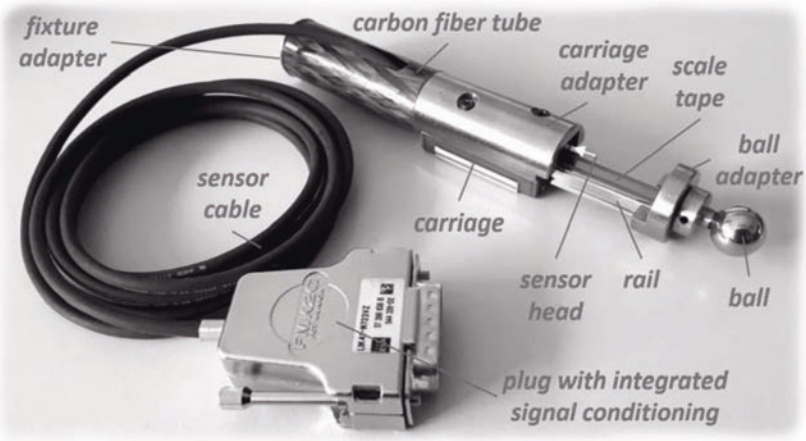


Fig. 5. IMD Double Ballbar

3 Path Generation

In work [12], Patel et al. have demonstrated that the DBB measurement possesses the potential to provide more information about the error source in kinematics with a set of well-designed poses that differ from a horizontal circle. Since practical experience shows that the largest part of the position error is caused by kinematic errors [13], it is of interest to design or generate such a spatial path that is feasible, collision-free, smooth and at the same time as sensitive as possible to the kinematics of the machine concerned.

In our former work [14], we developed a systematic and universally applicable approach to generating measuring paths so that they are able to facilitate DBB measurement with continuous data capture on various kinematics. For the sake of simplicity and also for comparison with the conventional DBB circular test according to [4], we generated in [14] a path with spherical curve, although the approach could also be applied to generate paths with varying measuring radius. For the same reason, we will also restrict the topic in this section to generating a path with spherical curve. Given the kinematic model of the machine concerned, the approach principally consists of the following four steps:

1. The sensitivity information are derived and the most sensitive poses within the working space (Fig. 6.1) are sought.
2. The sequence for the poses in step 1 is determined, in which the machine's tool center point (TCP) passes the sensitive poses found in step 1 (Fig. 6.2). Criteria, that should be taken into account, include the curve length, the kinematic feasibility and kinetic constraints.
3. Equidistant support points for the translational movement of TCP (also the ball on the spindle) are generated along a smooth curve crossing the poses in step 2 (Fig. 6.3).
4. The orientation for support points in step 3 is interpolated and optimized, so that the resulting path is feasible, free of collision and as smooth as possible in joint space (Fig. 6.4).

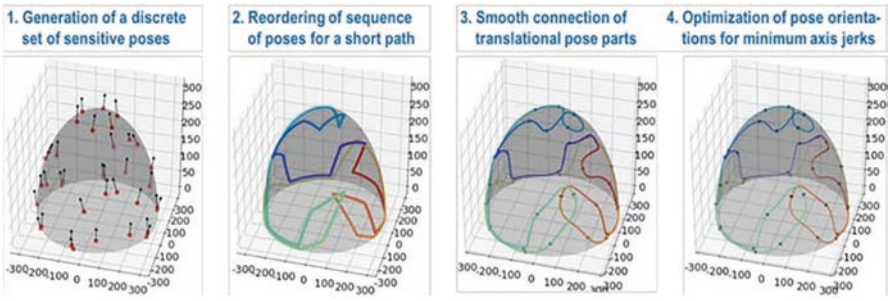


Fig. 6. Generation of 6D path with constant nominal radius

In our work [14], we also compared various indices for the sensitivity analysis of the generated path and concluded that the condition number of the identification Jacobian

matrix derived from the kinematics of the concerned machine is a proper index to describe the sensitivity in the case of continuous DBB measurement. Moreover, from a mathematical point of view, a well-conditioned measuring path could also reduce the influence of the nonkinematic factor on the calibrated kinematic parameters, which potentially makes the calibration more precise, as we also analyzed in our above work.

4 Experimental Results

In order to confirm the universal applicability of the measuring instrument and associated approach (see Sect. 2), we experimentally validated our concept on machines of different kinematics. For experimentation platforms we have chosen:

- Felix – a 6–6 Stewart-Gough platform designed in IMD (Fig. 7 left) – as an example of parallel kinematic machine (PKM) and
- DMU 80 evo linear (Fig. 7 right) as a representative of conventional 5-axis machine tool with serial mechanism.

Fig. 7. Experiment Platforms: left: Felix; right: DMU 80 evo linear

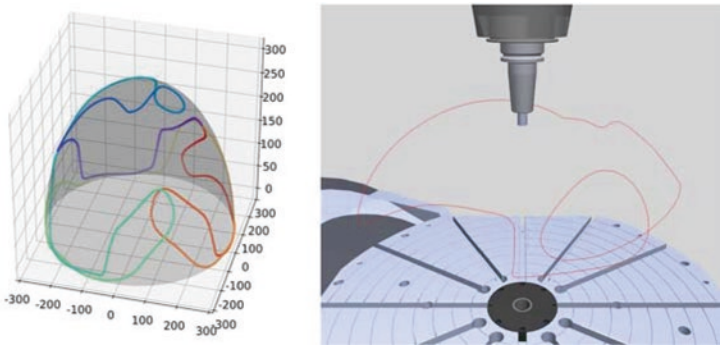


Fig. 8. Spherical curve of generated paths: left: Felix, right: DMU 80 evo linear

Figure 8 illustrates the spherical curve of the generated paths for Felix in its reference coordinate system (rainbow curve in the left subfigure) and for DMU 80 evo linear in the local table coordinate system (red curve in the right subfigure). Both paths have a nominal radius of 300 mm. We carried out the measurement along the generated path and subsequently the conventional DBB circular test according to [4] with identical nominal radius, for each machine respectively, so that the environment was kept as constant as possible during all measurements and the measuring results are thereby comparable.

Figures 9 and 10 compare the time sequence of the DBB measuring errors (Δr in legend) along both paths in the left subfigure and visually summarize the value range in the right subfigures on Felix and DMU 80 evo linear, respectively. Herein, 'circular' stands for circular test and 'spatial' for the generated path. In order to make the difference clearer, the y-axes of both time sequences within each figure were scaled to the same range. As shown, all the measurements were carried out successfully without collision.

On both machines, the measurement along the generated paths possesses significantly larger value range compared to the circular test (2.3153 mm versus 0.2268 mm on Felix and 0.6599 mm versus 0.1071 mm in the case of DMU 80 evo linear) and simultaneously could end within 5 min, which makes it suitable even for time-demanding tasks.

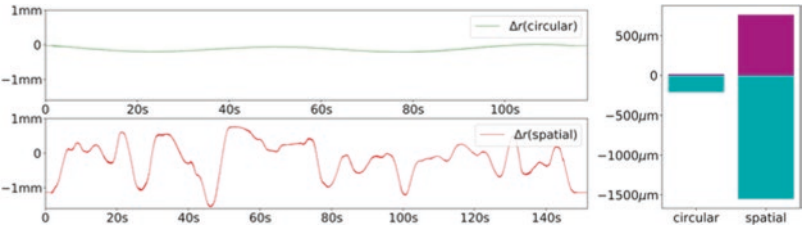


Fig. 9. Comparison of measurement on Felix, left: value-time curve, right: summary of value range

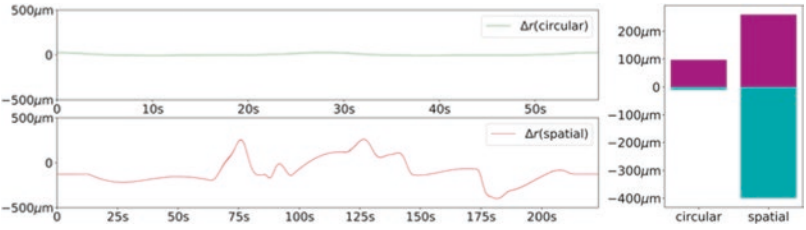


Fig. 10. Comparison of measurements on DMU 80 evo linear, left: value-time curve, right: summary of value range

The same conclusion could be drawn in the frequency domain. Figure 11 visualizes the DBB measuring values in lower frequency region on both machines after

FFT. In both cases, the frequency curve of circular test attenuates very fast with increasing frequency, which means it could only collect the error signal in very low frequencies and hence provides only very restrictive information about the kinematic error, as we stated in Sect. 1. In contrast, the frequency curve of the measurement along the generated path has a much larger amplitude and attenuates much slower in the corresponding frequency region, which not only implies it is more sensitive to the kinematics of the machine concerned but also reveals it possesses higher information content. For qualitative applications such as rapid accuracy test, the generated measuring path includes the most sensitive poses and therefore provides a more reliable assertion about the accuracy of a multi-axis machine. Aside from that, the optimized path makes more kinematic errors observable in the measuring values and at the same time makes them more distinguishable from errors caused by other sources such as mechanic vibration, temperature, etc. This might improve the identifiability of kinematics with DBB measurement, which we would like to validate in further research.

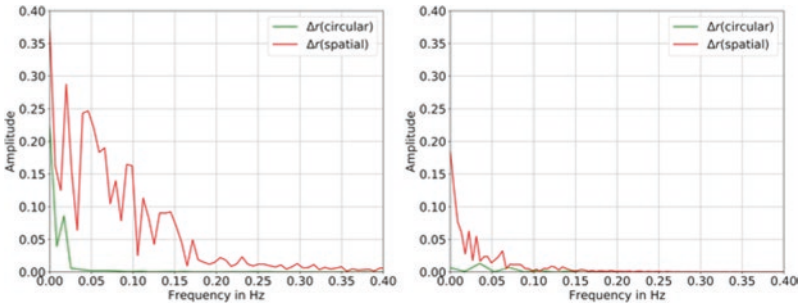


Fig. 11. Measuring values in frequency domain, left: Felix, right: DMU 80 evo linear

5 Conclusion

In this paper, we validated the concept of continuous Double Ballbar measurement by applying the instrument and methods we developed to machines with different kinematics. The experimental results verified the universal applicability of our approach. One could also see that the measuring approach could end within minutes. Additionally, considering that the setup of Double Ballbar on machines takes much less effort and time than measuring instruments based on other principles such as laser, we treat our approach as a competitive candidate for universal applicable rapid test under production condition.

Furthermore, as we pointed out in our work [14], from the theoretical point of view, a measuring path with well-conditioned identification Jacobian matrix is also more suitable for application such as calibration than otherwise designed paths, which we will try to verify through experiments in our future research.

Acknowledgement. This research is funded by DFG Project IH 124/3-1: Vermessung und Bewertung der räumlichen Bewegungsgenauigkeit an mehrachsigen Werkzeugmaschinen unter Betriebsbedigungen.

References

1. ISO 10791:2001 – machine tools – test conditions for machining centres
2. National Aerospace Standard Washington 913: Bahngesteuerte Fräsmaschinen (1974)
3. NC-Gesellschaft: Prüfwerkstück für die 5-Achs-Simultan-Fräsbearbeitung (2005)
4. ISO 230:2005 – test code for machine tools
5. Reinishaw AxiSet Check-Up. https://www.renishaw.de/de/axiset-check-up--11353. Accessed 25 Apr. 2020
6. Reinishaw QC20-W ballbar system. https://www.renishaw.com/en/qc20-w-ballbar-system--11075. Accessed 25 Apr. 2020
7. Soons, J.A.: On the Geometric and Thermal Errors of a Hexapod Machine Tool. Parallel Kinematic Machines, pp. 151–169. Springer, London (1999)
8. Khalil, S., Besnard, W.: Identifiable parameters for parallel robots kinematic calibration. In: 2001 ICRA. IEEE International Conference on Robotics and Automation (Cat. No. 01CH37164), vol. 3, pp. 2859–2866 (2001)
9. Pahk, H.J., Kim, Y., Moon, J.H.: A new technique for volumetric error assessment of cnc machine tools incorporating ball bar measurement and 3d volumetric error model. Int. J. Mach. Tools Manuf. **37**, 1583–1596 (1997)
10. Driels, M.R., Pathre, U.S.: Significance of observation strategy on the design of robot calibration experiments. J. Rob. Syst. **7**(2), 197–223 (1990)
11. Kauschinger, B., Friedrich, C., Zhou, R., Ihlenfeldt, S.: Fast evalution of the volumetric motion accuracy of multi-axis machine tools using a double-ballbar. J. Mach. Eng. **20**, 24–42 (2020)
12. Patel, A.J., Ehmann, K.: Calibration of a hexapod machine tool using a redundant leg. Int. J. Mach. Tools Manuf. **40**(4), 489–512 (2000)
13. Stark, G., Benz, E., Hüttenhofer, M.: Calibration Experiences in Industry. Robot Calibration, pp. 283–295. Chapman & Hall, Cambridge (1993)
14. Zhou, R., Kauschinger, B., Ihlenfeldt, S.: Path generation and optimization for DBB measurement with continuous data capture. Measurement **155**, 107550 (2020)

Additive Processes

Evaluating the Cumulative Energy Demand of Additive Manufacturing Using Direct Energy Deposition

S. Ehmsen(✉), L. Yi, and J. C. Aurich

FBK - Institute for Manufacturing Technology and Production Systems, TU Kaiserslautern, Kaiserslautern, Germany
svenja.ehmsen@mv.uni-kl.de

Abstract. Additive manufacturing [AM] is often claimed as an environmentally friendly technology that also offers great potential for the industry. However, material and energy efficiency depend on a large number of influencing factors. Recent studies have focused on the quantitative evaluation of the environmental impact and energy demand of AM processes as well as the investigation of their impact factors. For powder production as well as post-processing there are only few studies available so far. This paper introduces an evaluation model to quantify and analyze the cumulative energy demand [CED] from cradle to gate using direct energy deposition [DED]. During the analysis, the process steps that have a significant impact on the CED are identified. It is observed that the proposed evaluation model is a powerful tool to analyze the energy performance of DED technology

Keywords: Additive manufacturing · Cumulative energy demand · Sustainability

1 Introduction

Additive manufacturing [AM] is a technique that adds material layer by layer to create parts [1]. Due to this, only the material required for the part is processed, which reduces the amount of process scrap [2–4]. The resulting resource efficiency is further supported by the fact that AM allows free design of part geometry and topologically optimized part design. This leads especially to material savings during manufacturing and energy savings during the use phase due to lightweight design and improved functional performance as well as to other environmental benefits [5–7]. Apart from these advantages, AM is also interesting for the industry due to the reduced production time and unit costs through fast and efficient manufacturing for small batch sizes and customized parts, which leads to competitive advantages [4, 7–10]. Furthermore, AM enables processing of materials that may be difficult to machine such as titanium [8].

It is expected that the importance of AM for the industry will increase significantly in the future. At the same time, climate change requires the development and

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 357–367, 2021.
https://doi.org/10.1007/978-3-662-62138-7_36

adoption of sustainable and energy-efficient manufacturing technologies. Therefore, it is necessary to quantify and minimize the environmental impact of AM at an early stage, before industry adapts the technology on a large scale. One of these promising AM-technologies is Direct Energy Deposition [DED], which melts the metal as it is being deposited to create parts layer by layer. In the context of the sustainability, the analysis of the environmental performance of DED is insufficient. Therefore, this paper aims to propose and validate an assessment model to determine the cumulative primary energy demand for a DED process from cradle to gate.

2 State of the Art

Several additive manufacturing technologies (e.g. selective laser melting) cause high energy demands due to long processing times up to several days. DED has compared to other metal-based processes a higher build-up rate which reduces the production time. Therefore, it is predicted to have great potential for the industry. In addition, DED allows for larger assembly volumes and creates parts with high density, strength and high dimensional accuracy. However, post-processing to achieve the desired surface quality and dimensional accuracy is still necessary [11–13].

A common approach to investigate the environmental impact of AM processes is a Life Cycle Assessment [LCA] (e.g. [14] and [15]). Approaches such as Baumers et al. [16] and Baumers et al. [17], that focus on energy efficiency only, usually consider the unit process and neglect the remaining process chain. Moreover, most of them concentrate on powder bed processes such as electron beam melting, selective laser sintering and selective laser melting (e. g. [18]).

In the existing literature, the discussion on the environmental and energy-related issues are still insufficient. Le Bourhis et al. [19] develop an approach to evaluate the environmental impact of a DED process based on the Eco-Indicator 99 method. In the inventory analysis, they consider not only electrical energy but also the demand for raw materials and supplies. Since the approach is mainly based on the AM process itself, the supply chain of raw materials in particular is not sufficiently considered. Liu et al. [13] also consider all environmental impacts of DED, quantified by a LCA. In contrast to Le Bourhis et al., they do not only concentrate on the AM process, but rather examine the process from Cradle to Cradle. They concluded that the DED process has the highest specific energy consumption [SEC] = MJ/kg, which describes the energy consumed to produce one unit, e.g. one kilogram, within the entire process chain. Besides, compared to conventional manufacturing methods DED has a greater environmental impact. A further comparison of DED and milling is carried out by Morrow et al. [3]. The authors compare the energy demands of three different parts, in which each part is manufactured with DED and milling. Within their system boundaries they consider the powder or raw blank production and the manufacturing processes. They observe that it depends on the solid-to-cavity ratio, whether DED or milling is preferable from an energy point of view. The solid-to-cavity ratio defines the ratio of the mass of a part and the mass that would be contained in the restrictive volumetric envelope of a part DED requires less energy for small ratios, milling

is more beneficial for large ratios. Bambach et al. [20] recognized that DED requires less energy than conventional manufacturing if the manufactured powder consists of 85% powder produced by secondary material. Therefore, the choice of raw materials has a significant impact on the SEC.

In assessing the literature, it is observed that the presented approaches do not sufficiently consider the raw material production and its further processing, the powder production as well as the post-processing of AM parts. Additionally, there is no sufficient identification of the factors that significantly influence the energy demand for those process steps. Although there are some studies focusing on the evaluation of the energy demand of DED, these studies are specific for the investigated part and depend on the respective assumptions and set system boundaries. Due to this specific character, these approaches do not provide a general model to assess the energy demand of AM parts. Thus, it is necessary to develop a generally valid model that covers the entire process chain, and calculate the energy demand for individual process steps as well as the total energy demand.

3 Modeling the Cumulative Energy Demand for the Direct Energy Deposition Process Chain

In the presented approach, the cumulative energy demand $[CED] = \text{MJ}$ is modeled, since the environmental performance of AM is significantly influenced by energy demand and an LCA is very complex [21]. In contrast to the LCA, the CED assessment only considers energy-related impacts, which leads to a reduction in complexity while giving a good approximation of the results of an LCA. The CED is defined as the entire primary energy demand that arises during the production, use and disposal of a product or service, or which may be attributed respectively to it in a causal relation [22]. The indicated system boundary requires a cradle to grave approach. However, for a comprehensive analysis from cradle to grave, the use phases of AM parts are case-specific and cannot be generally described. For this reason, a cradle to gate approach is chosen, in which only the process steps from raw material production to the finished part are considered, as shown in Fig. 1. The CED for this process chain can be calculated according to Eq. 1: Here, the subscripts represent the process steps that are considered in this research.

$$CED_{total} = CED_{transport} + CED_{raw\,material} + CED_{atomization} + CED_{DED} + CED_{post\,processing} \tag{1}$$

To assess the CED of a part manufactured by DED, the sub-processes and energy-relevant input factors must be identified for each process step. To achieve a model, which can calculate the CED as a function of the part, the specific CED, $[ced] = \text{MJ/kg}$ must be calculated, by multiplying the previously captured SEC of an input factor $[i]$ with its respective primary energy factor $[E]$ as shown in Eq. 2.

$$ced_i = SEC_i \cdot E_i \tag{2}$$

The primary energy factor indicates how much primary energy is needed to generate a unit, e.g. electricity or inert gas. The calculation of the CED for a specific part is shown in Eq. 3.

$$\begin{aligned} CED_{total} = & \sum_{t=1}^{T} \left(ced_{transport,t} \cdot m_{transport,t} \cdot s_t \right) + ced_{raw\,material} \cdot m_{raw\,material} \\ & + ced_{atomization} \cdot m_{atomization} + ced_{DED} \cdot m_{DED} \\ & + \sum_{z=1}^{Z} ced_{post\,processing,z} \cdot m_{post\,processing,z} \end{aligned} \tag{3}$$

where $[m] = kg$ represents the mass of the respective material or part, which differs between each process step. Besides, $[t]$ represents the different transport operations and $[z]$ the different post-processing treatments.

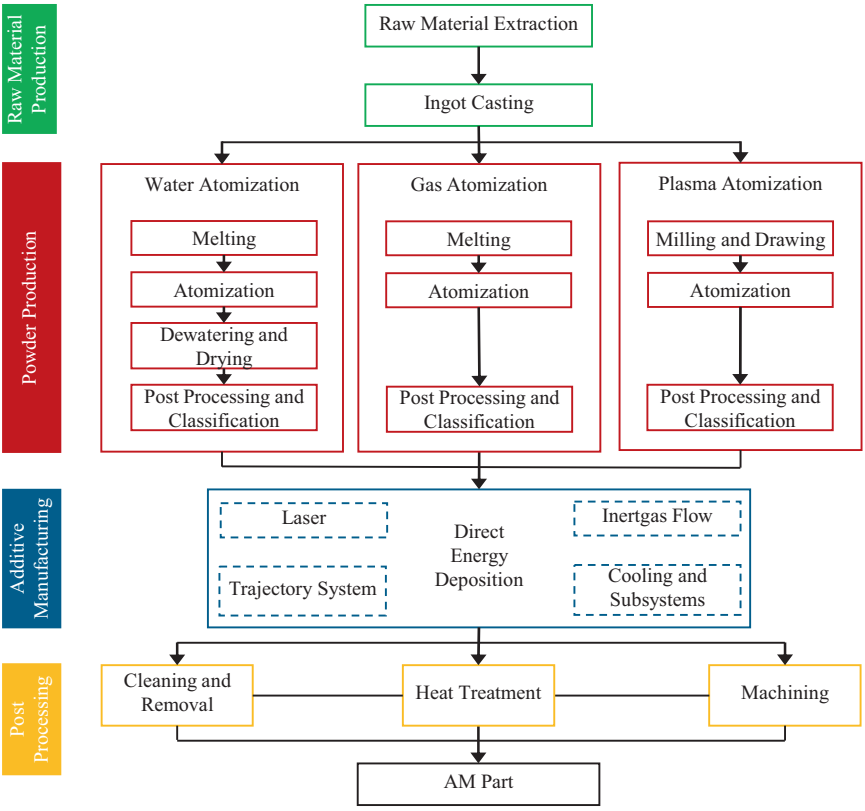


Fig. 1. Process chain of DED from Cradle to Gate

3.1 Transport

Transportation is mainly necessary between raw material extraction, powder production and AM. To determine the CED [$CED_{transport}$], all transport operations [t] must therefore be summed up. Previous assessments determined the energy demand of transportation to be very low, so that it is often neglected [13–15]. However, transportation is an explicit element in the CED approach and is therefore considered in the model. The CED depends on the mass of the transported good [$m_{transport,t}$], the transport distance [s_t] and the selected means of transport and its *ced* [$ced_{transport,t}$].

3.2 Raw Material Production

The exact process of raw material production differs depending on the raw materials and the country of extraction. Hence, the CED can only be calculated approximately. The CED of this process step, includes firstly the *ced* for resource extraction [$ced_{extraction}$]. This requires not only electricity but also the energy sources coal, crude oil or natural gas. After the extraction the first processing such as ingot casting [$ced_{casting}$] is carried out, in which mainly electricity consumption is required [13]. Hence, the CED of this process step depends on the *ced* for resource extraction [$ced_{extraction}$] and ingot casting [$ced_{casting}$], as shown in Eq. 4, and the extracted and casted mass [$m_{raw\ material}$].

$$ced_{raw\ material} = ced_{extraction} + ced_{casting} \tag{4}$$

3.3 Powder Production

There are three common methods of producing powder for AM: Water atomization, gas atomization and plasma atomization. For water and gas atomization, ingots are melted in an induction furnace and then transferred into the atomization chamber. As the melt falls through the chamber, water or gas jets atomize the melt. For water atomization afterwards de-watering and drying of the powder is necessary [12, 23, 24]. Most powders for AM are produced by gas atomization. To avoid oxidation, inert gas like argon or nitrogen is used during the process. During plasma atomization the raw material is fed directly to the atomization chamber as coarse powder or wire, where it is simultaneously melted and atomized by plasma torches and gas jets [23, 25]. In order to obtain the powders in the desired size, the powders are classified after the atomization. This can be done either by sieving, slurring or layering. Optionally, an oxidized particle surface, as occurs in particular with water atomization, can be counteracted by an annealing treatment [26].

Due to the several processes during atomization, many influencing variables have an effect on the CED. To identify factors that significantly influence the CED it is therefore necessary to consider the CED of different process systems [s] separately such as furnace or electrode, vacuum, water and inert gas pump, as well as subsystems. In addition to the *ced* of those systems, the *ced* of operating materials [o],

e.g. inert gas, and subsequent processing steps [p] have also to be considered. The *ced* for the atomization can be calculated as given in Eq. 5.

$$ced_{atomization} = \sum_{s=1}^{S} ced_{s,atomization} + \sum_{p=1}^{P} ced_{p,atomization} + \sum_{o=1}^{O} ced_{o,atomization} \tag{5}$$

3.4 Direct Energy Deposition

The CED of DED is related to different systems that are implemented in the AM machine. Firstly, energy is required for both the laser beam and for its regulation [19]. Secondly, the inert gas flows transport the powder and create a powder cone that allows the powder to be applied precisely to the part. Furthermore, inert gas also prevent oxidation [3, 19]. Thirdly, energy is required to move the deposition head or the building platform [11]. Lastly, the cooling system and supporting systems such as the control cabinet, heating systems, hydraulic components nearly have a constant energy demand [19, 27]. Therefore, the *ced* for DED can be calculated by summing up the *ced* s of the outlined systems as in Eq. 6.

$$ced_{DED} = ced_{laser} + ced_{inert\ gas} + ced_{trajectory\ system} + ced_{cooling\ system} + ced_{support\ systems} \tag{6}$$

Important parameters which can be set individually and which influence the *ced* are e.g. the layer thickness, the building speed and the component size.

3.5 Post-Processing

To meet industrial requirements, post-processing of AM parts is often necessary due to the staircase effect, low surface quality and insufficient dimensional accuracy. The required post-processing treatment depends on the AM-process, material and the intended application for the AM part [14, 28]. The post-processing options for DED can be divided into cleaning and removal, machining as well as surface finishing.

First of all, it is necessary to remove the part from the building platform. For example, this can be done by sawing. Afterwards, Ultrasonic cleaning or air blasting can be used to remove excess powder. If desired, for large parts, a heat treatment can be carried out to reduce residual stresses. Finally, to improve surface quality, shot peening can be used and to achieve sufficient dimensional accuracy manufacturing methods such as milling are suitable [28–32]. For each post-processing method [z], the *ced* can be divided into the *ced* for the process [s] itself and the *ced* for operating materials [o] such as cutting fluids if necessary (Eq. 7).

$$ced_{post\ processing,z} = \sum_{s=1}^{S} ced_{s,post\ processing,z} + \sum_{o=1}^{O} ced_{o,post\ processing,z} \tag{7}$$

4 Case Study

To implement the evaluation model, Microsoft Excel is used, as it has a feasible usability and wide application in the industry. As a first step, machine-specific parameters, embodied energies of raw materials and operating materials as well as primary energy factors are collected in a specially created database. These data are then linked by the presented equations in Sect. 3 and thereby the model is developed. By entering process- and part-specific data afterwards, the model then calculates the CED for each process step as well as the total CED.

In the case study, a cam disc, shown in Fig. 2 on the left side, is made of aluminum with a volume of 19.78 cm^3 and a mass of 0.14 kg. It is assumed that the aluminum is mined and initially processed into ingots in China. Afterwards, it is shipped from China to Europe by container ship, which corresponds to a distance of about 20,000 km, and then transported 500 km by truck to the point of powder production. The ingots are melted and then atomized by gas atomization with argon. After atomization, the powder is again transported 500 km by truck to the place of DED and post-processing. For the DED process a DMG MORI Lasertec 65 3D is used. The process time at a build-up rate of 72 cm^3/h takes about 16.5 min. Here argon is also used as an inert gas. Powder losses during DED were neglected, as the powder can be reused almost completely without any loss of quality [8]. After DED the part is then removed from the building platform by sawing and is processed by milling with a material removal rate of 5% by a Mori Seiki milling machine.

Table 1. CED calculation

Process step	Sub-process	CED [MJ] sub process	CED [MJ] process step	Ratio	Ref.
Transport	Ship	0.49	0.71	2%	[33]
	Truck	0.23			
Raw material	Extraction	8.19	8.55	20%	[34, 35]
	Casting	0.36			
Atomization	Process	2.23	3.50	8%	[19, 31, 36]
	Inert Gas	1.27			
DED	Laser	12.41	23.84	57%	[19, 36, 37]
	Inert Gas	0.13			
	Trajectory system	0.02			
	Cooling system	10.24			
	Supporting systems	1.02			
Post-processing	Sewing	0.05	5.09	12%	[31, 38]
	Milling	5.04			
Total			**41.69**		

The resulting CED for each considered sub-process as well as the CED summed up for the process steps are given in Table 1. To produce a part with DED, a CED of 41.69 MJ is required. Although the actual DED process has the largest share of CED with 57%, the process steps of raw material production with 20%, post-processing with 12% and powder production with 8% are also significant. Transport has the lowest share of the CED with 2% (see Fig. 2 (right)). These numbers indicate that a sole focus on the DED process itself is not sufficient for the energetic evaluation of a part manufactured with DED. So far there are no sufficient studies of powder production to determine how much energy is required for each sub-process. Future research must therefore analyze the CED of the individual sub-processes and input factors as well as their correlations.

However, each part has its individual CED, which can only be transferred to other parts to a limited extent. However, parallels in correlations and influencing factors of energy can be identified, so that a generally valid evaluation model is possible and necessary for a comprehensive energetic analysis of the process. Therefore, in subsequent works, further machine- and process-specific parameters as well as additional material data are included into the model, to achieve a wider applicability and a greater choice of parameter configurations. For this purpose, further investigations such as a detailed analysis of the powder production, the DED process and post-processing are conducted. In addition, a transfer of the Excel model to Matlab or Python is being striven. The aim is to develop a generally applicable CED-calculation-tool that calculate and predict the CED of a part manufactured by DED and thereby identify the key factors that have a significant impact on the CED.

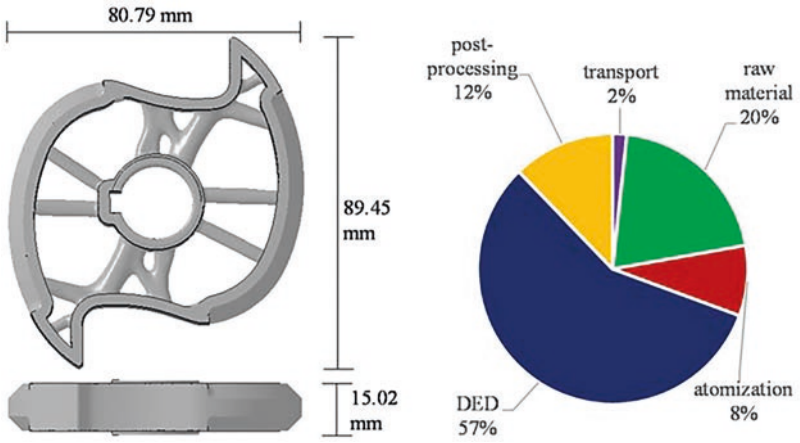


Fig. 2. Part analyzed in the case study (left) and percentage distribution of the CED of the process steps in the total CED (right)

5 Conclusion and Outlook

This paper presents a model to quantify the CED of a part manufactured by DED from cradle to gate. In the presented case study, it becomes evident that a sole consideration of the DED process is not sufficient. In particular, the process steps of powder production need to be investigated more closely. Future work shall initially analyze in detail the processes and individual steps of powder production, DED processing and post-processing in order to allocate and relate the energy-related input factors, within each process step and across the process steps. For example, it is necessary to investigate how different energy-related process conditions during powder production affect the powder and how this powder in turn influences the energy requirements during the AM process. Subsequently, a generally applicable CED-calculation tool should be developed, which allows the simulation and prediction of the CED of a part manufactured by DED.

Acknowledgements. This research was funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 252408385 – IRTG 2057

References

1. ISO: Additive manufacturing - general principles - terminology, Geneva (DIN EN ISO/ASTM 52900) (2015)
2. Campbell, T., Williams, C., Ivanova, O., Garrett, B.: Could 3D Printing Change the World? Technologies, Potential, and Implications of Additive Manufacturing. Strategic Foresight Report. Atlantic Council, Washington (2011)
3. Morrow, W.R., Qi, H., Kim, I., Mazumder, J., Skerlos, S.J.: Environmental aspects of laser-based and conventional tool and die manufacturing. J. Clean. Prod. **15**, 932–943 (2007)
4. Slotwinski, J.A., Garboczi, E.J.: Metrology needs for metal additive manufacturing powders. JOM **67**, 538–543 (2015)
5. Horn, T.J., Harrysson, O.L.A.: Overview of current additive manufacturing technologies and selected applications. Sci. Prog. **95**, 255–282 (2012)
6. Huang, R., Riddle, M., Graziano, D., Warren, J., Das, S., Nimbalkar, S., Cresko, J., Masanet, E.: Energy and emissions saving potential of additive manufacturing: the case of lightweight aircraft components. J. Clean. Prod. **135**, 1559–1570 (2016)
7. Huang, Y., Leu, M.C., Mazumder, J., Donmez, A.: Additive manufacturing: current state, future potential, gaps and needs, and recommendations. J. Manuf. Sci. Eng. **137**, 525 (2015)
8. Clayton, J.: Optimising metal powders for additive manufacturing. Met. Powder Rep. **69**, 14–17 (2014)
9. Tuck, C.J., Hague, R.J.M., Ruffo, M., Ransley, M., Adams, P.: Rapid manufacturing facilitated customization. Int. J. Comput. Integr. Manuf. **21**, 245–258 (2008)
10. Despeisse, M., Ford, S.: The role of additive manufacturing in improving resource efficiency and sustainability, vol. 460, pp. 129–136 (2015)
11. Herzog, D., Seyda, V., Wycisk, E., Emmelmann, C.: Additive manufacturing of metals. Acta Mater. **117**, 371–392 (2016)
12. Klocke, F.: Fertigungsverfahren 5. Gießen, Pulvermetallurgie, Additive Manufacturing, 4th edn. VDI-Buch. Springer Vieweg, Berlin (2015)

13. Liu, Z., Jiang, Q., Cong, W., Li, T., Zhang, H.C.: Comparative study for environmental performances of traditional manufacturing and directed energy deposition processes. Int. J. Environ. Sci. Technol. **15**, 2273–2282 (2018)
14. Priarone, P.C., Ingarao, G., Di Lorenzo, R., Settineri, L.: Influence of material-related aspects of additive and subtractive Ti-6Al-4V manufacturing on energy demand and carbon dioxide emissions. J. Ind. Ecol. **21**, 191–202 (2017)
15. VDI-Zentrum Ressourceneffizienz: Ökologische und ökonomische Bewertung des Ressourcenaufwands. Additive Fertigungsverfahren in der industriellen Produktion (2019)
16. Baumers, M., Duflou, J.R., Flanagan, W., Gutowski, T.G., Kellens, K., Lifset, R.: Charting the environmental dimensions of additive manufacturing and 3D printing. J. Ind. Ecol. **21**, 9–14 (2017)
17. Baumers, M., Tuck, C., Wildmann, R., Ashcroft, I., Hague, R.: Energy inputs to additive manufacturing: does capacity utilization matter. Eos **1000**, 30–40 (2011)
18. Fredriksson, C.: Sustainability of metal powder additive manufacturing. Procedia Manuf. **33**, 139–144 (2019)
19. Le Bourhis, F., Kerbrat, O., Hascoet, J.Y., Mognol, P.: Sustainable manufacturing: evaluation and modeling of environmental impacts in additive manufacturing. Int. J. Adv. Manuf. Technol. **69**, 1927–1939 (2013)
20. Bambach, M.D., Bambach, M., Sviridov, A., Weiss, S.: New process chains involving additive manufacturing and metal forming – a chance for saving energy? Procedia Eng. **207**, 1176–1181 (2017)
21. Yi, L., Glatt, M., Sridhar, P., de Payrebrune, K., Linke, B.S., Ravani, B., Aurich, J.C.: An eco-design for additive manufacturing framework based on energy performance assessment. Addit. Manuf. **33**, 101120 (2020)
22. Verband Deutscher Ingenieure (VDI): Cumulative energy demand (KEA) - Terms, definitions, methods of calculation (VDI 4600) (2012)
23. Dawes, J., Bowerman, R., Trepleton, R.: Introduction to the additive manufacturing powder metallurgy supply chain. Johnson Matthey Technol. Rev. **59**, 243–256 (2015)
24. Hoeges, S., Zwiren, A., Schade, C.: Additive manufacturing using water atomized steel powders. Met. Powder Rep. **72**, 111–117 (2017)
25. Boulos, M.: Plasma power can make better powders. Met. Powder Rep. **59**, 16–21 (2004)
26. 2Schatt, W.: Pulvermetallurgie. Technologien und Werkstoffe, 2nd edn. VDI-Buch. Springer, Berlin (2007)
27. Gu, D.D., Meiners, W., Wissenbach, K., Poprawe, R.: Laser additive manufacturing of metallic components: materials, processes and mechanisms. Int. Mater. Rev. **57**, 133–164 (2012)
28. Kumbhar, N.N., Mulay, A.V.: Post processing methods used to improve surface finish of products which are manufactured by additive manufacturing technologies: a review. J. Inst. Eng. (India): Ser. C 99, 481–487 (2018)
29. Kellens, K., Mertens, R., Paraskevas, D., Dewulf, W., Duflou, J.R.: Environmental impact of additive manufacturing processes: does am contribute to a more sustainable way of part manufacturing? Procedia CIRP **61**, 582–587 (2017)
30. Qian, M., Xu, W., Brandt, M., Tang, H.P.: Additive manufacturing and postprocessing of Ti-6Al-4V for superior mechanical properties. MRS Bull. **41**, 775–784 (2016)
31. Faludi, J., Baumers, M., Maskery, I., Hague, R.: Environmental impacts of selective laser melting: do printer, powder, or power dominate? J. Ind. Ecol. **21**, 144–156 (2017)
32. Choi, H., Byun, J.M., Lee, W., Bang, S.R., Kim, Y.D.: Research trend of additive manufacturing technology. J. Korean Powder Metallur. Inst. **23**, 149–169 (2016)
33. Ashby, M.F.: Materials and the Environment. Eco-informed Material Choice. Butterworth-Heinemann, Waltham (2012)

34. Umweltbundesamt (Probas): Prozessdetails: Metall Aluminiumbarren-Welt-2005. https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={0CF22677-9DBF-40D9-8A17-9A9BB9D03A4F}. Accessed 18 June 2020
35. Umweltbundesamt (Probas): Prozessdetails: Metall Aluminium-AU-2020. https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={760B714B-BBEE-4185-B1AB-75798F103911} (2020)
36. Umweltbundesamt (Probas): Prozessdetails: Xtra-generischArgon-DE-2005. https://www.probas.umweltbundesamt.de/php/prozessdetails.php?id={07D594A0-3BD2-444C-BF10-6D247E624169} (2020)
37. Wippermann, A., Gutowski, T.G., Denkena, B., Dittrich, M.A., Wessarges, Y.: Electrical energy and material efficiency analysis of machining, additive and hybrid manufacturing. J. Cleaner Prod. **251**, 119731 (2020)
38. Kara, S., Li, W.: Unit process energy consumption models for material removal processes. CIRP Ann. **60**, 37–40 (2011)

Building Blocks for Digitally Integrated Process Chains in PBF-Based Additive Manufacturing

M. Sjarov[1(✉)], N. Ceriani[2], T. Lechler[1], and J. Franke[1]

[1] Institute for Factory Automation and Production Systems, Friedrich-Alexander-University Erlangen-Nuremberg, Egerlandstr. 7-9, 91058 Erlangen, Germany
martin.sjarov@faps.fau.de
[2] Siemens AG, Gleiwitzer Str. 555, 90475 Nürnberg, Germany

Abstract. As applications of Additive Manufacturing (AM) extend from rapid prototyping into series production, new challenges of digital information integration arise. In order to obtain transparency of overall costs and part quality, the relevant information of each process step needs to be gathered from different sources and combined in a meaningful way. Thus, a digitally integrated process chain can provide added value for manufacturers and their clients.

This paper proposes building blocks for digitally integrated process chains in the context of powder bed fusion (PBF) based AM to form a basis for transparency of cost and quality. Solving requirements for the end-to-end digital chain leads into challenges of acquiring, storing, processing and routing information. Related technologies comprise domain-specific software, enterprise integration patterns, database and semantic technologies. The mapping of these enablers onto the AM chain leads to a target architecture that forms the fundament for future investigations.

Keywords: Additive manufacturing · Digital process chain · Application integration · Industry 4.0

1 Introduction

Additive Manufacturing (AM) comprises a set of manufacturing processes with distinct properties: In contrast to traditional manufacturing, no tooling is required as the final parts are directly manufactured by joining material layer by layer. At that, the AM machine consumes digital build job data in order to realize a desired 3D geometry. This direct manner of part generation allows the realization of complex geometric shapes as well as the integration of functions directly into the part. Thus, additively manufactured parts are able to cover a wide range of applications [1, 2].

The group of AM processes is standardized within the ISO/ASTM 52900 [3]. Among the AM processes, powder bed fusion (PBF) is suitable to produce parts from

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 368–377, 2021.
https://doi.org/10.1007/978-3-662-62138-7_37

polymers as well as different metals. Especially for the additive production of metal parts, PBF is by far the most common type of process [4]. The markets for metal AM machines as well as material supply have been characterized by steady growth, each marking peaks in terms of latest surveys [5].

The process chain of AM in general and metal PBF in particular can be divided into three sections: Pre-processing, In-processing and Post-processing [6]. Each of these sections contains numerous sub-steps, linked with properties and parameters influencing the final part quality as well as being associated with certain costs and value added. In order to gain transparency among these parameters, their relations and interdependencies, a digital integration mechanism is required. Its targets comprise integration on different levels, including applications, data and knowledge. The outline of the required building blocks i.e. foundational knowledge, tools and methods to design and implement a digitally integrated process chain will be subject of the ensuing paper, which is therefore structured as follows.

In the next section, considerations of a standardized process chain for metal PBF are presented. Based on related work, a structured view on process steps is provided. It is the basis for the derivation of further sub-steps, use cases, interfaces and data models. Section 3 covers AM-associated software tools. As they are numerous and equipped with heterogeneous interfaces, a need for integration can be derived. Thus, Sect. 4 analyzes digital integration in the context of AM from different perspectives: Applications, data and knowledge. Section 5 combines the findings into the concept of a digitally integrated AM process chain, serving as a blueprint for future implementation.

2 Metal PBF AM Process Chain

2.1 Related Work

Different scientific and normative work forms the basis to formalize the PBF process chain for further considerations. An introductory work covering AM processes is given by Gebhardt et al. [2]. Gibson et al. [7], again introductory, list an 8-step generic AM process. VDI 3405 provides an overview of various AM process technologies and related requirements, separating the process chain into pre-processing, in-processing and post-processing [6]. A more detailed view, using the VDI top-level grouping, is presented by Möhrle [8]: Based on 10 sources, a super set of PBF process steps is provided. A review on relevant process parameters is to be found in [9]. A recent working paper, still being under further development and covering the entire PBF process chain, is the automation roadmap AM of the VDMA [10].

Based on this related work as well as ongoing studies within the associated research project (see acknowledgements), a consolidated PBF process chain can be derived, that is briefly presented in this section. It is, at the top level, divided into

1. the primary process chain, comprising further the pre-, in- and post-processing and all process steps that add value to the final product, and

2. the secondary process chain, comprising the powder and substrate cycle as well as machine preparation and maintenance, thus not directly adding value to the final product but being of a supplemental nature.

The process chains themselves may be viewed and modeled on different levels of depth, according to the requirements of specific use cases. Table 1 exemplarily lists process steps that were detailed until a depth level of five (L5). This hierarchical analysis and structuring is a prerequisite for the digital modeling and representation.

Table 1. Structuring of exemplary AM process steps into different, hierarchical levels of detail according to the generalized PBF process chain

Level 0	Level 1	Level 2	Level 3	Level 4	Level 5
AM process	Primary chain	Pre-processing	Design	Direct CAD	Design part
AM process	Primary chain	Pre-processing	Design	Reverse engineer	From point cloud
…	…	…	…	…	…
AM process	Primary chain	Post-processing	Build level	Depowdering	Manual Depowdering
…	…	…	…	…	…
AM Process	Secondary Chain	Powder Cycle	-	-	Mix Powder

2.2 Primary Process Chain

The primary process chain (L1) comprises process steps that mainly contribute to the value added to the final product. It starts with the **Pre-Processing** (L2) which itself is to be subdivided into the following L3 and L4 steps, each having complex sub-steps as well as sequences, including iterations:

- Part design: Direct CAD (Computer Aided Design), reverse engineering
- Enrich CAD model: Design features, serialization codes
- Simulation: Structural strength, warpage
- CAM (Computer Aided Manufacturing): Apply offsets, file conversions, data preparation and data set fixing, order mixing, part arrangement
- Support generation: Automatic, semi-automatic, manual
- Slicing, hatching: Process parameters, built simulations

The pre-processing ends with the generation of corresponding print job data. Then, the **in-processing** (L2) follows, which is to be divided into initiation, the build process itself as well as the cool down and build chamber removal.

The subsequent **post-processing** (L2) is to be divided (L3) into steps concerning the built in its entirety – comprising depowdering, heat treatment and parts separation – as well as steps operating at the level of single parts: Remove support, dispose waste, hot isostatic pressing, blasting, machining, improvement of properties and testing (destructive, non-destructive). The primary chain ends with the quality assurance of the produced part, which is then eligible for dispatching.

2.3 Secondary Process Chain

The secondary chain (L1) comprises on L2 the substrate cycle, the powder cycle as well as machine-related activities. In contrast to the L2 steps of the primary chain, the required level of detail and therefore depth of hierarchy for the secondary chain is reduced. Both substrate and powder management are essential for the building of a part and share the common process steps of check-in, installation at the machine and disposal of waste material. Concerning the powder, the mixing, sieving, splitting and qualification of batches is of particular importance. The substrate requires measurement, machining and blasting before being installed into the machine.

In conclusion, the identification and detailed modeling of generic process steps is a major requirement and therefore building block of the digitally integrated chain. As the exhaustive listing and modeling of process steps would exceed the scope of this paper, it is the subject of future research and publications.

3 Additive Manufacturing and Software Tools

The tool chain of AM – especially in the pre-processing – depends on specialized software components, each having their focus on different steps and sections, some covering a more narrow part and others being of an overarching nature (see Fig. 1). The AM process chain starts with the part design, which is enabled by CAD software. It is complemented by and iteratively used with CAE tools e.g. simulation software, where one area of interest is the part deformation due to process-induced heat distributions. In order to generate print jobs, CAM tools are required. The implemented tool chains are manifold regarding their possible sequences and iterations. Furthermore, they depend on specific functionality of the used software e.g. the possibility of integrated operation within one tool or the need to export and import data between applications. Regarding the in-processing, machine operation is guided by MES (Manufacturing Execution System) and SCADA (Supervisory Control and Data Acquisition) systems. Furthermore, complex data sets generated during a print job are oftentimes stored internally on the machines themselves and accessible via machine interfaces.

Software tools acting as a backbone for the process chain comprise Product Data Management (PDM) and Enterprise Resource Planning (ERP). Each of the mentioned tools enables certain process steps, collecting necessary data. Thus, in order to gain AM-specific knowledge covering an end-to-end perspective including costs, time and quality, an integrating mechanism is required.

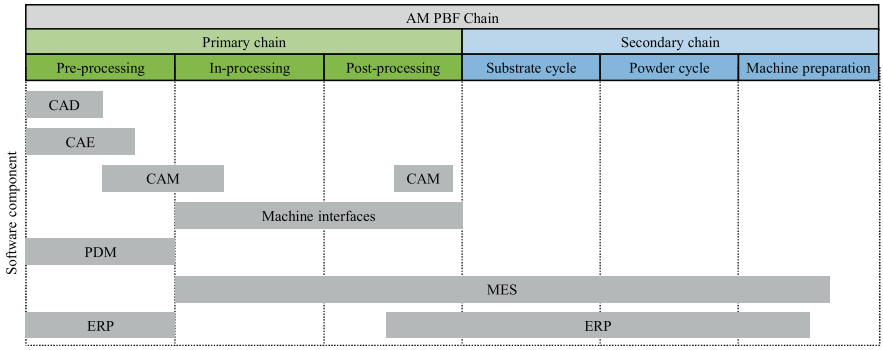


Fig. 1. Software components and their main regions of concern within the AM process chain

4 Additive Manufacturing and Digital Integration

As shown in the previous section, different types of applications as well as resulting data sets are to be integrated, when it comes to an integrated AM chain. Thus, this section provides building blocks targeting different levels of digital integration.

4.1 Application Integration

When digital application integration is required, several integration styles can be differentiated. According to Hohpe and Woolf these comprise at least file transfer, shared databases, remote procedure invocation and messaging. The stated order of these integration styles implies an order of increasing complexity when it comes to the realization of the concepts, from file transfer being the simplest one to messaging as a more sophisticated method [11].

Despite messaging being the most complex of the four stated integration approaches, it holds advantageous properties: The data exchange between the sending and receiving applications is asynchronous, avoiding the blocking that might occur when remote procedure calls would be utilized. This decouples sender and receiver, which is an important property for distributed systems. Furthermore, the messaging mechanism, often referred to as Message Oriented Middleware (MOM) in this context, can transform the message to comply with specific data formats [12].

Multiple implications for the AM chain ensue: Depending on the application or system containing the data (ref. Sect. 3) a suitable integration style needs to be applied. Due to the mentioned advantages, a central integration mechanism should be MOM-based. File-based approaches can complement this in the pre-processing, as it mostly relies on the generation and exchange of particular files.

4.2 Data Integration

Scheideler and Villmer investigated requirements of integrated AM chains focusing on different data formats and their suitability [13]. They concluded that consistent

data models are a major prerequisite for an integrated end-to-end chain. Likewise, Dedenka et al. stated that the standardization and enrichment of data models is of particular importance for standardized digital AM chains [14].

Complementary to file-based perspectives on AM data, different database technologies need to be taken into account. Besides traditional relational database systems, document-oriented and graph databases fit the accruing AM data: Document-oriented databases are capable of storing heterogeneously structured data that e.g. occurs within the AM pre-processing. As they do not depend on fixed schemas, a gradual realization is facilitated. Graph databases focus extensively on the relations between entities, therefore being suitable to e.g. store the process history of an AM part including iterations. Furthermore, graph databases are closely related to knowledge representation (see Sect. 4.3).

4.3 Knowledge Integration

Before knowledge in general and AM knowledge in particular can be integrated, i.e. productively connected across process steps, it needs to be formalized to be machine-readable. Ontologies represent one possibility to accomplish this [15]. Recent work of knowledge representation using ontologies in the context of AM covers the design for AM [16], the AM product life cycle [17] and AM in general [18]. Concerning the integration of AM knowledge, Saal outlined the definition of design features and their mapping onto part quality data [19]. To finally utilize the formalized knowledge, inferencing resp. reasoning comes into play.

When harnessing ontologies for knowledge integration with the end-to-end AM chain, the combination of a neutral top-level ontology together with one or more domain-specific ontologies holds advantages over constructing a custom ontology from scratch. Furthermore, as the design and use of ontologies is an open-end process, a small initial setup can serve as the basis to which future extensions can be added [20].

5 Digitally Integrated AM Process Chain

This section summarizes and combines aforementioned facets of a digitally integrated AM process chain, resulting in a blueprint of a software-based integration mechanism (IM) as the proposed solution. Being software-intensive, the 4 + 1 view model of Kruchten [21] is applicable to describe it. In the following, a physical view, a logical view as well as two use cases are outlined as a subset of this model. Furthermore, a method to realize the desired solution is proposed.

5.1 Physical View: Target Architecture

Figure 2 provides a physical view on the AM IM, which is the central component connecting both existing applications and machines. To realize this, the AM IM provides a multitude of different interfaces in order to consume incoming data. Internally, the

data needs to be harmonized, routed and stored accordingly. The enterprise integration patterns (ref. Sect. 4.1) provide a formal basis for tackling this [11]. To integrate existing enterprise software products, each needs to be assessed regarding the provided interfaces. Then, dedicated adaptors connect each component. The same holds true for machines to couple.

The acquired heterogeneous data is stored in corresponding database types (ref. Sect. 4.2). A common AM ontology provides the basis for AM knowledge integration and knowledge-based reasoning. This implies, that incoming data is contextualized, corresponding to the current state of the underlying ontology.

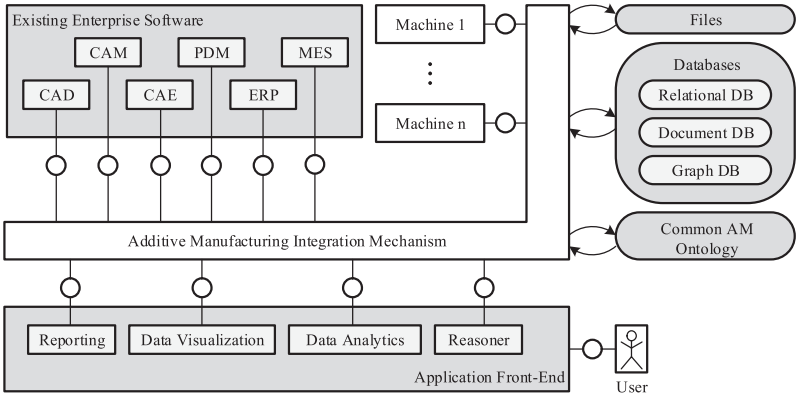


Fig. 2. Target architecture of the AM integration mechanism and its surroundings

To provide the added value for the end user, dedicated services with corresponding interfaces connect to the AM IM in order to visualize and analyze available data and combine it with AM knowledge. As the system is inherently distributed, it is implied that the user interaction can take place at any location where access, predominantly over the internet, is possible.

5.2 Logical View

In the context of the AM IM, the logical view comprises the end user-related services as well as design requirements for databases and ontologies. Basis for the data modelling is the generalized process chain (ref. Sect. 2) represented in the form of generic process steps (ref. Fig. 3) that can be related to each other according to an actual part-specific process chain. Each generic process step contains a set of associated input and output parameters. Further, it defines how cost and added value are generated and which internal and external parameters potentially influence the resulting part quality. A common AM ontology defines and relates involved entities, their properties and interdependencies. It is dynamic and needs to be enriched over time according to the identified needs.

Basic end user services comprise the visualization of actual part-specific and order-specific process chains together with the associated data sets e.g. actual versus planned time and cost. Advanced services allow for data analytics and artificial intelligence-based approaches to correlate arbitrary data with target properties as e.g. quality-relevant part design features. To further specify the needed services, use cases can provide guidance.

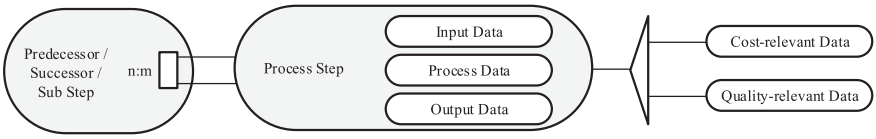


Fig. 3. A generic process step with its associated data, noted in FMC

5.3 Use Cases

To further motivate for the AM IM, two exemplary use cases (UC1, UC2) are outlined, which originated within the context of the research project (see acknowledgements) but are still to be implemented. Both rely on the outlined architecture and leverage the data of an integrated process chain.

UC1, part history: After a part passed the quality assurance, a quality engineer (QE) wants to track its process history and powder genealogy. Dedicated services provide, beginning with the pre-processing, a complete graph of process steps the part passed through, including iterations and idle times. Each process step can be investigated for relevant cost and quality data as well as further meta data. As it is a customer requirement for this certain part, the QE provides a detailed report as an attachment to the delivery.

UC2, price forecast: A customer requests an offer for the production of a given part. The sales engineer (SE) utilizes a service offered by the AM IM to search for similar historic part geometries and their resulting overall production times, costs and potentially relevant quality issues. This implies the identification and semantic mapping of critical part geometries as well as parameters that are mainly relevant for the resulting cost. Using the gathered information on costs, lead times, and expectable quality, the SE sends a suitable offer to the customer.

Additional use cases envisioned by means of expert workshops within the research project comprise the analysis of machine and equipment utilization, knowledge and feature-based analysis of part-specific pre-processing decisions and their consequences for the final part quality.

5.4 Realization Method

In order to transform the building blocks into a functioning AM IM, the following steps are proposed as guidelines:

1. Conceptualization of a general AM process chain containing all possible process steps with their associated data sets, cost and quality items
2. Definition of use cases that utilize information of a connected AM chain and derivation of corresponding service requirements
3. Customization of the AM IM architecture according to the specific needs
4. Definition and interface analysis of software products and machines to connect
5. Commitment to specific software products implementing the different components of the AM IM (databases, messaging, interfaces)
6. Implementation of a smallest possible executable system
7. Stage-by-stage expansion and refinement of the system

As the AM IM is a piece of software, established software development methods can be applied. Especially for the step-by-step expansion, a Scrum-like proceeding is advisable.

6 Conclusion and Outlook

This paper presents building blocks for digitally integrated AM process chains. A conceptualized PBF process chain is the basis for the identification of generic process steps, involved software and hardware components as well as resulting interfaces. In order to gain insight into the physical process chain, e.g. for tracing part costs or make statements regarding the resulting part quality, the digital acquisition, storage and correlation of different heterogeneous data is required. Therefore, mechanisms for different aspects of digital integration are introduced. A central integrator is message-oriented middleware in combination with different database and knowledge management technologies. This leads to the AM IM software component, which is then characterized using a subset of Kruchten's 4 + 1 views.

For future work, next steps can be derived from the further proposed method for implementation. As the process chain is fully conceptualized and first use cases are defined, now the selection and commitment to a suitable software stack using state-of-the-art technology needs to be made. Then, as the origin for further gradual expansion, a minimal functioning unit of the AM IM is to be implemented.

Acknowledgements. This paper emerged within the scope of the publically funded research project "IDEA". Funding agency: BMBF. Funding identifier: 13N15003.

References

1. DebRoy, T., Wei, H.L., Zuback, J.S., Mukherjee, T., Elmer, J.W., Milewski, J.O., Beese, A.M., Wilson-Heid, A., De, A., Zhang, W.: Additive manufacturing of metallic components – process, structure and properties. Prog. Mater. Sci. **92**, 112–224 (2018). https://doi.org/10.1016/j.pmatsci.2017.10.001
2. Gebhardt, A., Kessler, J., Thurn, L.: 3D Printing. Understanding Additive Manufacturing, 2nd edn. Carl Hanser, Munich (2019)

3. DIN EN ISO/ASTM 52900:2017-06: Additive Fertigung - Grundlagen - Terminologie (ISO/ASTM 52900:2015); Deutsche Fassung EN_ISO/ASTM 52900:2017. Beuth, Berlin
4. Zeyn, H.: Industrialisierung der Additiven Fertigung. Digitalisierte Prozesskette - von der Entwicklung bis zum einsetzbaren Artikel, 1st edn. Industrie 4.0. Beuth & VDE, Berlin (2017)
5. Associates, W.: Wohlers Report 2019. 3D Printing and Additive Manufacturing State of the Industry. Wohlers Associates, Fort Collins (2019)
6. VDI 3405:2014-12 Additive manufacturing processes, rapid manufacturing - Basics, definitions, processes. Beuth, Berlin
7. Gibson, I., Rosen, D., Stucker, B.: Additive Manufacturing Technologies. Springer, New York (2015)
8. Möhrle, M.: Gestaltung von Fabrikstrukturen für die additive Fertigung. Springer, Berlin (2018)
9. Stavropoulos, P., Foteinopoulos, P.: Modelling of additive manufacturing processes: a review and classification. Manuf. Rev. **5**, 2 (2018). https://doi.org/10.1051/mfreview/2017014
10. VDMA AG AM: Roadmap Automatisierung AM (2018)
11. Hohpe, G., Woolf, B.: Enterprise Integration Patterns. Designing, Building and Deploying Messaging Solutions, 19th edn. The Addison-Wesley Signature Series. Addison, Boston (2015)
12. Razzaque, M.A., Milojevic-Jevric, M., Palade, A., Clarke, S.: Middleware for internet of things: A survey. IEEE Internet Things J. **3**(1), 70–95 (2016). https://doi.org/10.1109/JIOT.2015.2498900
13. Scheideler, E.M., Villmer, F.J.: Anforderungen an integrierte Prozessketten in der Additiven Fertigung. In: Kynast, M., Eichmann, M., Witt, G. (eds.) Rapid.Tech – International Trade Show & Conference for Additive Manufacturing, pp. 10–24. Carl Hanser, München (2017)
14. Denkena, B., Dittrich, M.A., Henning, S., Lindecke, P.: Investigations on a standardized process chain and support structure related rework procedures of SLM manufactured components. Procedia Manuf. **18**, 50–57 (2018). https://doi.org/10.1016/j.promfg.2018.11.007
15. Qi, Q., Pagani, L., Scott, P.J., Jiang, X.: A categorical framework for formalising knowledge in additive manufacturing. Procedia CIRP **75**, 87–91 (2018). https://doi.org/10.1016/j.procir.2018.04.076
16. Kim, S., Rosen, D.W., Witherell, P., Ko, H.: A design for additive manufacturing ontology to support manufacturability analysis. J. Comput. Inf. Sci. Eng. **19**(4), 1107 (2019). https://doi.org/10.1115/1.4043531
17. Mohd Ali, M., Rai, R., Otte, J.N., Smith, B.: A product life cycle ontology for additive manufacturing. Comput. Ind. **105**, 191–203 (2019). https://doi.org/10.1016/j.compind.2018.12.007
18. Sanfilippo, E.M., Belkadi, F., Bernard, A.: Ontology-based knowledge representation for additive manufacturing. Comput. Ind. **109**, 182–194 (2019). https://doi.org/10.1016/j.compind.2019.03.006
19. Saal, C.: Durchgängige Feature-ID von der 3D-CAD Konstruktion bis zur Qualitätssicherung. In: DFX 2019: Proceedings of the 30th Symposium Design for X, 18–19 September 2019, Jesteburg, Germany. 30th Symposium Design for X, 18th–19th September 2019. The Design Society (2019). https://doi.org/10.35199/dfx2019.19
20. Arp, R., Smith, B.D., Spear, A.D.: Building Ontologies with Basic Formal Ontology. MIT Press, Cambridge (2015)
21. Kruchten, P.B.: The 4+1 view model of architecture. IEEE Softw. **12**(6), 42–50 (1995). https://doi.org/10.1109/52.469759

Correlation of Spatter Quantity and Speed to Process Conditions in Laser Powder Bed Fusion of Metals

E. Eschner[1,2(✉)], K. Schwarzkopf[1,2], T. Staudt[1,2], and M. Schmidt[1,2]

[1] Institute of Photonic Technologies, Friedrich-Alexander-Universität Erlangen-Nürnberg, Konrad-Zuse-Str. 3/5, 91052 Erlangen, Germany
eric.eschner@fau.de
[2] Erlangen Graduate School in Advanced Optical Technologies (SAOT), Friedrich-Alexander-Universität Erlangen-Nürnberg, Paul-Gordan-Str. 6, 91052 Erlangen, Germany

Abstract. The results of spatter measurements within laser-based powder bed fusion of metals are presented. A stereoscopic imaging setup and corresponding reconstruction algorithm are used to determine three-dimensional measures of the spatter characteristics from experiments with 1.4404 stainless steel. The spatter characteristics are correlated to the process zone morphology and evaporation behavior. The evolution of the spatter count over consecutive tracks is investigated and shows a decrease and convergence towards a constant value. Experiments with the process gases argon, helium and nitrogen reveal that the spatter count decreases with the gas density, whereas the spatter speed stays unaffected. This confirms the key role of the process gas in the entrainment of powder particles and the associated spatter generation. The results indicate that macroscopic spatter characteristics contain relevant information about microscopic process behavior. This makes spatter characteristics a designated process feature for new sensing approaches for the observation of industrial applications.

Keywords: Additive manufacturing · Stereoscopic imaging · Powder bed fusion of metals

1 Introduction

The process of laser powder bed fusion of metal (PBF-LB/M, acronym according to DIN EN ISO/ASTM 52900) is emerging to become an essential part of the modern production technology portfolio. First industrial applications show their applicability within a cost effective production chain. Due to its advantages, e.g. freedom of design, realization of light-weight constructions and tooling independent of the specific part, PBF-LB/M enables the production of highly specialized and complex parts.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 378–386, 2021.
https://doi.org/10.1007/978-3-662-62138-7_38

However, to this date the technology still has to overcome some obstacles regarding quality assurance and process control in order to be applied in a wide spectrum and number of industrial applications. Due to multifactorial process parametrization, a generalized understanding of the complex defect mechanisms is challenging but necessary to identify and prevent alteration of part quality. Consequently, sensor data is a key element to evaluate and ensure a constant process result. Therefore, multiple approaches are under investigation, potentially allowing for the detection and ideally the avoidance of process irregularities. This work focuses on the correlation of three-dimensional (3D) spatter trajectories to the changes in process parameters and the vicinity of the process zone.

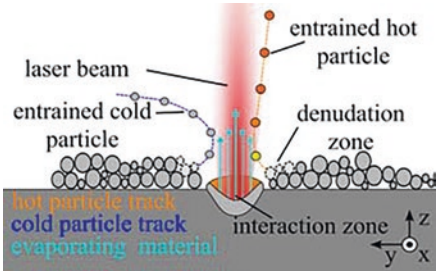


Fig. 1. Illustration of the process zone and the spatter generation caused by particle entrainment from the denudation zone.

The main cause of spatter in PBF-LB/M is entrainment of powder towards the process zone, which is illustrated in Fig. 1 [1]. To understand the underlying mechanisms, one has to understand the process zone evolution. Within PBF-LB/M a laser-beam is irradiated onto a powder bed, with a layer height within the magnitude of the laser spot size. The laser light is absorbed and transferred into heat which leads to melting of the powder and a temperature increase in the surrounding material. This causes the formation of a molten pool. If the incident laser intensity exceeds a material dependent threshold intensity, a pronounced vapor plume forms and the vapor speed caused by evaporation increases linearly with power [2]. The recoil pressure of the evaporating material causes the formation of a depression, respectively an onset of a keyhole, on the melt pool surface [3]. The increasing evaporation pressure leads to a transition of a forward directed vapor plume, which points away from the already processed track, to a backward directed vapor plume (compare Fig. 4b) [2]. Due to viscosity, momentum of the fluid flow within the vapor plume is transferred into the surrounding atmosphere. This induces a secondary fluid flow of the atmospheric gas directed towards the process zone, leading to the erosion of the powder bed and the entrainment of powder particles. The depletion of powder around the process zone leads to the formation of the so called denudation zone [4]. If the particles interact with the vapor plume, they get molten and expelled, forming spatter particles. Yet, not all entrained particles are transformed to spatter, therefore redistribution to other parts of the powder bed or part surface may occur. However, solidified spatter particles, can lead to defects later on in the process [5]. While the entrainment of particles is the main cause for spatter formation [6], spatter can also be generated by ejection of molten material from the process zone [1].

Since the evaporation driven spatter generation is the dominating mechanism, spatter characteristics are physically linked to evaporation behavior and therefore allow to obtain insights into microscopic process zone features, as we have shown in previous work [7] and as the results in [8] imply. Spatter measurements within highly magnified process zone observations using a single camera have proven to provide better insight and extend general process understanding [1]. But neither do they allow the reconstruction of 3D particle trajectories, since the projection onto the camera results in a projection error within the measurements, nor are they applicable in a conventional industrial PBF-LB/M process which would necessitate the imaging over the whole build plate. However, it has been shown, that 3D spatter information can be obtained with low cost hardware within industrial machinery, without the need to optically resolve the spatter particles [9].

Within this work we use a stereoscopic high-speed camera setup to image spatter particles. A specialized algorithm is applied to reconstruct their 3D trajectories. The acquired data provides insight into (i) the quantitative evolution of spatter count over multiple tracks and (ii) the dependence of spatter speed on process conditions. The understanding of the correlation between certain spatter properties and the present process zone conditions is a key element of a process knowledge-based sensing approach. The results can enhance the understanding of in-situ spatter measurements of the PBF-LB/M process and their correlation to process errors.

2 Materials and Methods

2.1 Experimental Setup

The experimental setup is shown in Fig. 2. A two-axis stage system is used to move the sample relative to the laser beam. It comprises a fast axis along x which can move the sample with up to several m/s (Aerotech act115Dl-1000) and a slower axis along

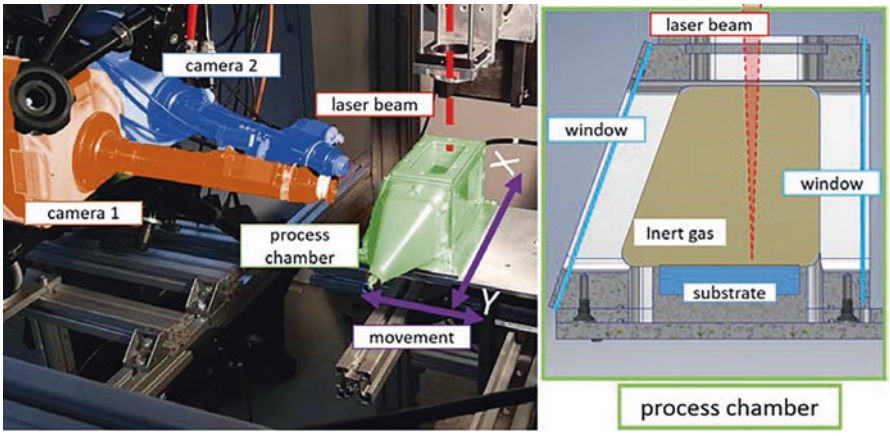


Fig. 2. Experimental setup (left) and cross-section of the used process chamber with inert gas atmosphere (right).

y to move the sample perpendicular to the laser scanning direction, allowing for the realization of consecutive tracks with defined hatch spacing (PI M-605.2DD). The coordinate system used in this paper is chosen to be parallel to the laser scanning direction in x and the laser beam along z. The laser beam source, a Trumpf TruDisk 6001, is used in combination with a 100 µm fiber and self-build optics (magnification 1:1, focal length 200 mm). The focal point of the laser beam is set to lie on the surface of the specimen, resulting in an spot diameter of d_{spot} of 100 µm and a top hat beam profile. The specimen is placed inside a process chamber, which is flooded with inert gas (Fig. 2b). The chamber has two glass windows for process observation and a transparent top window with an antireflective coating for the laser wavelength.

For the experiments, 2 mm thick 1.4404 stainless steel plates were sandblasted and a 70 µm powder layer of 1.4404 powder was manually spread onto it using distance gages and a scraper. The powder particle size ranges between 20 and 53 µm. Measurements with a laser scanning microscope validated good reproducibility of a consistent layer height. For the experiments five consecutive tracks with a hatch spacing of 80 µm are produced for each set of parameters. Each set of tracks is produced five times resulting in a sample size of five. If not explicitly stated differently, we used argon as inert gas for the experiments presented within this paper.

2.2 Computation of Statistical Spatter Properties

The process is imaged by a stereoscopic camera-system comprised of two high-speed cameras (Vision Research Phantom v1210) with the resolution set to 512×256 pixels and a recording frequency of 60,000 fps. No illumination is used, therefore only parts of the process zone with significant thermal emission, respectively the interaction zone and spatter particles are imaged. The stereoscopic camera-system is calibrated in advance of each set of experiments. Within the acquired images potential spatter particles are identified and 3D positions are reconstructed by particle matching within the two camera images. By integration of a priori knowledge into a tracking algorithm, the spatter paths are detected. An in-depth description of the applied methodology is provided in our previous work [10]. The applied algorithm delivers the 3D track of each detected individual spatter particle, enabling the determination of its e.g. statistical spatter speed and direction over multiple consecutive frames. For the purpose of comparability, the tracks of the individual spatter particles are further processed to compute statistical spatter measures for each experiment. Since each spatter is tracked individually the measurement is not altered by multiple detections of the same spatter particle in consecutive frames. Within this work we use the spatter count $n_{spatter}$ to describe the overall number of detected individual spatter particles within a single track. Furthermore, the median absolute speed $v_{spatter}$ of the spatter particles is computed. The median is chosen over the mean value since the distribution of the absolute speed cannot be normally distributed but has to be rather skewed, which is confirmed by our measurements and exemplarily shown in Fig. 3b. For each set of samples, the standard deviation and mean value is computed. The intermediate processing results – detected particles, reconstructed spatter tracks and speed histogram are visualized in Fig. 3 for an exemplary data set.

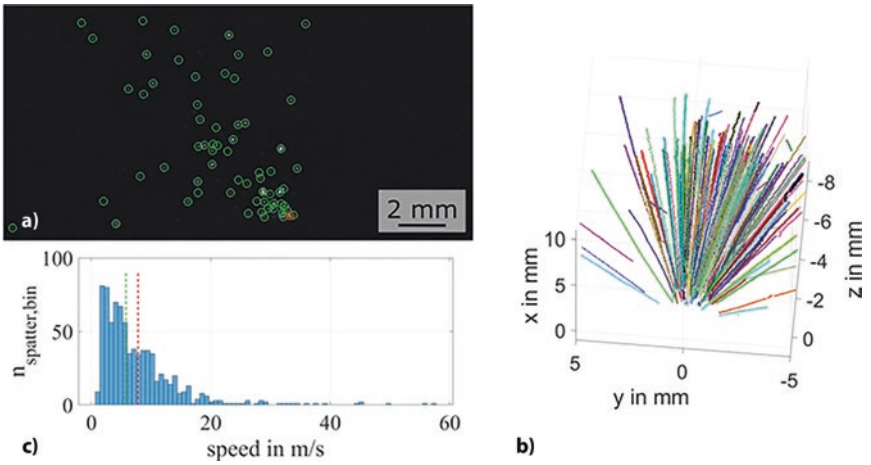


Fig. 3. Exemplary data set showing (a) raw image with detected potential spatter particles (green) above the process zone (red), (b) the reconstructed spatter tracks, with each individual spatter particle marked in one color and (c) the spatter speed histogram with the respective mean (red) and median (green) spatter speed.

3 Results and Discussion

The process behavior observable within the chosen range of experiments can be associated with two distinct process regimes. Lower intensities cause a process behavior in which the penetration depth can be mainly attributed to heat conduction into the melt pool, called heat conduction regime (HCR) in the following. In contrast, an increased intensity leads to evaporation which causes a depression at the melt pool surface. This leads to an increased penetration depth of the molten pool, as convective flows increase and the position of the absorbing melt pool surface relative to the specimen's surface changes. This characteristic process behavior is called evaporation regime (EVR) in the following. In order to make the experiments comparable for several feed velocities v_{feed}, the power is scaled in accordance with a respective constant line energy l_e, which leads to a constant energy input per track:

$$l_e = \frac{P}{v_{feed}} \tag{1}$$

3.1 Transfer of Vapor Momentum to Spatter Particles

Figure 4 shows the evolution of the spatter speed with an increasing line energy, respectively power, for three different feed speeds v_{feed}. Within the chosen parameter range the process zone shifts from HCR to EVR [7]. The onset of evaporation is evident by a shift from a constant spatter speed to a more or less linear increase at roughly 450 W. This shift is also present in high speed process observations of the process zone (Fig. 5), in which a change from a steady molten pool, to the formation

of a depression and introduction of dynamic fluctuations on the molten pool surface can be observed.

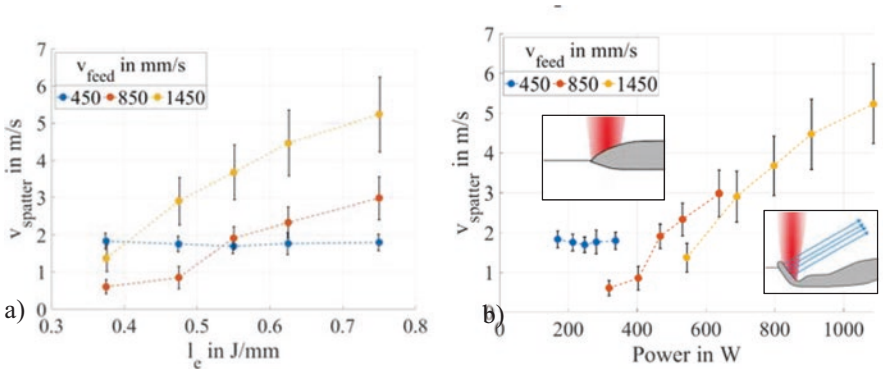


Fig. 4. Spatter speed $v_{spatter}$ over line energy l_e (a) and power (b), showing a shift from a heat conduction driven process regime at lower feed velocities to a evaporation driven process regime at higher feed velocities (n = 25).

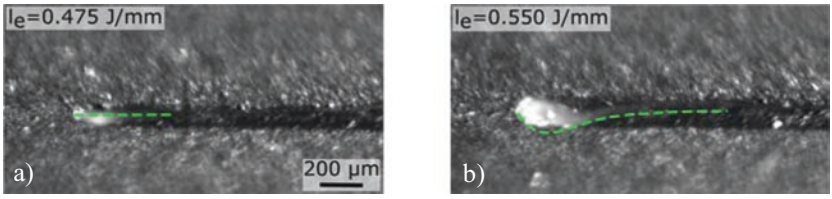


Fig. 5. Observed melt pool topography (green dashed line) on a bare plate sample of 1.4404 within the heat conduction regime at 0.475 J/mm (a) and the evaporation regime at 0.550 J/mm (b) close to the transition point with a feed velocity v_{feed} of 850 mm/s.

Within the HCR the spatter speed decreases with increasing feed velocity at a constant line energy. This can be explained by an increased interaction time of the laser beam with a respective part of the track. A longer interaction time leads to an increased overall momentum that is transferred to the surrounding atmosphere. Within the EVR, the amount of evaporated material will increase linearly with power, after exceeding the material and interaction time specific evaporation threshold intensity. However, similar to the observations within the HCR, a decreased interaction time leads to a decreased transfer of momentum.

3.2 Influence of Consecutive Tracks

While the median absolute spatter speed $v_{spatter}$ has shown to be constant over multiple consecutive tracks with a defined hatch spacing of 80 µm, the spatter count not only depends on the chosen process parameters. It is also highly influenced by the previous neighboring tracks, which is evident from the measurements shown in Fig. 6.

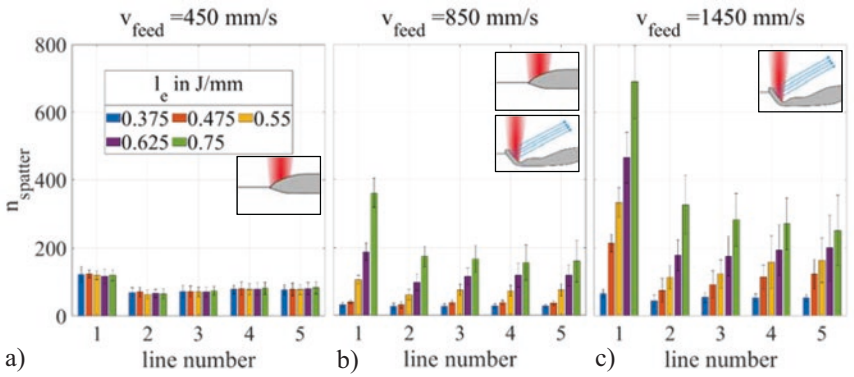


Fig. 6. Spatter count of consecutive 10 mm tracks for selected line energies and feeding velocities (n = 5).

The evolution of the spatter counts of the first line, which is produced on a powder bed which has not been altered by any previous processing, directly correlates to the observations of the spatter speed evolution in Fig. 4. Within the HCR the spatter count is not affected by an increase in line energy. Whereas the measurements within the EVR show a linear increase in spatter count. The transition from the HCR to EVR can also be observed in the spatter count measurements in Fig. 6b, for the first two line energies the spatter count is relatively constant and then increasing linearly. The evolution of the spatter count over multiple consecutive lines shows that for all measurements the spatter count decreases and approaches a constant value over multiple consecutive lines. In the HCR measurement in Fig. 6a - after the first line - the spatter count decreases by half and stays from there on effectively constant. One could expect this, since powder is only adjacent to one side of the current track and while the previously processed track is adjacent to the other side, from the second track on. However, this effect is not present for the first two line energies in Fig. 4b, which are also assigned to the HCR. Those stay rather at a constant value from the first line on. This can indicate that proportionally less powder from the side of the track is transformed to spatter than from the front of the process zone. Within the EVR the spatter count is higher for all measurements in comparison to the HCR. Additionally, a constant spatter count is approached slower with regard to the number of tracks.

3.3 Atmospheric Conditions

As depicted before, the atmosphere is plays an essential role in the powder entrainment mechanism. In [11] simulations have shown, that the particle entrainment is affected by the density of the used inert gas. Yet the effect on spatter count has not been experimentally investigated, whereas this might provide relevant information for process parametrization. Therefore, we processed the samples under helium, argon and nitrogen with a feed velocity of 650 mm/s with the same sets of parameters. Figure 7 shows the results of spatter count measurements of the first track.

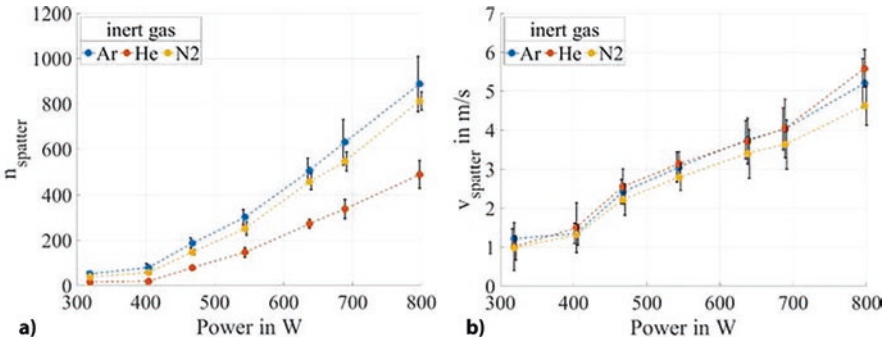


Fig. 7. Evolution of spatter count $n_{spatter}$ (a) and $v_{spatter}$ (b) at $v_{feed} = 650mm/s$ with increasing power under helium, nitrogen and argon atmosphere. (n = 5)

From the measurements it can be seen, that the spatter count significantly increases with the density of the selected process gas, which is 1.78 kg/m^3 for argon, 1.25 kg/m^3 for nitrogen, 0.18 kg/m^3 for helium at room temperature. Similar to the trend identified in 3.2 the spatter count increases with power, respectively line energy. The spatter count measurements contradict the simulative results in [11], which show that a lower process gas density results in an increased particle entrainment. An explanation for the discrepancy could be a lower proportion of entrained particles being molten and thereby converted to spatter due to a more confined vapor plume, which however was not further investigated. The spatter speed measurements do not show a significant process gas depended evolution, which suggests that the spatter speed is mainly influenced by interaction of the particles with the vapor plume, and not the atmosphere adjacent to it.

4 Conclusion

Within this paper we have investigated the effect of changing process conditions and periphery onto spatter count and spatter velocity. We have found a link between a transition of a HCR to EVR and spatter velocity, which has been validated by high-speed images of the molten pool topography. Furthermore, measurements revealed that within the HCR the interaction time mainly influences the spatter count and therefore the spatter count decreases with an increasing feed velocity. However, with the onset of evaporation, the spatter count increases linearly with power and interaction time. In addition to this, measurements over multiple consecutive lines have shown that the spatter count continuously decreases and approaches a constant value. The evolution of the spatter speed correlates directly to the spatter count. Additionally, the investigation of different inert gases revealed that the spatter count decreases with the gas density.

The results show that quantitative measurements of spatter characteristics provide new insight into process behavior. In this context, the used approach of a fully automated image processing algorithm has proven to be a beneficial tool, within the

presented work. However, the observations also show that the measurements can be well correlated with observations from specialized experimental setups. These setups cannot be realized in or transferred to an industrial PBF-LB/M machine, in contrast to spatter measurements. Future work will further investigate the correlation between the spatter characteristics and process conditions and their potential to be used for process control.

References

1. Ly, S., Rubenchik, A.M., Khairallah, S.A., Guss, G., Matthews, M.J.: Metal vapor micro-jet controls material redistribution in laser powder bed fusion additive manufacturing. Sci. Rep. **7**(1), 4085 (2017)
2. Bidare, P., Bitharas, I., Ward, R.M., Attallah, M.M., Moore, A.J.: Fluid and particle dynamics in laser powder bed fusion. Acta Mater. **142**, 107–120 (2018)
3. Cunningham, R., Zhao, C., Parab, N., Kantzos, C., Pauza, J., Fezzaa, K., Sun, T., Rollett, A.D.: Keyhole threshold and morphology in laser melting revealed by ultrahigh-speed x-ray imaging. Science **363**(6429), 849–852 (2019)
4. Matthews, M.J., Guss, G., Khairallah, S.A., Rubenchik, A.M., Depond, P.J., King, W.E.: Denudation of metal powder layers in laser powder bed fusion processes. Acta Mater. **114**, 33–42 (2016)
5. Darvish, K., Chen, Z.W., Pasang, T.: Reducing lack of fusion during selective laser melting of CoCrMo alloy: effect of laser power on geometrical features of tracks. Mater. Des. **112**, 357–366 (2016)
6. Zheng, H., Li, H., Lang, L., Gong, S., Ge, Y.: Effects of scan speed on vapor plume behavior and spatter generation in laser powder bed fusion additive manufacturing. J. Manuf. Proces. **36**, 60–67 (2018)
7. Eschner, E., Staudt, T., Schmidt, M.: Correlation of spatter behavior and process zone formation in powder bed fusion of metals. CIRP Ann. **69**(1), 209–212 (2020)
8. Yin, J., Wang, D., Yang, L., Wei, H., Dong, P., Ke, L., Wang, G., Zhu, H., Zeng, X.: Correlation between forming quality and spatter dynamics in laser powder bed fusion. Addit. Manuf. **31**, 100958 (2020)
9. Barrett, C., Carradero, C., Harris, E., Rogers, K., MacDonald, E., Conner, B.: Statistical analysis of spatter velocity with high-speed stereovision in laser powder bed fusion. Prog. Addit. Manuf. **4**(4), 423–430 (2019)
10. Eschner, E., Staudt, T., Schmidt, M.: 3D particle tracking velocimetry for the determination of temporally resolved particle trajectories within laser powder bed fusion of metals. Int. J. Extrem. Manuf. **1**(3), 35002 (2019)
11. Mayi, Y.A., Dal, M., Peyre, P., Bellet, M., Metton, C., Moriconi, C., Fabbro, R.: Laser-induced plume investigated by finite element modelling and scaling of particle entrainment in laser powder bed fusion. J. Phys. D Appl. Phys. **53**, 075306 (2019)

Investigation on Structural Integration of Strain Gauges using Laser-Based Powder-Bed-Fusion (LPBF)

M. Link[1(✉)], M. Weigold[1], J. Probst[2], R. Chadda[2], C. Hartmann[2], M. Hessinger[2], M. Kupnik[2], and E. Abele[1]

[1] Institute of Production Management, Technology and Machine Tools (PTW), Technische Universität Darmstadt, Otto-Berndt-Straße 2, 64285 Darmstadt, Germany
m.link@ptw.tu-darmstadt.de

[2] Measurement and Sensor Technology, Technische Universität Darmstadt, Merckstraße 25, 64283 Darmstadt, Germany

Abstract. We present a disruptive method for manufacturing structure-integrated force sensors based on the LPBF process by inserting a strain sensing steel plate in the additive manufacturing process of the deformation element. In order to investigate the strain transmission from the printed part to the integrated steel plate, we built two prototypes which differ in terms of geometry. Both of them were investigated focusing on strain transmission and differences in the thermal stress applied. Measurements revealed a high impact of induced thermal stress from the manufacturing process. It can be reduced by having a lower exposed area in the built-up process. A good linearity (about 1.5%FS) for the first prototypes with a nominal load of 50 N is achieved, due to a good strain transmission from the printed part to the integrated steel plate. Thus, the suitability of the LPBF method for the structural integration of force sensors can be confirmed.

Keywords: Laser-based powder-bed-fusion · Structural integration · Force sensors

1 Introduction

Structural health monitoring (SHM) enables damage detection and improves maintaining services, and, thus, is increasingly required in fields such as plant construction. Evaluation of structural health requires sensor integration of mostly mechanical sensors (e.g. force, torque) into existing process chains and according data analysis. Due to the complex geometries of the structures, a simplified integration of general-purpose sensors is not possible or only possible with considerable effort. Using laser-based powder-bed-fusion (LPBF) of metals, as an additive manufacturing

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 387–395, 2021.
https://doi.org/10.1007/978-3-662-62138-7_39

(AM) process, is a promising approach for the production of structure-integrated sensors which can be easily customized to the structure given [1].

LPBF is currently the most frequently used AM technology in the metal processing industrial sector [2]. The advantages over other AM processes are its comparatively high resolution and homogeneous process conditions.

Furthermore, the process is particularly suitable for component integration, such as the integration of sensors [3]. Typical for this manufacturing process are the layer-by-layer process sequence (see Fig. 1) for the generation of complex component structures and the production without component-specific tools [4].

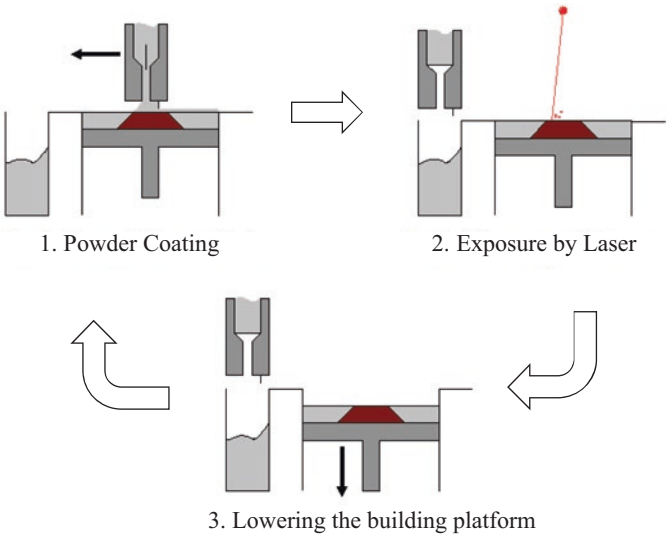


Fig. 1. Working principle of LPBF additive manufacturing technology [5]

A promising method for manufacturing structure-integrated force sensors is the integration of strain gauge sensor elements into the LPBF manufactured deformation body. Our concept is to apply strain gauges in full-bridge configuration to a conventionally manufactured stainless steel plate (1.4310) as a measuring element carrier and integrate it in the volume produced by LPBF (stainless steel 1.4404) by interrupting the additive manufacturing process (see Fig. 2). In this integration, LPBF is used to create material connection. This approach is beneficial in terms of complex geometry variants, individualization of the sensor and increased reliability through better protection against environmental influences. Furthermore, the sensors can be miniaturized to a greater extent through the functional integration of encapsulation and deformation body. These features are particularly advantageous in areas such as structural and system monitoring as well as in medical technology.

In order to ensure the sustainable production of mechanical sensors manufactured using LPBF based on this disruptive concept, main issues such as the assurance of a sufficient and reproducible strain transmission of the LPBF-manufactured

connection point and the control of the thermal processes during manufacturing need to be addressed. In this work we built up basic deformation elements of a one-dimensional cantilever-based force sensor by integrating a steel plate in the LPBF process and applied the strain sensing elements after manufacturing.

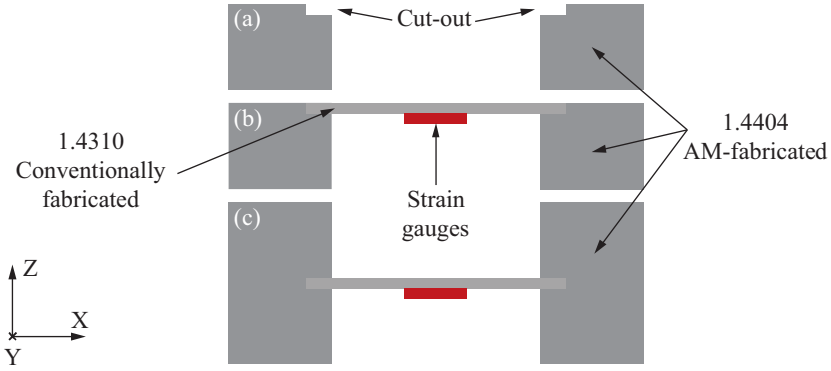


Fig. 2. Build-up process with LPBF-manufactured deformation body; production of base body (a); insertion of the steel plate as the measuring element carrier (b); completion of the prototype and application of the strain gauge (c)

2 Research Methodology

As first prototypes of our force sensors two cantilever-based deformation elements with the dimensions 80 mm × 31 mm × 25 mm were produced using a commercially available LPBF (model EOS M290, EOS GmbH, Krailling, DE) with a maximum build volume of 250 mm × 250 mm × 325 mm. This system has a 400 W single-mode Yb fiber laser with a wavelength of 1064 nm, which generates a focus diameter of 83 µm. Gas-atomized EOS 316L stainless steel powder was used for all experiments in this work. A scanning electron microscopic (SEM) analysis of the powder was used to determine the particle distribution. Particles in the range 11 to 54 µm and a mean diameter d50 of 25 µm were determined. For the layer-by-layer construction of the prototypes, the parameter set "316L_Surface_M291 1.10" with a layer thickness of 20 µm, qualified by the system manufacturer, was used.

First, the base body was built up to a height of 16.5 mm using LPBF (see Fig. 2a). Then, the building process was interrupted and the powder was removed at the corresponding recess. A cut-out of 5.1 mm × 10.1 mm × 2 mm was provided at this height to insert the steel plate as the measuring element carrier (stainless steel 1.4310). This steel plate (50 mm × 10 mm × 2 mm) has been cut by a CO_2 laser and deburred. After inserting this steel plate (see Fig. 2b), a layer of powder was applied. The LPBF process was then continued and a material connection between the steel plate and the base body was created (see Fig. 2c). For the second prototype, a cut-out with a width of 0.4 mm in the x-direction and 2 mm in the y-direction was produced above the steel plate. This reduces the laser energy applied due to the smaller exposure area (see Fig. 3) and there is no

exposure to powder in the edge area between the base body and the measuring element carrier. After completion the LPBF process, the top and bottom sides were reworked by milling and M5 threaded holes are drilled for fastening to the measuring system.

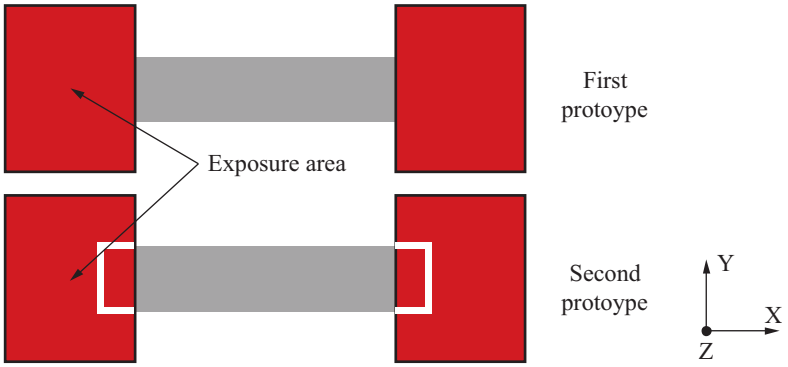


Fig. 3. Exposed areas from the two prototypes

In order to find a valid region for the strain sensing element for a nominal load, a finite element analysis (COMSOL Multiphysics 5.5, Burlington, USA) with the built-up deformation elements was carried out. The criterion is a proper strain in the range from 500 µm/m to 1000 µm/m for the application areas. For the boundary conditions, we set one side of the printed body to a fixed constraint and the other part of the printed body was loaded with a force. This resulted in a cantilever behavior.

The simulation shows a proper strain of up to 1400 µm/m at the surface of the integrated steel plate for a nominal load of 50 N (see Fig. 4). The simulation results for both deformation elements are nearly equal. However, singularities at the edges of the integrated steel plate result in exaggerated strains. We chose the center area of the steel plate as application area for the first. The mechanical stress is below the yield strength.

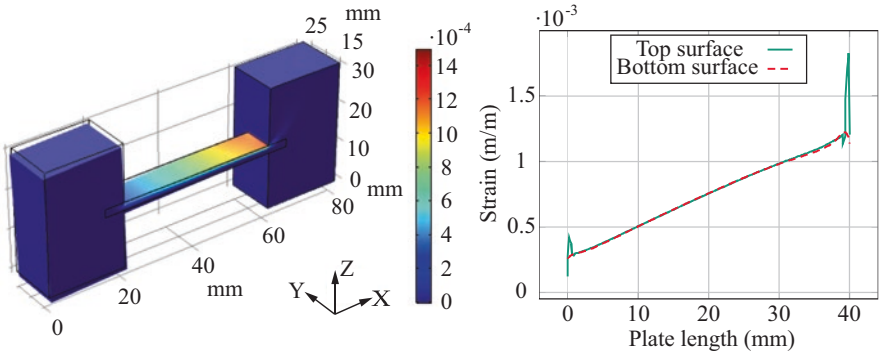


Fig. 4. Simulated deformation element 1 in COMSOL Multiphysics. The strain in x-direction shows the expected cantilever behavior with a valid region for the strain gauge application with roughly 750 µm/m at the mid-position of the inserted steel plate (left). Furthermore, the strain on the top and bottom surface is nearly equal (right). Note, that the values for the bottom surface are multiplated with −1 for a better comparability

As strain sensing element, we chose metal foil strain gauges. These enable high accuracies with small linearity errors and they are available in various types for specialized applications [6].

Therefore, single linear strain gauges (model BF350 – 2AA(11)N4, Zemic Europe, Etten-Leur, NL) with a gauge factor of $2.11 \pm 1\%$ and a base resistance of 349.9 ± 0.1 W were used for our first prototypes.

Mechanical connection of the strain gauges was done by using a hot-curing cyanoacrylate (model M-Bond 200, Micro-Measurements, North Carolina, USA), after preparing the application area with sandpaper and subsequent cleaning. The single strain gauges were connected to a full bridge circuit using enameled copper wire. For further protection against environmental influences, e.g. humidity or dirt, a silicone seal (model SG250, HBM/HBK, Darmstadt, DE) was used (see Fig. 5).

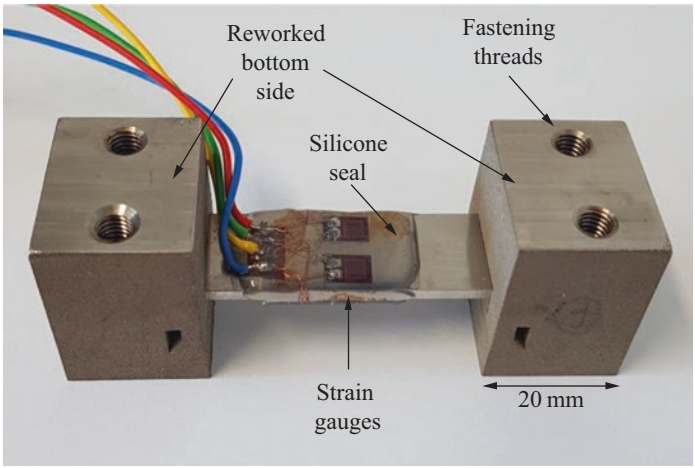


Fig. 5. Built-up prototype force sensor. The strain gauges are connected in a full bridge configuration and sealed with silicon

3 Experimental Setup

The two prototypes of force sensors (called prototype 1 and prototype 2) were characterized concerning their linearity with respect to a best-fit line. This best-fit line goes through the initial point, such that the maximum deviations (upward/downward) from the measurement signal have the same amount according to definition of the linearity deviation in [7]. Furthermore, the hysteresis of the force sensors was measured. The deviations to the best-fit lines were referred to the nominal value (reduced error). These properties were investigated at room temperature (21 °C).

A simple setup with weight plates for loading the deformation elements with a force was used (see Fig. 6). We attached a weight support plate on one side of the prototype where the weight plates were placed. The other side was fixed mounted. The nominal load of the sensors is 50 N. Loading was done with calibration weights

(2×2 kg, 1×1 kg) with an accuracy of ± 10 mg. Regarding the gravity of 9.81 m/s^2, this corresponds to a force of 49.05 N.

The measuring process was carried out following the guideline *OIML R60* (Metrological regulation for load cells), which included three full load cycles for pre-stressing followed by a three time repeated stepwise loading and unloading.

For investigating the linearity, loading and unloading steps of 20%, 40%, 60%, 80% and 100% of the nominal load were chosen. In case of the hysteresis the loading and unloading steps were 20%, 60% and 100% of the nominal load.

The bridge output signal of the force sensor (supplied with 5 V) was forwarded in a 4-wire circuit to a bridge amplifier (model EL3356-0010, Beckhoff Automation, Verl, DE) with an accuracy of $<\pm 0.01\%$. This bridge amplifier was connected to a data acquisition system (model cRIO-9074, National Instruments, Texas, USA).

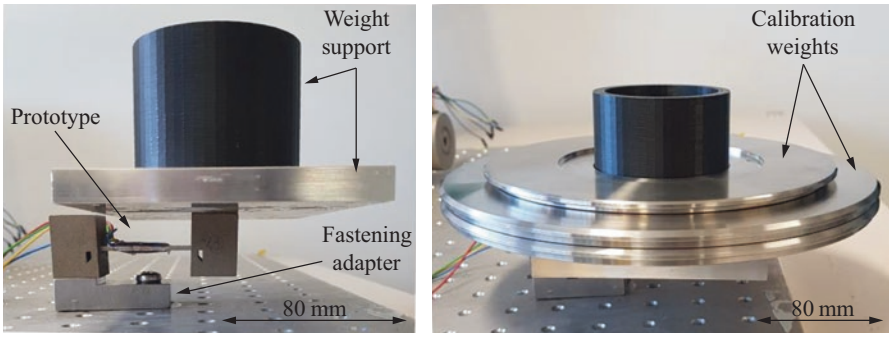


Fig. 6. Test setup for loading the sensor with the calibration weights. The sensor is fixed mounted (left) and weight plates are laid on weight support structure (right)

4 Results and Discussion

The construction process was successfully completed and a complete material connection between the base body and the measuring element carrier was achieved (see Fig. 7). The unmachined surfaces showed LPBF typical characteristic values of Ra = 15 µm. Due to the thermally induced energy, stresses occured in both prototypes. This resulted in a bending of the deformation elements, prototype 1 of 6° and prototype 2 of 4°, in relation to each other. The lower deformation of prototype 2 was a superposition of several factors. The smaller exposure area of prototype 2 in contrast to prototype 1 led to a lower introduced thermal energy via the laser beam. Consequently, a less thermal heating of the component was achieved, and, thus, a less deformation during cooling. In prototype 1 the powder was exposed between the base body and integrated steel plate. Due to the process-related surface roughness of the LPBF process, cavities were created between the base body and the steel plate. These were filled with powder particles at the start of the third phase (see Fig. 2c) of the process. Due to the increased thickness of the powder layer in these areas, heat dissipation was reduced, and defects were formed, which can lead to a higher induced thermal stress. Furthermore, no thermal post-treatment to reduce the induced stresses was carried out in these investigations in order to examine the as-built condition in the first step.

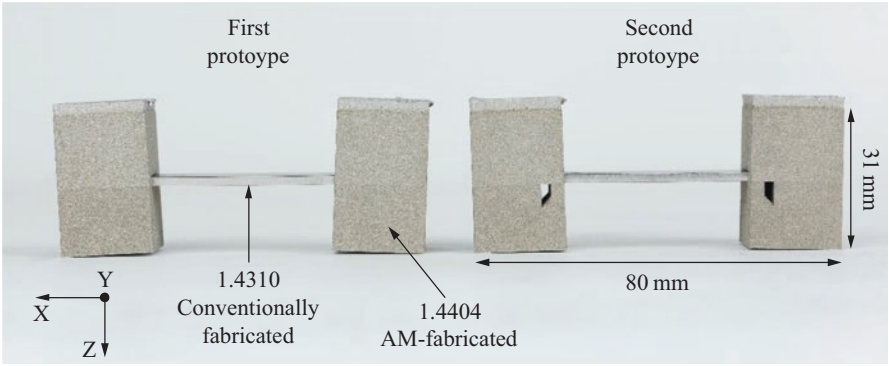


Fig. 7. Prototypes after separation from the LPBF building platform

The bridge offset voltages of the force sensors are 14 µV/V and 34.7 µV/V, respectively. These offset values resulted from the application process, the integrated steel plate, which was pre-stressed, and, therefore deformed as well as the resistance tolerance of the strain gauges. Assuming an optimal application of the strain gauges, an offset value of 157.2 µV/V caused by the resistance tolerance can occur.

The effect of the pre-stress seems to be noticeable in the measurements as well. Evaluation of the measurement results led to linearity errors (see Fig. 8) of up to ±1.5%FS for prototype 1 and ±1%FS for prototype 2, respectively. Nearly the same results occured for the hysteresis error (see Fig. 9). A pre-stress and eventually pre-deformation can impact the linear behavior. Nevertheless, both errors are in this initial state of this disruptive concept for manufacturing structure-integrated sensors in a good range. So the impact of influences of the application process of the strain gauges can be treated subordinated in comparison to the mechanical pre-stresses and the manufacturing process.

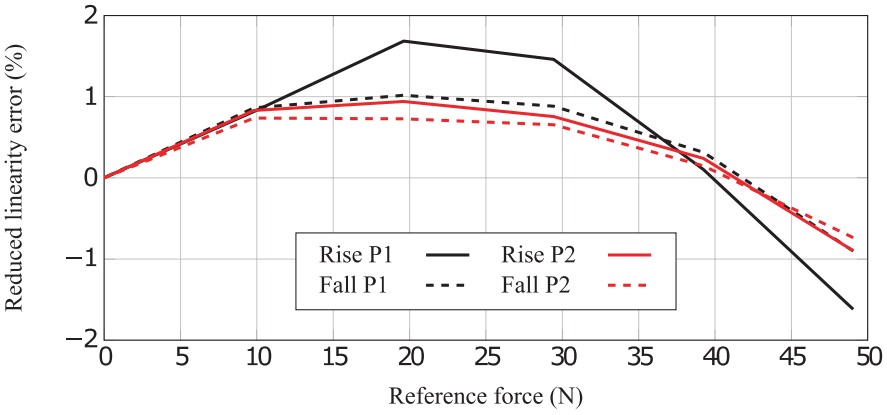


Fig. 8. Reduced linearity errors for increasing and decreasing force of both sensors for the third load cycle (P1 = Prototype 1, P2 = Prototype 2)

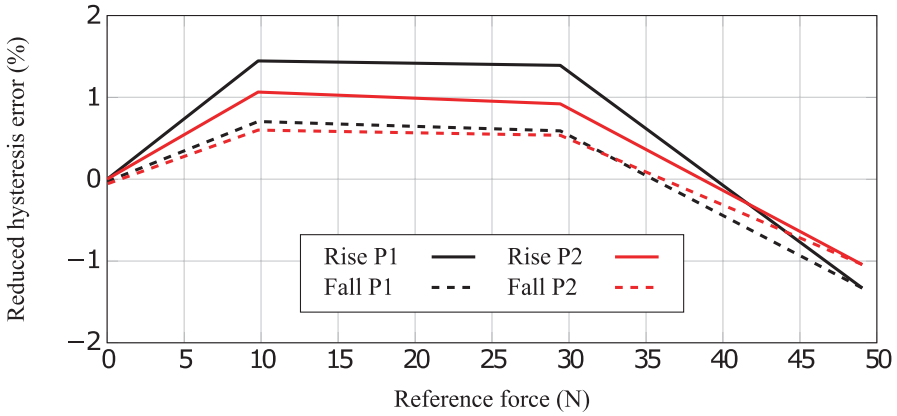


Fig. 9. Reduced hysteresis errors for increasing and decreasing force of both sensors for the third load cycle (P1 = Prototype 1, P2 = Prototype 2)

Further evaluations revealed a significantly higher strain for both prototypes, and, thus, higher output signal in the measurements in comparison to the simulation results (see Table 1). This might also be due to the pre-deformation of the prototypes, which induces additional forces acting on our deformation element that have to be investigated further. Thus, the simulation results based on the basic Euler–Bernoulli beam theory may not be valid. Modeling the cantilever behavior based on second order beam theory will be examined for future work.

Table 1. Summary of simulation and measurement results

	Strain (µm/m)	Output Signal (µV/V)	Sensitivity (µV/V/N)
Simulation	744	1569.8 ± 15.7	32 ± 0.3
Prototype 1	1036 ± 10.3	2185 ± 21.9	43.5 ± 0.4
Prototype 2	881 ± 8.8	1858 ± 18.6	37.1 ± 0.4

5 Conclusion and Outlook

This work presents a disruptive concept for manufacturing structure-integrated sensors based on additive manufacturing. Despite the induced thermal stress, a nearly linear strain transmission from the printed part to the inserted plate was achieved. Therefore, the suitability for manufacturing force sensors with this concept is confirmed. Our two prototypes showed that considering an additional cut-out for the insertion of the steel plate results in a stress relief in the manufacturing process, and, thus, in a better sensor performance. In future investigations the induced stress has to be reduced. This can be done by varying the process parameters in the LPBF process or by further modifying the geometry. Furthermore, a complete encapsulation of the strain gauges is the main goal. For this purpose, the strain gauge has to be applied to

the steel plate before integration into the deformation body. A heat treatment to reduce the stresses in the component will also be investigated. Damage to the strain gauges as a result of the temperature has to be prevented.

Acknowledgment. This research received support from the Deutsche Forschungsgemeinschaft (DFG) under grants KU 3498/4-1 and AB 133/108-1.

References

1. Hessinger, M., Kniepkamp, M., Lotichius, J., Hatzfeld, C., Werthschützky, R., Abele, E., Kupnik, M.: Strain transmission characteristics of steel substrates for additive sensor manufacturing with selective laser melting. In: Sensor 2017 – 18th International Conference on Sensors and Measurement Technology, AMA, Nürnberg (2017)
2. Wohlers, T.T.: Wohlers report – 3D printing and additive manufacturing state of the industry. Wohlers Associates (2018)
3. Abele, E., Kniepkamp, M., Link, M.: Additive manufacturing – Challenges and chances. In: 22nd International Seminar on High Technology, Piracicaba (2017)
4. Gebhardt, A.: Additive Fertigungsverfahren – Additive Manufacturing und 3D-Drucken für Prototyping – Tooling – Produktion, 5th edn. Hanser, München (2016)
5. Sander, J.: Selektives Laserschmelzen hochfester Werkzeugstähle. Saechsische Landesbibliothek- Staats- und Universitaetsbibliothek Dresden. Technische Universität, Dresden (2018)
6. Hoffmann, K.: Eine Einführung in die Technik des Messens mit Dehnungsmessstreifen. Hottinger Baldwin Messtechnik GmbH, Darmstadt (1987)
7. HBM Linearity deviation. https://www.hbm.com/en/0798/tips-and-tricks-torque-linearity-deviation. Accessed 14 May 2020

3D Printing Technology for Low Cost Manufacturing of Hybrid Prototypes from Multi Material Composites

L. Penter[1(✉)], J. Maier[2], B. Kauschinger[1], T. Lebelt[2], N. Modler[2], and S. Ihlenfeldt[1]

[1] Chair of Machine Tools and Control Engineering, TU Dresden, 01062 Dresden, Germany
lars.penter@tu-dresden.de
[2] Institute of Lightweight Engineering and Polymer Technology, TU Dresden, 01062 Dresden, Germany

Abstract. Combining fiber-reinforced composites with classic construction materials has tremendous potential for lightweight design but requires expensive equipment such as dies for injection molding. Additive manufacturing is cost-efficient for small volumes from fiber-reinforced materials. The MM3D (multi-material 3D printing) project analyses what technologies regarding interfaces, printer technology and identification of process parameters make economic fusion of 3D printing and established production processes possible.

Results from pull-off tests suggest that adhesion between surfaces doubles with optimal substrate temperature and quadruples with appropriate plasma treatment. Because hybrid structures require printing on free forms, the authors employ a hexapod machine for rotational motion of an extruder head. The paper presents solutions for resulting challenges such as referencing pre-existing structures in the workspace, generating the printing path and coupling extrusion rate to printing speed. Finally, the identification of process parameters is addressed.

Currently the technology is tested for manufacturing a nature inspired bike saddle.

Keywords: 3D printer · Hybrid parts · Fiber-reinforced composites

1 Introduction

Striving for weight reduction in every branch from car manufactures and aerospace industry to machine tool manufacturer requires intelligent and economic lightweight designs. Fiber reinforced composites unite low density with tailored high strength properties. Combining composites with classic construction materials such as steel and aluminum has tremendous potential for achieving afore mentioned goals. Multi material structures allow for integrating additional functionality, improving

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 396–405, 2021.
https://doi.org/10.1007/978-3-662-62138-7_40

mechanical properties and offer new design possibilities. Their economic application requires efficient production technologies for different production volumes. As of now, fusing these materials requires expensive production equipment such as pricey dies for injection molding, which makes hybrid structures unprofitable for middle or small production volumes.

Additive manufacturing (3D printing) could provide cost efficient and flexible solutions for manufacturing prototypes and small batches from fiber-reinforced materials and help to merge them with today's established production processes and classic materials. The research paper elaborates on three critical issues for printing 3D elements such as fins and joints on preexisting structures from metal or fiber-reinforced plastic sheets: interfaces, printer technology and parameter identification.

2 Interfaces

Intrinsic hybrid structures from sheets of metal or fiber-reinforced plastics in modern car bodies are specialized to specific applications by adding functional elements through injection molding [1, 2]. Crucial in the production of such assemblies is joining the semi goods. Typical approaches fall into one of the two categories: in-mold assembly (IMA) and post-molding-assembly (PMA). IMA joins components from different materials during forming; additional joining operation are obsolete. Only high part volumes make IMA profitable, as tooling prices can be between 130 to 170% of standard molds [3]. PMA assembles parts after injection molding outside the mold [4]. A broader variety of joining techniques such as interlocking, welding, or adhesive bonding make PMA more flexible regarding process and design changes than IMA.

Now, the MM3D approach combines PMA with fused-layer modelling (FLM). The approach promises full flexibility, as costs for tooling are low and most design changes do not entail mold changes. FLM uses a heated extruder to melt thermoplastic material, which is then deposited horizontal layer by horizontal layer to form the printed shape. In order to make full use of the lightweight potential, a proper connection between the components must be guaranteed.

2.1 Methods to Optimize Adhesion

Optimal adhesion between textile-reinforced thermoplastic components requires individual pre-treatment of the base substrate in the contact area. An established process for increasing the adhesive strength between thermoplastic components is plasma treatment [4]. Plasma pre-treatment alters the surface energy of polymers, which splits its long molecule chains. The increase in surface energy of the material leads to more efficient surface wetting and improves adhesion. The degree of wetting on the surface can be determined by measuring the contact angle (Fig. 1). The OWRK model [5] is used for calculating the surface energy from the contact angle. Small contact angles correlate with high surface energy.

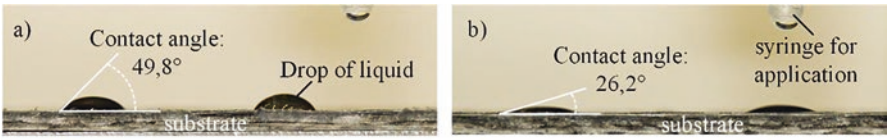


Fig. 1. Determination of wetting degree through contact angle measurement; a) low surface energy, b) high surface energy

A significant increase in adhesive strength can also be achieved by pre-heating the base substrate [6]. In order to understand the influence of different pre-treatments on the adhesive strength, the Institute of Lightweight Engineering and Polymer Technology (ILK) investigated semi-finished products from textile-reinforced PA6-GF (Tepex® dynalite 102-RG600(x)/47% Roving Glass) at different substrate temperatures with and without plasma activation. In addition, the experiments studied if the distance between FLM nozzle and base substrate surface affects the adhesive tensile strength. A robot cell by Plasmatreat GmbH with cold atmospheric pressure plasma and rotary nozzle was used for plasma activation. After plasma pre-treatment, the substrate was handled in airtight containers and directly transferred to the Ultimaker FLM system. Once the individual test temperature of the substrate was achieved, the test structure made from PA6 was printed on the substrate. Because FLM produces highly anisotropic structures, due to the layer-by-layer building process, a tensile load orthogonal to the layers provokes delamination within the additive structure and not in the contact zone with the substrate. In order to guarantee fracture in the contact zone, the new test specimen has a trapezoidal shape (Fig. 2a)). The interlocking load application prevents failure in individual layers. The ILK adapted the frontal tear-off test for metal-polymer composites according DIN 14916 [7] to the new geometry (Fig. 2b)).

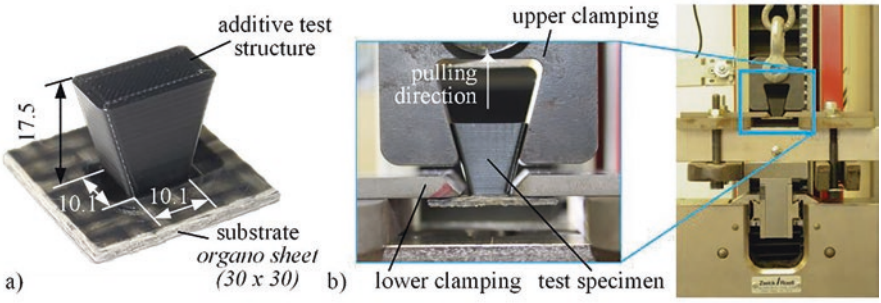


Fig. 2. a) Hybrid-test specimen; b) setup for tensile test

2.2 Results

The test results show that both preheating of the substrate and the distance of the FLM nozzle influence the adhesive strength (Fig. 3). Increasing temperature of the base

substrate up to 140 °C can significantly improve the adhesive tensile strength. The failure modes support this observation. In case of a 100 °C preheated base substrate, pure adhesive failure is observed. Further increasing the base substrate temperature to 120 °C leads to a mixed fracture of the composite. Tempering at 140 °C and 160 °C shows failure of the fiber-matrix adhesion,adhesion in the joining zone is much higher.

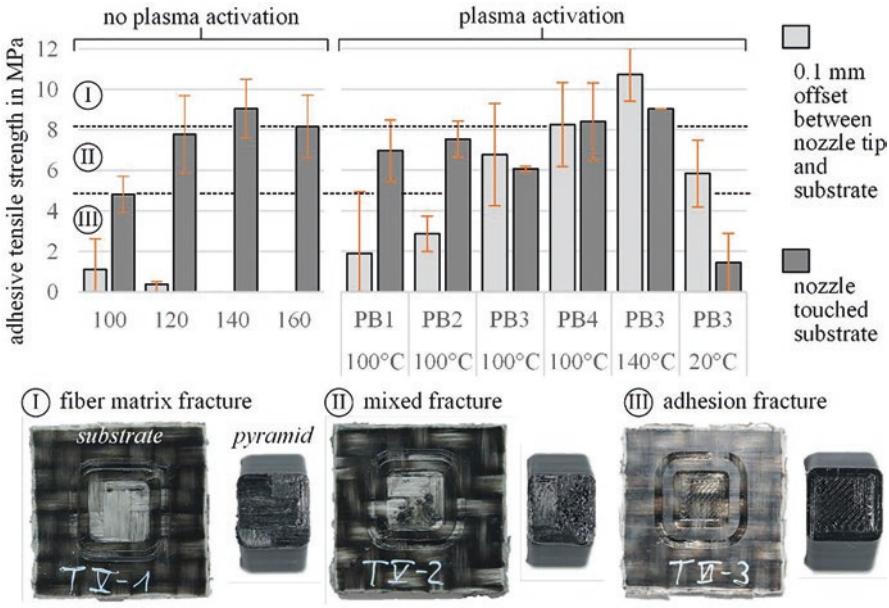


Fig. 3. Results of the adhesive tensile strength for the stages of substrate preheating and plasma treatment; contact angle after plasma treatment: PB1 – 53.9°, PB2 – 28.6°, PB3 – 25.2°, PB4 – 24.3°

The distance between the FLM nozzle tip and substrate surface influences the adhesion as well. At an offset of 0.1 mm between substrate and nozzle, the hot material melts the substrate surface only slightly when it first touches. The direct contact of the FLM nozzle with the surface leads to melting through both the FLM nozzle and the deposited material. Resulting structural changes in the contact area cause more intense diffusion of the polymer chains, which partially increases adhesion [8].

Suitable plasma pre-treatment of the substrate surface positively influences the strength of the connection (Fig. 3). The surface energy rises (as indicated by the decreasing contact angle) for parameter settings PB1 through PB4. The molecular adhesion behavior improves due to higher surface energy and enables a firm connection between substrate and FLM structure.

For a constant base substrate temperature of 100 °C and a nozzle-substrate distance of 0.1 mm, a suitable plasma configuration increases the adhesive tensile strength by a factor of four (Fig. 3, PB1 to PB4). Direct contact of the FLM nozzle with the preheated substrate surface also shows an increase in adhesive tensile strength. In addition, plasma pretreatment of the 20 °C substrate leads to higher adhesive tensile strength at a nozzle-substrate distance of 0.1 mm. It is, therefore particularly, suitable for generative hybridization processes with classic FLM manufacturing strategies.

3 Advanced Printer Technology

Printing on top of pre-existing free forms requires new printer motion features for three main reasons. First, current 3D printers move horizontal layer by horizontal layer and do not allow for depositing material with a constant offset on free forms. Second, there must be no objects in the printing area because common printers do not feature collision avoidance. Neither do slicer programs. Third, uninterrupted printing around free shapes requires velocity changes of the extruder head forward motion. Because material deposition of today's printers is not linked to the current velocity of the feed axis, reducing speed during printing leads to material overflow.

Besides pure motion and printing features, referencing and scanning functionality is essential for knowing the currently existing structure as a prerequisite for adjusting motion path. This is particularly important for addressing thermal deformations due to cooling. The cooling rate largely depends on the printed geometry.

3.1 Hexapod Printer with Screw Extruder

The printer technology in this paper features a hexapod structure, which allows for precise volumetric motions of a technology platform (Fig. 4a, b). Beside the translational motion of conventional printers, the platform can rotate 45-degree around all axes, which allows for an advanced collision avoidance with pre-existing structures (Fig. 4c). The technology platform carries a screw extruder, the filament, and measuring equipment. The extruder molts filament and deposits the material to the heated bed or the base structure (above called substrate). It is capable of extruding significantly larger volumes (0.48 kg/h at a rotational speed of extruder screw (rss) of 150 $\mathrm{min}-^{1}$, 0.6 kg/h at 165 min^{-1}) than current filament FDM printer.

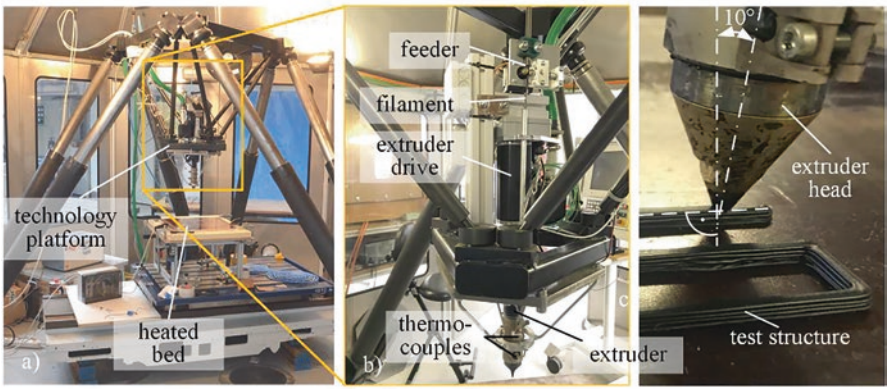


Fig. 4. a) hexapod with technology platform, b) technology platform with extruder, c) extruder on inclined technology platform

The 6D motion makes new printing strategies such as material extrusion across different horizontal planes possible. For referencing pre-existing parts, the hexapod printer is equipped with measuring units such as laser scanner and tactile measuring gage. Thermocouples constantly measure temperatures of bed and extruder.

The amount of material deposition must be constant for all extruder positions and printing paths. This is particularly important when moving around corners and printing interlocking features. Therefore, the amount of material extrusion must adjust to the current feed axis speed or vice versa. The study tested, if an empirical model can describe the non-linear correlation between potential input such as rotational speed of the screw and output such as amount of deposited material and if it enables a closed loop process control.

The potential 45-degree inclination of the extruder is most likely limited in reality. The study investigated limitations on the example of simple geometric shapes.

3.2 Results

Because the screw extruder of the hexapod printer extrudes large volumes, the resulting material flow at the nozzle reacts with a significant delay to changes of the screw's rotational speed. Consequently, the adaption of rotational speed to feed axis speed changes requires a feed forward control. The rss must be adjusted at a certain distance before the extrusion volume needs adaption. In order to maintain a constant layer width, the feed axis speed is then gradually adjusted to the current extrusion rate (material flow rate $\dot{V}_{total}$). Since a direct measurement of the extrusion rate was not possible, an empirical model was implemented (Fig. 5a). The correlation between extrusion rate and rotational speed of the screw extruder is non-linear and depends mainly on the thermal and rheological condition. The thermal influence decreases with running extrusion time and vanishes under steady state conditions. The rheological influence results mainly from delays in pressure build-up. The parameter identification of the correlation model is based on experimental data. The actual extrusion volume was determined via optical measurement of the printed layer width multiplied by the set layer height.

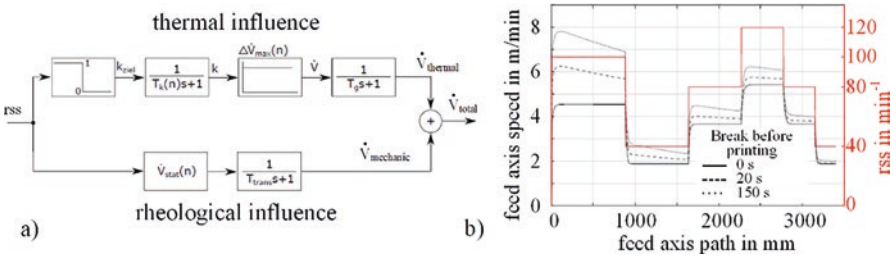


Fig. 5. a) Empirical model for calculating the extrusion rate depending on rotational speed of the screw; b) feed axis speed for constant printing layer width of 5 mm

Based on the model, the rotational speed of the extruder screw can serve as a quantity for the current extrusion rate. The feed axis speed is then adjusted to the current rss to maintain a constant layer volume. The rss graph and the corresponding feed axis speed based on the model prediction is shown in Fig. 5b.

The application is demonstrated for a rapid change of rss from 100 to 40 min^{-1}. The results in Fig. 6 show that adjusting the feed forward speed with current rss leads to a constant material deposition.

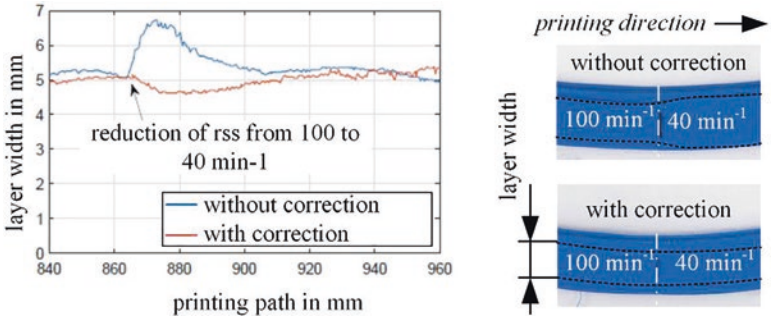


Fig. 6. Model based correction of feed axis speed (material: polypropene, nozzle diameter: 2 mm, temperature at nozzle tip: ca. 230 °C)

Planning the printing path entails to link rotational positions of the extruder head to the translational position. The investigations of the potential inclination of the extruder show that there is a limitation to less than 8 degree for printing on flat surfaces. Extruding the material requires counter pressure between nozzle and substrate for a correct position of the layer. Is there no resistance, the inner extruder pressure pushed the material of target (Fig. 7a).

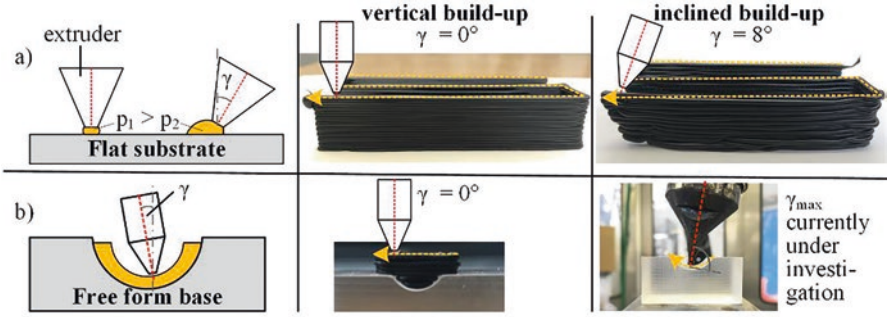


Fig. 7. Influence of extruder inclination on printing quality

Increasing the extruder inclination reduces counter pressure between nozzle and substrate and amplifies the material push-off. When printing on free forms (Fig. 7b)), the printer nozzle should point perpendicular to the surface. If so, there are no pressure related problems expected.

4 Process Parameter Identification

The quality of 3D printed elements depends on a multitude of parameters. Besides material properties and adjustable process parameters, the environmental conditions and printer properties affect the appearance of the printed layer. As a result, the

process commissioning requires extended trail & error especially for new materials and material combinations. The goal of the following investigation was to find a method to limit the parameters, which need identification, to speed up commissioning.

4.1 Sensitivity Analysis for Parameter Selection

In order to select the parameters most influential on the mechanical properties for identification, a sensitivity analysis was conducted. The tensile specimens were printed on a conventional 3D printer and were evaluated through tensile tests under different environmental conditions. The variation of input parameters (nozzle and bed temperature, feeding speed) is based on a fractional factorial design. For process steps "3D printing" and "tensile test" relevant temperatures, velocities and mechanical properties were monitored. Infill strategy and density as well as set layer thickness were constant.

4.2 Influences on Mechanical Properties

The properties determined through tensile test represent material inherent properties and the process parameters set for the printer Fig. 8. The nozzle temperature positively correlates with Young's modulus and the true layer thickness. The test bed temperature shows only a weak correlation with the true thickness. The printing velocity seems to have no influence on any of the mechanical properties. The results also show a dominant positive correlation for tensile test temperature with the ultimate tensile strength and a strong negative correlation with the yield stress. Based on the information only the nozzle temperature is essential for the mechanical properties of the part and would need identification during commissioning. A data set that is more detailed is currently under investigation.

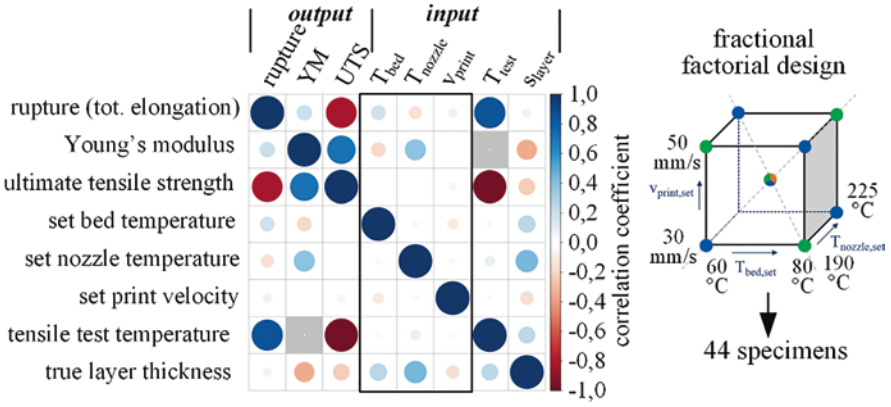


Fig. 8. Correlation analysis for printing parameters and mechanical properties on test specimen (Data analysis was conducted with Detact by Symate GmbH)

5 Conclusion

The technology package is currently applied to the process chain for manufacturing a bike saddle made from organic sheets and various 3D-printed metals as well as PA6 parts in the TU Dresden labs Fig. 9.

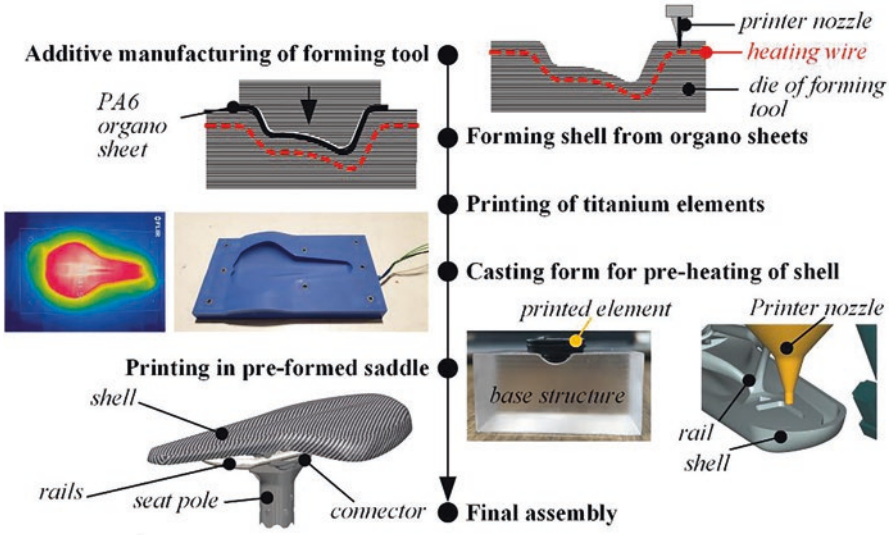


Fig. 9. Current process chain for manufacturing hybrid bike saddle

The process chain starts with additive manufacturing of a forming tool for producing the saddle shell from organo sheets. The saddle rails and the connector to the seat poll are printed from titanium. To connect the rails with the sheet, the hexapod printer is used to fill the pre-form with PA6. Due to its flexibility, the hybrid production technology is a potential process for producing cost efficient multi material parts for prototyping, individual products and low volume production.

Acknowledgements. The Sächsische Aufbaubank GmbH and its InnoTeam program support the MM3D project. The European Union and the Free State of Saxony co-finance the InnoTeam program.

References

1. Krahl, M., Bürkle, E., Gude, M., Stegelmann, M.: Kunststoff und Metall – Ein unzertrennliches Paar. Kunststoffe **12**, 32–37 (2017)
2. Luft, J., Troschitz, J., Krahl, M., Gude, M.: Thermoplast-Sandwichstrukturen aus recycelten C-Fasern. Lightweight Des. **5**, 20–25 (2018)

3. Knight, M.: In-mold assembly: the new frontier for multi-shot molding. https://www.ptonline.com/articles/in-mold-assembly-the-new-frontier-for-multi-shot-molding. Accessed 16. Mai 2020.
4. Mandolfino, C., Lertora, E., Gambaro, C.: Influence of cold plasma treatment parameters on the mechanical properties of polyamide homogeneous bonded joints. Surf. Coat. Technol. **313**, 222–229 (2017)
5. Owens, D., Wendt, R.: Estimation of the surface free energy of polymers. J. Appl. Polym. Sci. **13**, 1743 (1969)
6. Al-Sheyyab, A.: Hybrid structures – process integration and optimized performance. PHD thesis, Friedrich-Alexander-Universität Erlangen-Nürnberg (2008)
7. Bauer, A., Bischof, C.: Haftfestigkeitsuntersuchungen an Metall-Polymer-Verbunden. II. Stirnabreissversuch – Eine Methode zur Bestimmung der Verbundfestigkeit. Plaste Kautsch. **30**(4), 208–212 (1983)
8. Habenicht, G.: Kleben – Grundlagen, Technologie, Anwendungen. Springer, Berlin (2009)

Automated Assembly and Robotics

Value Stream Kinematics

E. Mühlbeier(✉), P. Gönnheimer, L. Hausmann, and J. Fleischer

wbk – Institute of Production Science, Karlsruhe Institute of Technology,
76131 Karlsruhe, Germany
edgar.muehlbeier@kit.edu

Abstract. The trend towards individualized products and the increasing demand for a greater variety of variants create new challenges for existing production environments and require a re-thinking of production. Established manufacturing systems that provide the desired flexibility are associated with significant productivity restrictions and are therefore unable to compete economically with production from rigid production lines. They are therefore often limited to serving niche markets. Consequently, an approach is needed that combines high productivity with high flexibility. For this purpose, this paper presents a new approach to manufacturing with an equally high productivity and flexibility, so-called value stream kinematics. The basic idea of value stream kinematics is to combine the advantages of specialized machines with the versatility of industrial robots. The vision behind this is to be able to realize entire value streams with uniform robot-like kinematics and no need for special machines.

Keywords: Flexible manufacturing system (FMS) · Kinematic · Reconfiguration

1 Introduction

Industry 4.0 is meanwhile known worldwide as a success story and stands like no other term for the innovativeness and creativity of research and industry. However, Industry 4.0 is essentially a digital innovation that merely adds additional sensors and actuators to conventional physical production systems. In order to utilize and further develop the full potential of Industry 4.0, the idea beneath it must be transferred to the next level, the hardware level. With Industry 4.0, greater flexibility and productivity in production is achieved by collecting and processing data. Thinking further ahead, now the machines and systems must also be adapted to the production of the future in order to make full use of this created potential. It is not enough to equip machines with more sensors and connect them together. The production machines and systems of tomorrow must be fundamentally rethought. In this paper, a novel production concept, which proposes to design entire production lines out of universal robot-like kinematics, will be presented.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 409–418, 2021.
https://doi.org/10.1007/978-3-662-62138-7_41

2 State of the Art

In recent years, driven in particular by Industry 4.0, a number of new approaches have emerged to increase the flexibility and versatility of production. Besides the networking of machines with sensors, the cooperation of robots plays a major role. An example from the automotive industry is a setup of a production system consisting of 45 cooperating robots for the rear axle assembly for a mid-range vehicle by the Daimler AG. Up to six robots work together in a robot network. This means that even moving parts can be processed and, in addition, dead times due to workpiece transfers are eliminated. Additional variants were quickly integrated into the production system and thus demonstrated that a production system with flexible robots also provides improved versatility [1]. However, the production system is only designed for assembly and is not capable of carrying out more demanding manufacturing processes such as machining operations. The challenges that prevent the qualification of robots for machining processes and the unused potential are analyzed and summarized in [2]. The authors come to the conclusion that the direction-dependent cutting forces must be considered more closely in relation to the pose dependent stiffness of the robots. This requires better process models and compliance models of the robots.

A model-based approach for the offline correction of motion paths is investigated in [3]. With a preceding process force simulation the deviation of the robot TCP is calculated and a corrected motion path is defined. In [4] a fuzzy controller based on a stiffness model is used for deflection compensation. To achieve a high reaction speed of the model, the fuzzy controller switches between different control strategies depending on the state of the machining process. In [5] a static deflection model is also used to compensate for the displacement of the TCP of an industrial robot used to machine large aerospace parts. However a spindle holder equipped with sensors allows the measurement of real process forces during operation and therefore the setup of more precise models. In [6] the methodology is enhanced by a process simulation and following path optimization (e.g. reduced feed rate in critical positions) implemented in form of an extended CAD-CAM interface. For their investigation and validation named approaches use heavy and massively constructed industrial robots (KR 500, KR240) for machining aluminum parts. Machining steel parts is not possible and for different production processes the chosen robots would be greatly oversized, which reduces the flexibility.

A versatile production requires not only the use of capable flexible robots but also extensive software assistance to support the planning and reconfiguration of production systems. This requires good information models of the production resources, especially of the robots and end effectors, as well as of the products to be manufactured and their specific requirements. An approach to describe production resources and the added value of a standardized information model is presented in [7]. The model, which is comprehensible for machines, enables the automated search for suitable production resources. However, this requires that a large number of manufacturers are willing to provide such detailed information models for their products. In [8], the product requirements are additionally presented as an information model. A matchmaking algorithm is used to determine a suitable manufacturing solution for

each product requirement from the available production resource pool. However, the system is not suitable for setting up an optimized production system, rather for limiting the search radius as well as checking the general manufacturability of a product with the given resources. Limiting the production machines to a uniform universal kinematic could simplify the design of a detailed information model. The authors of [9] present in their article a software platform for the creation and reconfiguration of assembly systems. The system calculates a system configuration on the basis of production scenarios, varying in the number of product variants to be manufactured and a fluctuating demand on the market, whereas the adaptability to expected changes is optimized. By taking into account the changes and associated costs, the configuration with the lowest total costs is still found. However, the system only considers product families and not a switch to a completely different product and furthermore is also limited to pure assembly.

A demonstration system, which also takes other manufacturing processes into account, is presented in [10]. The hybrid robot production cell combines additive processes (e.g. laser cladding), machining and quality assurance. Due to the near-net-shape production a high resource efficiency is achieved. The robots are qualified for the different manufacturing processes by changing end effectors. Another example known from the industry is the matrix production of Kuka. However, as explained in [11], the production islands are optimized for a specific manufacturing group and are therefore not completely flexible.

The approaches mentioned in this chapter all have in common that they only consider a partial aspect of the product production process. However, a holistic concept of a flexible production system which considers the whole production instead of just a production cell has not yet been developed.

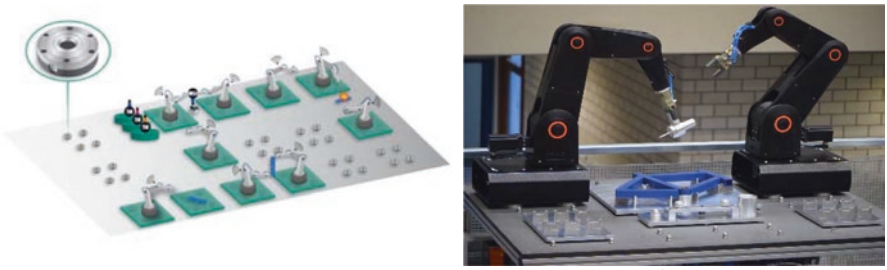
3 Solution Approach

3.1 Vision of the Value Stream Kinematics

In the following the new concept of this paper is introduced. The novel production concept intends to design entire production lines out of universal robot-like kinematics. The value stream will be built entirely from these flexible kinematics and with the complete abandonment of special machines. Figure 1 shows a visualization of the production concept (a) as well as a first simple demonstrator (b). A production system created by such an approach has the versatility to fully utilize the potential created by current Industry 4.0 developments.

The unified freely configurable robot-like kinematics carry out the handling tasks common in robotics and are also capable of carrying out acute processes such as assembly, additive manufacturing, cutting and joining processes as well as machining tasks and quality assurance using various manufacturing tools. The replacement of expensive, nonflexible specialized machines within a value stream through versatile kinematics could greatly increases the flexibility of the production chain. Today even high flexible machines are restricted to a limited amount of different manufacturing processes. Machining centers for example are often capable of executing turning and

milling processes and sometimes even additive processes. But obviously they cannot change to completely different manufacturing processes e.g. laser cutting. Contrary to special machines robot kinematics can be empowered relatively simple to do a completely different manufacturing processes by changing the end effector of the kinematic. Assuming that robots are also able to execute processes with high requirements to precision or high operating forces, the reconfiguration of a production line consisting entirely of such robots would be easier and faster compared to existing production systems. Only the end effectors have to be exchanged and the robots have to be repositioned according to their new function. The challenge here will be to redesign state of the art robots to empower them to do many additional sophisticated manufacturing processes without creating an over-developed expensive robot design. The primary focus here is on cutting and joining processes. The main groups of forming and shaping processes are not further considered, with the exception of additive processes and bend-forming. Neither are techniques of process engineering, e.g. from chemistry, considered due to completely different requirements.



(a) (b)

Fig. 1 Concept of a value stream of robot kinematics (a) and first demonstrator (b)

3.2 Coupled and Cooperative Kinematics

Especially machining processes represent a group of these sophisticated manufacturing operations with particularly high requirements for accuracy and applied forces. Specialized robots are available for these processes. Such cutting robots are very massive and heavy built and would thus be completely oversized for simple handling tasks. The adaption of the robot design to the process enables it to carry out a specific process, but is associated with a reduction in its flexibility.

Value stream kinematics follow a different approach. Instead of achieving a higher performance of the kinematic unit by adapting the design, it is achieved by cooperation of several kinematic units. For example, particularly heavy or difficult-to-handle objects can be gripped by several kinematics simultaneously. The level of cooperation varies from joint handling to a mechanical coupling of several kinematic units.

In order to investigate the potential for increasing the stiffness by coupling several kinematics, a simplified, fully parameterized FEM model was set up using the software Abaqus FEA. A single kinematic is modeled by a 3D beam structure with arbitrary cross section and a linear elastic isotropic material. For each individual joint

the respective non-linear axial, radial, torsional and tilting gear stiffness is considered by special connectors. In order to calculate the static compliance in the workspace, a force of 1000 N was applied at the tool center point (TCP) in each pose and each global positive and negative x-, y-, z-direction. Based on all the TCP deflections for one pose the major axis of deflection respectively the compliance can be calculated by eigenvalue decomposition. The whole construction of the model as well as the post processing is realized by a Python script exploiting the Abaqus scripting interface. To calculate the compliance in the entire workspace fully automatically and efficiently, the Python script was further extended by a self-written inverse kinematics solver for general serial robots with joint constrains taking the TCP position as input and calculating the gear rotation angles. By looping over the gridded workspace the pose dependent compliance and the reachable workspace can be estimated.

The simulative preliminary investigations have shown that the pose dependent overall stiffness of a system of coupled kinematics is significantly higher compared to a single kinematic unit. In addition, the direction-dependent compliance in the different poses is more uniform as can be seen from the comparison in Fig. 2. In future simulative investigations the coupling element in particular has to be given more importance.

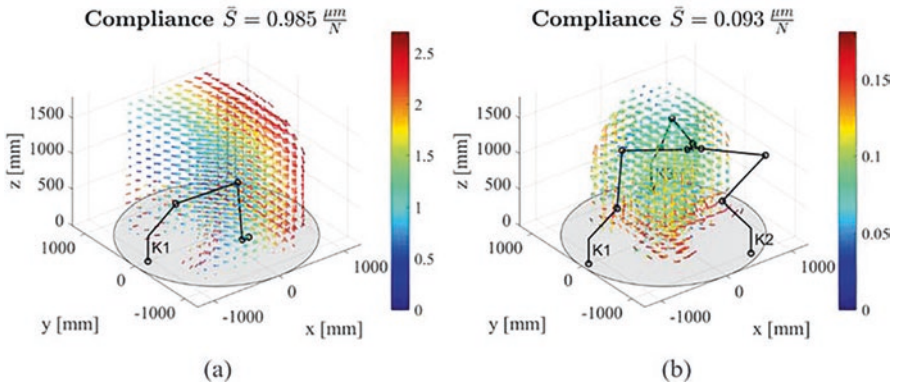


Fig. 2 Comparison of the simulated TCP compliance for (a) single and (b) three coupled kinematics, visualized using the numerical computing software MATLAB

By interconnecting several kinematic systems, the performance can be scalably adapted to the process and thus an individualized modification of the units is not required. The restriction to one uniform kinematic or at least to a small number of variants is necessary in order to achieve a high output quantity in production of kinematic systems in the future and therefore lower per-unit costs, as such kinematics will effectively only be used for complex and at the same time modest processes if they can compete with conventional handling solutions in terms of price.

The increased stiffness of a coupled system is achieved at the expense of mobility and the available workspace. Here the main influencing factors are the number of coupled units and their alignment. A practical use of the coupling mechanism requires a methodical scientific investigation of the mobility and stiffness of coupled systems. In

the long run, therefore, a tool for calculating the optimal configuration of the units for given process conditions must be developed.

The coupling also introduces the requirement of a new control architecture of the units. In order to ensure that the kinematics work independently, each unit must be equipped with its own control system. Furthermore, the controllers must be able to detect a coupling process and thereby switch into a master–slave state, in which a single controller directs the movement of the coupled system. The master controller must have sufficient computing capacity to calculate the inverse kinematics, interpolation and path planning of the entire system as well as ensuring synchronous movement. Alternatively, it is possible to distribute the computing load among all participating controllers or even to a central computing unit. In the latter case, however, real-time-capable communication between the central computer and decentralized controllers must be ensured. Within the scope of this research project it has to be investigated whether wireless data exchange, e.g. with 5G technology, can be used for this purpose or whether established wired communication must be used to maintain real-time capability. In addition to the additional computing load, the synchronization of the coupled units is a major challenge and research content of the project.

3.3 Reconfiguration

The concept of value stream kinematics focuses the production of small and medium sized production volumes with a high number of variants. Standardized products with very high volumes will continue to be manufactured in the future with rigid production lines consisting of special machines, which allows the lowest possible unit costs to be achieved. In recent years, however, there has been an increasing trend towards more individualized products in all areas of production. It can therefore be assumed that in the future a large number of conventionally manufactured products will be produced in smaller batch sizes with a greater variety of variants. As a result, production concepts such as the one presented will become more relevant for a variety of products.

Production in smaller quantities inevitably results in more frequent reconfigurations of the value stream. Due to the versatile kinematics presented in this concept, a reconfiguration can be done almost without investing in new hardware. Instead of replacing entire production machines, in many cases it will be sufficient to replace end effectors and reconfigure the production layout. For this purpose, the concept provides a special factory floor following the Plug & Play principle of a Lego board. The entire production area is covered with a grid of zero point clamping systems. The counterparts sit at the base of the kinematics. Media, power and data cables are routed underneath the factory floor. This means that the entire production area is equipped with clamping technology for attaching the kinematics and access points to media, power and information supply. The aim of this initial extra effort is to significantly accelerate the repositioning of the kinematics. With a load crane, a kinematic unit can be moved to the correct position and placed on the clamping systems where it is connected to one of the supply points. If the kinematics and the clamping systems on the production area are exactly measured, the exact positioning of the kinematics after

insertion into the zero-point clamping pots is also known. The Plug & Produce principle will simplify the conversion of production to other products and accelerate the production ramp-up. However, it must be investigated whether mounting the kinematics with a zero-point clamping system has a negative effect on machining accuracy compared to a rigid mounting on the floor.

3.4 Empowering Kinematic Units

A major challenge of the research project is the empowerment of the kinematics. In addition to the coupling mechanism, the concept envisages achieving greater rigidity and accuracy through further optimization. For this purpose, the kinematic joints will be equipped with new gears technology. These gears have no backlash as well as a significantly higher torsional stiffness and also enable the transmission of higher torques compared to conventional planetary or cycloid gears through the surface contact of the teeth [12]. The influence of such gears were also investigated with the simulation approach already presented.

Last but not least the compliance of the robot kinematics can be compensated by advanced control algorithms. The deflection of the TCP can thus be regulated during operation. To cover these approaches, the research project is supported by industrial partners in the implementation of the kinematic qualifying measures.

3.5 Engineering Platform

The installation and constant reorganization of a production system is a complex and time-consuming process. A comprehensive engineering platform will be necessary to help production engineers to reduce planning and commissioning time of production systems significantly. Various software tools for the development of production systems are already established in the industry. Siemens, for example, covers a large part of production system planning with Tools like NX, Mechatronics Concept Designer and Line Designer, just to name a few. The engineering platform is supposed to bring those functionalities into a common platform and to be extended by additional software components. In the following, the basic idea and functionality of such a future platform are described.

Figure 3 shows the main modules of the virtual platform. In the first module the CAD model of the end product is designed. The properties of the product and the required manufacturing processes are derived from the model in the second module. In order to implement the processes, in the next module, the quantity, arrangement and positioning of the kinematics as well as necessary couplings and end effectors are defined. In the last module the production sequence can be validated by simulating each individual process and the production system as a whole. The simulation data obtained are fed back into the hardware configuration and process design for the purpose of optimization. This iterative process is repeated until the production simulation results yield a prediction of a product quality that meets the requirements of the end product. In the last step, control programs for the kinematics are extracted from the

simulation. The automated generation of control programs is an essential element of the whole concept allowing rapid reconfigurations.

With the digital platform a significant part of planning and commissioning can be carried out on an entirely virtual basis. This makes it possible to plan future production systems in a time- and cost-optimized manner right from the start.

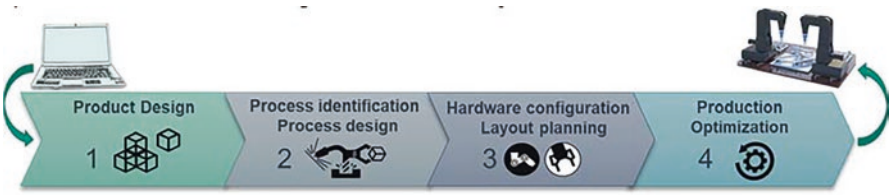


Fig. 3 Digital platform supports on the way from product model to production system

4 Implementation and Outlook

The feasibility of the concept presented in this paper shall be demonstrated with the help of a representative industrial demonstrator production. Within the scope of the research project, two preliminary demonstrator stages with a lower technology readiness level are planned as can be seen in Fig. 4. The first stage of the demonstrator is a test bench consisting of two smaller handling robots arranged on a plate with zero-point clamping systems (see Fig. 1(b)). The robots are each equipped with grippers to grip a press-in tool with two gripping points and move it jointly. Loading the robots with real process forces is not yet planned in this demo stage. The demonstrator serves only for the preliminary investigation of the coupling concept. For this purpose, a control architecture is currently being developed in a first step, in which a central controller controls both robots, triggers movement commands and continuously monitors the synchronization of the robots and readjusts it at control level. In the next development step, one of the two controllers is supposed to take over control as master and thus also controls the other robot. The central unit will then only serve as an interface to enable communication between the two robots. The final goal of the preliminary investigations is to derive the control architecture for the second demonstrator stage.

In the second stage, a production line consisting of a total of four industrial robots (cf. Kuka KR180) will be set up. Two of these robots will be equipped with Galaxie® gears and used as coupled kinematics. The second demo stage pursues three main goals:

- Integration of new gears and demonstration of improved performance
- Upgrade of robot controllers to allow synchronous coupled movements
- Functional proof of the coupling concept by executing a real production process

In the final stage, a prototype plant for the representative manufacturing of products will be established. Based on the research results of the second demonstration stage, a new six-axis kinematic unit will be developed in a first step. During the development of the new kinematics, the previously investigated optimization potentials with regard

to stiffness must be fully exploited. In contrast to conventional industrial robots, the kinematic unit will be designed for coupling and thus use an appropriately adapted robot controller. Supported by the research work, the development of the kinematics and control will be carried out by industrial partners. Once the basic infrastructure of the prototype plant has been set up, a first simplified production start-up will be carried out. Possible sample products are electric motor components and battery module housings with module assembly. Producing different types of products allows to test reconfiguration time when switching between products.

With the demonstrator system, key performance indicators of production can be determined and compared to conventional production systems. It is expected that throughput and cycle times will tend to decrease compared to rigid production systems. However, the production concept of value stream kinematics does not attempt to outperform highly automated production systems in terms of productivity, but rather plays out its strengths in case of large product variety and frequent reconfiguration. However, it is basically conceivable to manufacture products with no diversity and high quantities in order to cover peak demands, for example. The prototype plant will further be used as a basis for developing the digital platform. The system will serve as a test-stand to evaluate and further develop the various functions of the platform.

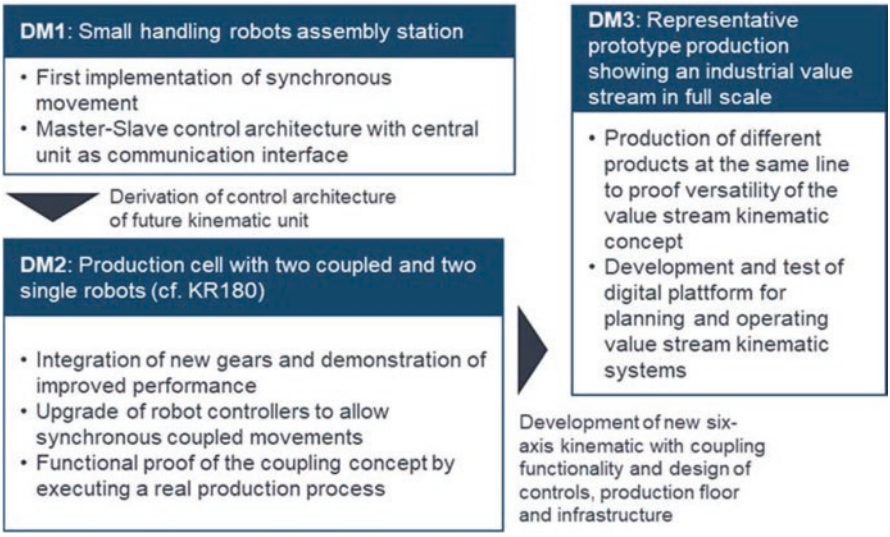


Fig. 4 Three-stage implementation strategy

5 Conclusion

The growing establishment of Industry 4.0 applications in production is pushing the development of more flexible production systems and plants. The concept presented in this paper addresses this development and attempts to go one step further by focusing on the versatility of the entire production system. Small and medium quantities are to be produced competitively and the production of large quantities is also

basically feasible on a value stream kinematic system. Value stream kinematic facilitates a fast reaction to suddenly appearing or increasing demands for certain products. The current developments with regard to global crises show again the necessity of such systems.

However, the implementation of the presented concept still requires great efforts in development and research. The realization of a kinematics capable of coupling requires the development of new technologies, both on the hardware and software side. Especially the real-time capable synchronization of the kinematics and the development of adapted connectable end effectors is a research focus. This also requires a detailed analysis of the requirements (e.g. in the case of cutting: chip removal, coolant supply and thermal behavior) for the different manufacturing processes. Another target is the development of the digital platform and the associated network infrastructure. Process and production simulation with coupled units form the core of the research work.

References

1. Steegmüller, D., Zürn, M.: Wandlungsfähige Produktionssysteme für den Automobilbau der Zukunft. In: Bauernhansl, T., ten Hompel, M., Vogel-Heuser, B. (eds.) Industrie 4.0 in Produktion, Automatisierung und Logistik. Springer Vieweg, Wiesbaden (2014)
2. Iglesias, I., Sebastián, M.A., Ares, J.E.: Overview of the state of robotic machining: current situation and future potential. Procedia Eng. **132**, 911–917 (2015)
3. Schnoes, F., Zaeh, M.F.: Model-based planning of machining operations for industrial robots. Procedia CIRP **82**, 497–502 (2019)
4. Zaeh, M.F., Roesch, O.: Improvement of the machining accuracy of milling robots. Prod. Eng. Res. Dev. **8**(6), 737–744 (2014)
5. Denkena, B., Lepper, T.: Enabling an industrial robot for metal cutting operations. Procedia CIRP **35**, 79–84 (2015)
6. Denkena, B., Brüning, J., Windels, L., Euhus, D., Kirsch, S., Overbeck, D., Lepper, T.: Holistic process planning chain for robot machining. Prod. Eng. Res. Dev. **1**(6), 715–722 (2017)
7. Siltala, N., Järvenpää, E., Lanz, M.: Value proposition of a resource description concept in a production automation domain. Procedia CIRP **72**, 1106–1111 (2018)
8. Järvenpää, E., Siltala, N., Hylli, O., Lanz, M.: Capability matchmaking procedure to support rapid configuration and re-configuration of production systems. Procedia Manuf. **11**, 1053–1060 (2017)
9. Colledani, M., Yemane, A., Lugaresi, G., Borzi, G., Callegaro, D.: A software platform for supporting the design and reconfiguration of versatile assembly systems. Procedia CIRP **72**, 808–813 (2018)
10. Baier, C., Hähn, F., Tepper, C., Weigold, M.: Robot-based hybrid production concept. In: Wulfsberg, J., Hintze, W., Behrens, B.A. (eds.) Production at the Leading Edge of Technology. Springer Vieweg, Berlin (2019)
11. Plewe, M.: Smarte Matrix-Produktion: Vision wird bei Kuka Realität. MECHATRONIK: Synergien in Industrie 4.0 (7–8), (2018)
12. Wahlster, W., Beste, D.: HERMES AWARD – Internationaler Technologiepreis der HANNOVER MESSE. Innovationen für die industrielle Produktion - die ersten zwölf Jahre. 1. Aufl. Springer Vieweg, Wiesbaden (2016)

Simulation-Based Robot Placement Using a Data Farming Approach

T. Lechler[1(✉)], G. Krem[2], M. Metzner[1], M. Sjarov[1], and J. Franke[1]

[1] Institute for Factory Automation and Production Systems, Friedrich-Alexander-University Erlangen-Nuremberg, 91058 Erlangen, Germany
tobias.lechler@faps.fau.de
[2] Friedrich-Alexander-University Erlangen-Nuremberg, 91058 Erlangen, Germany

Abstract. Increasing flexibility in production systems is driving the use of robotic solutions. During their planning, robots must be placed according to their future operations. Thereby, influences such as space limitation, mechanical reach or cycle time must be taken into account. This paper introduces a concept based on the data farming methodology aiming at the optimal robot positioning for a given set of constraints. By simulating a defined sequence of robot operations with changing robot placement in a definable investigation area, each result data set is stored and analyzed. The simulation run with the best fitting robot position according to the defined key performance indicators is shown. For further evaluation, a clustering algorithm is used to evaluate the simulation results. The usage of the proposed method enables production planners to conveniently place robots in the optimal position according to their later application.

Keywords: Simulation · Data farming · Robot placement

1 Introduction

The economic market situation requires the industry to build and develop production systems faster due to increasing product variation and shorter product life cycles. The similarly increased flexibility of these systems can be supported by digital planning and simulation tools for the development of new production systems [1]. As innovative technologies e.g. advanced robotics lead to an increase in automation in industrial systems, the use of simulation gains relevance in production system engineering and design. Virtual commissioning (VC) in particular allows reliable studies of the production system [2].

VC is used as a common method for the evaluation of functionality, performance and safety aspects of production systems before their real i.e. physical commissioning [3]. For this, fully functional simulation models are required, which are nowadays used for a time and cost efficient development process [4]. Besides the application in

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 419–428, 2021.
https://doi.org/10.1007/978-3-662-62138-7_42

the planning of new production systems, simulation can also be used in retrofit applications. There, the automation of existing manual workstations is an often discussed task. It comprises the finding of the optimal robot position at the existing workplace. In this paper, this question is answered by exploiting a data farming (DF) based approach. With the use of DF, a large amount of information can be discovered using simulation [5]. This can be interpreted to gain new insight about the simulation model and its real representation. DF is mainly used for large scale and data intensive simulation experiments in order to cover a large variety of possible system behavior [5,6]. Here, a DF approach is used to generate a large amount of data to determine the optimal robot position.

The paper is separated into five sections: In Sect. 2, the state of the art in DF as well as robot placement is presented. Afterwards, the DF concept for the robot placement is explained. The application of the concept including the simulation model and the results are presented subsequently. Finally, a brief summary and outlook are given.

2 State of the Art

For a better understanding of the proposed concept, the state of the art in DF and robot placement are presented. Furthermore, the need for action is determined based on the introduced main topics.

2.1 Data Farming

Whereas data mining comprises the collection, preparation and analysis of existing data in order to discover new knowledge about a research object, DF includes the generation of the data based on a simulation [7, 8]. DF was first introduced at the 1999 Winter Simulation Conference in Phoenix [9]. The metaphor of farming in DF rests on the following steps: [9]

- Fertilization: Provision of expert surveys to gain expert knowledge about the research object.
- Cultivation: Identification of potentially relevant influencing factors on the planned simulation.
- Planting: Integration of the knowledge into simulation models and execution of the simulation using an investigation area based on the expert knowledge.
- Harvesting: Usage of data mining approaches to analyze the data and detect new correlations.

According to this procedure, a complete examination of all possible parameter combinations based on a full factorial experiment design can be performed. Despite the growing availability of high performance computers, the described brute force method is still limited concerning its performance. Therefore, especially the design of experiments is still focus of current research [8].

The DF method was initially developed for military applications. In this field, strategic simulations depending on a large variety of parameters are executed [10].

Furthermore, questions concerning medical care as well as the evacuation of people in earthquake-prone regions could be analyzed using DF [11]. In the context of production and logistics, Feldkamp et al. describe a DF approach for knowledge discovery in manufacturing simulations and present different applications such as manufacturing simulations and gold mining studies [12, 13]. Furthermore, applications for decision support based on DF in the context of supply chain management are elaborated [14]. In summary, most use cases base on a discrete event simulation, which is most often implemented using the software Plant Simulation by Siemens.

2.2 Robot Placement

In industry sectors, such as automotive or electronics manufacturing, robots are used to perform specific tasks on a workpiece in a production system. As these tasks are repeated numerous times, it is important, that they are performed as quickly as possible to reduce cycle times and therefore maximize throughput [15]. Here, the performance of a robot during a process depends on its placement relative to the positions that need to be reached and therefore its corresponding path. Robot placement describes the optimal robot positioning in a production system with reference to its cycle time operations [16]. In practice, the finding of a suitable robot position is usually based on expert knowledge. Consequently, there are different approaches to optimize the process time or the robot position. Table 1 gives an overview of different approaches for robot placement including the used software in chronological order.

Table 1. Implementations of robot placement

Year	Author	Field of application	Software	Source
1996	Feddema et al.	Algorithm for optimal robot position	Not specified	[17]
2008	Mitsi et al.	Algorithm for optimal robot position and axis angle	SolidWorks and Fortran	[18]
2009	Kamrani et al.	Movement path optimization of ABB industrial robots	VBA, Matlab and RobotStudio	[19]
2010	Baizid et al.	Optimal robot placement zone	SolidWorks with Application Programming Interface	[16]
2013	Boschetti et al.	Movement path optimization of robots	Not specified	[20]
2016	Spensieri et al.	Movement path and robot placement optimization	IPS Robot Optimization and Lua	[15]
2019	Siemens AG	Smart place	Process Simulate	[21]

The publications show that the existing approaches for the determination of the optimal robot position base on analytical descriptions and calculations. The extension of potentially already existing simulation models for later VC by a DF framework to determine the optimal robot position by simulating the later working procedure has not been considered to date.

3 Concept for Data Farming and Knowledge Discovery

Figure 1 shows the proposed concept for simulation-based robot placement based on the DF methodology. Starting with the experiment definition, the user defines the investigation area for the DF application. Then, the design of experiment using full- or fractional factorial experimental design defines the simulation runs stored in the experimental plan data base and needed to generate the required data. The correlation of the simulation results with the experimental plan allows the analysis of the research object. Here, analysis processes such as classification-, clustering- or regression-analyses can be used, which allows a deeper understanding of the research object. This improved knowledge can be used as new input data to find more suitable solutions and to perform an iterative overall process. In this paper, this concept is applied on the search for the optimal robot placement in a defined investigation respectively installation area of given simulation. Therefore, the approach aims for a user-friendly implementation of DF in the VC environment.

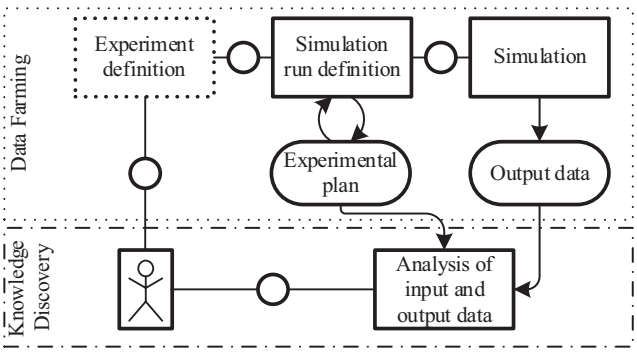


Fig. 1. Proposed concept in the Fundamental Modeling Concept (FMC) notation based on [8]

4 Prototypical Implementation and Validation for Robot Placement

This section introduces a prototypical validation of the proposed concept. Starting with the description of the robot placement use case, the implementation as well as execution of the concept are shown. In conclusion, the DF results are analyzed and discussed.

4.1 Use Case Description

An exemplary use case demonstrates the basic functionality of the proposed concept. Figure 2 shows the corresponding 3D scene on the right side and a top view on the left. The robot moves the workpiece from a central pick-up position to the placing position on the left side. In consequence, the workpiece pose at the place position is

fixed. The DF target is to find the optimal robot position regarding a minimal process time. The process starts and ends at a defined initial robot position and includes the described pick and place task. The manipulated variables in this scenario are in theory the six degrees of freedom of the robot basement. Since pick and place tasks and their chain-linking are common in robot applications, this use case is simple, but representative for extended future applications of the proposed concept.

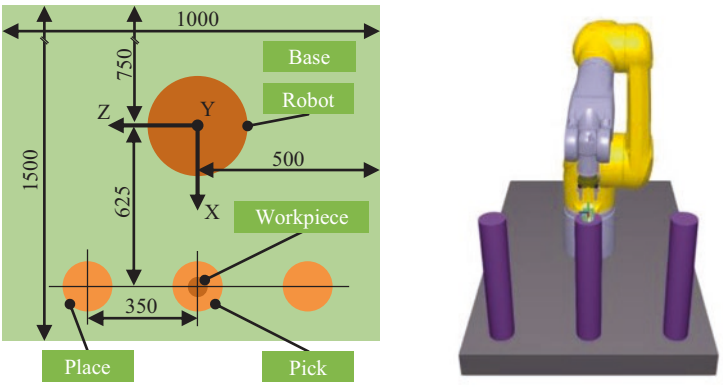


Fig. 2. Use case setup (left) and simulation model (right)

4.2 Implementation of the Concept

The prototypical implementation is based on the described state of the art (Sect. 2) and utilizes the proposed concept (Sect. 3). The resulting system uses DF as well as knowledge discovery. Figure 3 shows the crucial components of the prototypical implementation. Within the experiment definition, the parameter variation for the simulation experiment is provided. The input data is read via a script included in a C-Sharp project. There, the simulation runs for the later experiments are defined and stored in an experiment data file. An application programming interface (API) couples the simulation and the controller script. In the prototypical setup, the software industrialPhysics contains the simulation model shown in Fig. 2. The model contains a Stäubli TX2-90L with the built-in generic controller, the pick and place positions as well as a source and drain for the workpieces. Its behavior is recorded using key performance indicators (KPI), which are stored in an output data file for later data analysis. The KPIs in this use case are the robot axes, simulation time and resulting process time. Based on the experiment and output data, the analysis in the knowledge discovery sector takes place, where a Python script helps detecting correlations between the experiment and output data. Experiment definition and control are provided to the user via a corresponding graphical user interface.

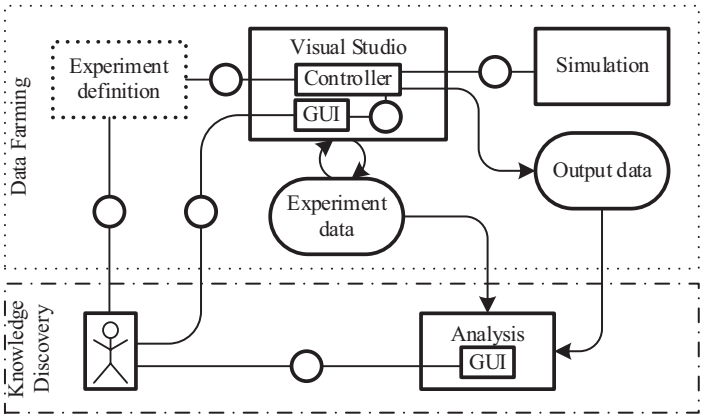


Fig. 3. Overview of the prototypical implementation in FMC notation

4.3 Data Farming

In order to find the optimal robot position for the given discretization, the valid area for the robot basement needs to be defined first. Theoretically, all six degrees of freedom of the robot basement pose, could be determined, though the exemplary implementation includes the translational coordinates only, to reduce computing time. In the proposed use case, the valid area is defined from −0.15 m to 0.15 m in steps of 0.05 m in every translational coordinate axis shown in Fig. 2 on the left side. The test planning is implemented fully factorial, which leads to a total of 343 simulation runs based on three parameters with seven entries each. The results are recorded in a resolution of 0.01 s and stored in the output data file as described earlier. The computing time for the 343 simulation runs is 27 min 6 s.

4.4 Knowledge Discovery

The knowledge discovery is divided into two parts. At first, the result data is presented. Then, k-means clustering is used to further analyze the simulation. Figure 4 shows the result of the 343 simulation runs. The process times differ from 3.47 s to 3.77 s in the defined investigation area. The highest can be found at x-positions of 0,15 m, which means that the closer the robot position is located to the pick and place positions, the longer the tasks takes. With declining x-position values, the process times decrease in general except for some outliers at high z-positions. The best robot position in the defined area from −0.15 m to 0.15 m is at $x = -0.10\ m$, $y = 0.10\ m$ and $z = -0.05\ m$ with a process time of 3.47 s.

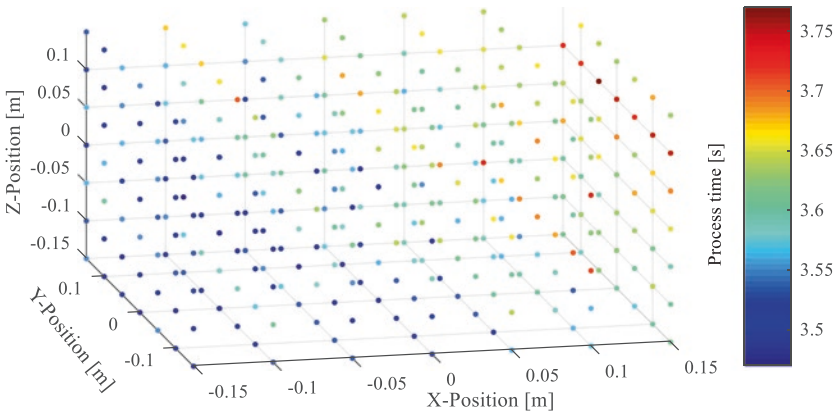


Fig. 4. Visualization of process times according to the robot position

To further evaluate the farmed data, the process times are clustered using k-means clustering with $k=5$ cluster. The number of clusters is chosen according to the Calinski-Harabasz criteria for $k=1$ to $k=5$ cluster. Figure 5 shows the distribution of the five clusters in a box plot on the left. Since cluster five has the lowest process times, the x-, y-, and z-position values of this cluster are plotted over the other clusters in Fig. 5 on the right in a parallel coordinates plot. It can be stated, that the x-position influences the process times the most. With lower x-position coordinates, y- and z-position values can vary without affecting process times much.

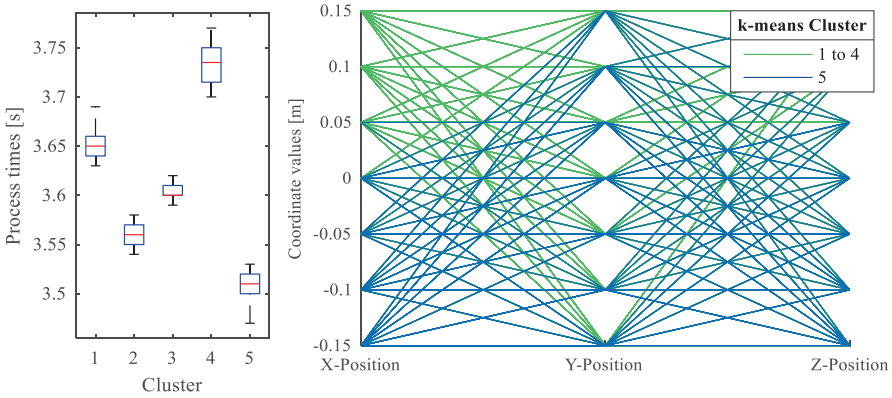


Fig. 5. Distribution of five process time k-means cluster (left) and visualization of robot position coordinates x, y and z of all clusters (right)

To explore correlations in the simulation results even further, a second $k=5$ k-means clustering is applied on the process time and the three position coordinates. Figure 6 shows on the top the results in a parallel coordinates plot. To visualize the relations between the process time and the position, all four parameters

are standardized using the zscore algorithm. Since the clusters are plotted on top of each other in a parallel coordinates plot, the distributions of the five clusters for each parameter are plotted in box plots for better visibility. As in the previous clustering, now, the clusters with low process times, cluster 2 and 3, contain low x-position values. They differ at the y- as well as the z-position values. Since cluster 2 contains the best process time, it can be assumed, that higher y-position and lower z-position values could have beneficial influence on the process times according to the generic robot controller used in the simulation.

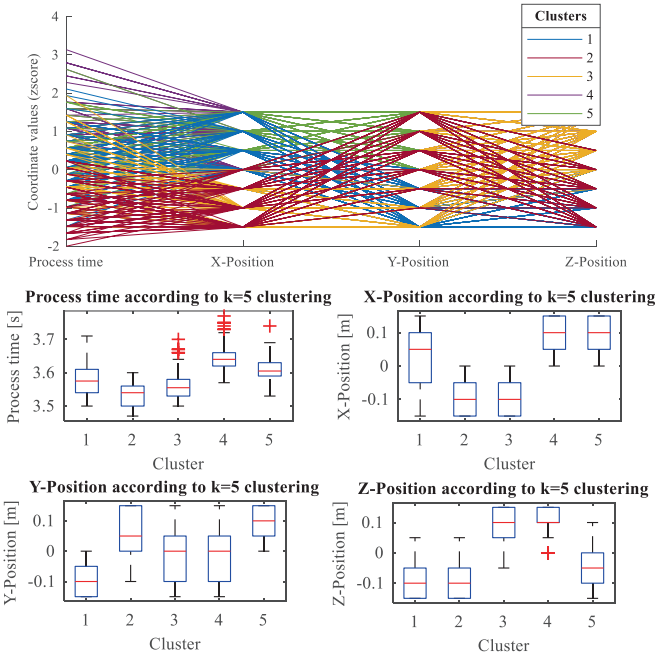


Fig. 6. Parallel coordinates plot of k-means clustering ($k=5$) over standardized process time, x-, y- and z-positions (top); distribution of clusters at process time, x-, y-and z-position

4.5 Discussion

The previous sections show, that the proposed DF concept can be used to determine the optimal robot placement in a defined investigation area based on a simulation model. Here, a full factorial experiment design leads to a data set, which is then further analyzed. Since the presented use case only includes a simple pick and place application, the expected results are reproduced by our method. However, the approach can be extended for more advanced applications to specify the impact of different robot positioning. Besides the presented clustering analysis, the, here predefined, investigation can be determined by a wide experiment design and a subsequent regression analysis to determine a target including the optimal robot position. Nevertheless, using a generic robot controller, the exact behavior of the real robot

cannot be determined. Therefore, a connection between the simulation model and a manufacturer-specific robot controller needs to be established. Furthermore, the used robot controller uses a fix axis configuration. For the implemented pick and place scenario, a configuration change of the robot is unnecessary and should be avoided during programming. For more complex processes however, the evaluation of multiple configurations and configuration changes is important for process time analysis as well. The simulation of the 343 experiments currently requires about 27 min. A reduction of the computing time can be achieved by deactivating the visualization and by the distribution of the experiments on different simulation instances.

5 Conclusion and Outlook

Robot placement is a common task in new production planning as well as retrofit applications. The current solution methods for robot placement barely use the benefits of simulation tools. With the increasing usage of simulation in production system planning, the extension of existing simulation models with a data farming framework allows the determination of the optimal robot position. Therefore, this paper introduces an augmented data farming concept combining the virtual commissioning software industrialPhysics with an automated experiment manager and a graphical user interface for the visualization of the results. The prototypical implementation of a pick and place use case including one robot shows, that the optimal robot position can be determined by using the data farming approach. Moreover, the implementation allows a straightforward exchange of the simulation model to analyze other, more advanced robot placement use cases. Besides robot placement, the presented approach allows the discovery of further knowledge about the production system based on the simulation model, which is shown by the implementation of k-means clustering. The future usage of a manufacturer-specific robot controller can further improve the accuracy of the simulation results. This, as well as the implementation of different axis configurations at the same robot position, are the topics for future research. Besides the improvement of the simulation and its control, the results should be validated with a real robot setup.

References

1. Kuhn, A.: Fast ramp up. Schneller Produktionsanlauf von Serienprodukten. Verl. Praxiswissen, Dortmund (2002)
2. Lee, C.G., Park, S.C.: Survey on the virtual commissioning of manufacturing systems. J. Comput. Des. Eng. **1**, 213–222 (2014)
3. Lechler, T., Fischer, E., Metzner, M., Mayr, A., Franke, J.: Virtual commissioning – scientific review and exploratory use cases in advanced production systems. Procedia CIRP (2019). https://doi.org/10.1016/j.procir.2019.03.278
4. Zäh, M.F., Wünsch, G.: Schnelle Inbetriebnahme von Produktionssystemen. Werkstattstechnik **95**(9), 699–704 (2005)

5. Feldkamp, N., Bergmann, S., Strassburger, S., Borsch, E., Richter, M., Souren, R.: Combining data farming and data envelopment analysis for measuring productive efficiency in manufacturing simulations. In: 2018 Winter Simulation Conference WSC, pp. 1440–1451
6. Horne, G.E., Schwierz, K.P.: Summary of data farming. Axioms (2016). https://doi.org/10.3390/axioms5010008
7. Aggarwal, C.C.: Data Mining. The Textbook. Springer, Cham (2015)
8. Feldkamp, N.: Wissensentdeckung im Kontext der Produktionssimulation. TU Ilmenau Universitätsbibliothek, Ilmenau (2020)
9. Brandstein, A.G., Horne, G.E.: Data farming: a meta-technique for research in the 21st century. Maneuver Warf. Sci., pp. 93–99 (1998)
10. Horne, G.E., Meyer, T.: Data farming and defense applications. In: Armstrong, R., McNamara, J., Pinelli, T.E. (eds.) MODSIM World Conference and Expo, Hampton, VA, USA. Hampton, VA: Langley Research Center, pp. 74–82 (2010)
11. Horne, G., Seichter, S.: Data farming in support of NATO operations - methodology and proof-of-concept. In: Proceedings of the Winter Simulation Conference 2014, pp. 2355–2363 (2014)
12. Feldkamp, N., Bergmann, S., Straßburger, S., Schulze, T.: Data Farming im Kontext von Produktion und Logistik. In: Wenzel, S., Peter, T. (eds.) Simulation in Produktion und Logistik. Kassel University Press, Kassel (2017)
13. Feldkamp, N., Bergmann, S., Strassburger, S., Schulze, T.: Knowledge discovery in simulation data: a case study of a gold mining facility. In: Proceedings of the 2016 Winter Simulation Conference, pp. 1607–1618 (2016)
14. Rabe, M., Scheidler, A.A.: Farming for Mining – Entscheidungsunterstützung mittels Simulation im Supply Chain Management. In: Rabe, Markus, Clausen, Uwe (eds.) Simulation in Production and Logistics. Fraunhofer IRB Verlag, Stuttgart (2015)
15. Spensieri, D., Carlson, J.S., Bohlin, R., Kressin, J., Shi, J.: Optimal robot placement for tasks execution. Procedia CIRP (2016). https://doi.org/10.1016/j.procir.2016.02.105
16. Baizid, K., Chellali, R., Bentaleb, T., Yousnadj, A., Meddahi, A.: Virtual reality based tool for optimal robot placement in robotized site based on CAD's application programming interface. In: Proceedings of Virtual Reality International Conference (2010)
17. Feddema, J.T.: Kinematically optimal robot placement for minimum time coordinated motion. In: Proceedings of IEEE International Conference on Robotics and Automation. IEEE International Conference on Robotics and Automation, Minneapolis, MN, USA, 22–28 April 1996, pp. 3395–3400. IEEE (1996). https://doi.org/10.1109/ROBOT.1996.509229
18. Mitsi, S., Bouzakis, K.D., Sagris, D., Mansour, G.: Determination of optimum robot base location considering discrete end-effector positions by means of hybrid genetic algorithm. Robot. Comput. Integr. Manuf. (2008). https://doi.org/10.1016/j.rcim.2006.08.003
19. Kamrani, B., Berbyuk, V., Wäppling, D., Stickelmann, U., Feng, X.: Optimal robot placement using response surface method. Int. J. Adv. Manuf. Technol. (2009). https://doi.org/10.1007/s00170-008-1824-7
20. Boschetti, G., Rosa, R., Trevisani, A.: Optimal robot positioning using task-dependent and direction-selective performance indexes: general definitions and application to a parallel robot. Robot. Comput-Integr. Manuf. (2013). https://doi.org/10.1016/j.rcim.2012.09.013
21. Siemens AG: Roboter - Intelligent platzieren. Process Simulate für Teamcenter 14.0. https://docs.plm.automation.siemens.com/tdoc/tecnomatix/14.0/PS_TC#uid:index_xid1015772:id12415110051341:KinematicsMenu_SmartPlace (2019)

Frequency-Based Identification of the Inertial Parameters of an Industrial Robot

L. Gründel[1(✉)], C. Reiners[1], L. Lienenlüke[1], S. Storms[1], C. Brecher[1], and D. Bitterolf[2]

[1] Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen, University, Aachen, Germany
l.gruendel@wzl.rwth-aachen.de
[2] Siemens, Frauenauracher Straße 80, 91056 Erlangen, Germany

Abstract. Due to disturbances or a lack of excitation during the measurements, conventional identification methods offer solutions with limited precision for the inertial parameters of industrial robots (IR). This paper introduces an approach to increase the rank of the identification matrix through additional equations from the frequency domain. In areas of lower frequencies, the total inertia that is affecting an axis is related to the amplitude of the frequency response of the rotational speed controlled system (RSCS). Another advantage of the presented method is the possible correction of friction effects via the phase information, which enables a higher identification accuracy. The frequency responses are measured during exciting trajectories, which stimulate low frequencies. Thereby, the approach generates additional equations, which enables the identification of more inertial parameters with a higher accuracy. In this paper, the measurement method and the identification algorithm are outlined.

Keywords: Identification · Modeling · Inertial parameters · Frequency response

1 Introduction

Machine tools represent an enormous investment risk for small and medium-sized enterprises (SMEs). Conventional IR offer the possibility of flexible machining of large components at significantly lower investments. However, due to the serial kinematics and the resulting high compliance, they usually achieve insufficient machining qualities. In the field of research, model-based simulation and control methods are used to increase the accuracy of IR in both planning and the process itself [14]. One approach is the derivation of equations of motion, which calculate the drive torques, that are needed to perform a certain movement, including the influence of intalk- and crosstalk forces within the mechanical structure of the IR. The torques can either be fed forward to the control loop during machining or used for the calculation

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 429–438, 2021.
https://doi.org/10.1007/978-3-662-62138-7_43

of predicted deviations during the process planning [5, 7, 10]. The accuracy of the equations of motion, which are usually derived from the recursive Newton–Euler approach, depends significantly on the inertial parameters.

This paper is structured as follows: In the following chapters, the current state of the art regarding the identification of inertial parameters and the basic idea of this paper is presented. Afterwards the theoretical background and the derivation of the necessary equations are explained. Then, the approach is validated followed by an evaluation and a conclusion.

2 State of the Art

In general, the identification topic for industrial robots is widely discussed in research journals over the past decades. The presented approaches in literature share the following similarities:

- Using a model that is linear regarding the inertial parameters
- Construction of an overdetermined system of equations that is generated with data points of a measured trajectory
- Parameter identification via linear regression techniques

Linear models following different methods like the energy model, the power model or the most commonly used inverse dynamics model were developed [3, 4, 8]. Those models are derived by calculating the systems energy, power or drive torques as outputs based on the linear relationships between the motion inputs (positions, velocities and accelerations) and the geometric and inertia parameters, while the geometric parameters are usually known. In order to construct an overdetermined system of equations, the IRs discrete motion values and drive torques for a given trajectory are measured during a predefined time interval and fed into the model equations, which leads to:

$$Y = W(q, \dot{q}, \ddot{q})\mathrm{X} + \rho \tag{1}$$

Where Y is the models output, W is the $(n \times m)$ observation matrix with $n \gg m$, X is the vector of inertial parameters and ρ is the residual error vector. A commonly used linear regression method is the least squares (LS) approach [12]. Besides LS, other regression techniques like the instrumental variable approach and output error methods were also used [2, 6]. The common goal of the various strategies is to identify inertial parameters which, when used in the selected model, provide results that best fit the measurements and are also physically plausible [15].

The system of equations in (1) can only be solved if W has a full rank. Due to a lack of excitation, a restricted workspace and invariant motion during the experiments, some equations provide redundant information because they linearly depend on other equations. Hence, (1) can be reduced to a set of so-called minimal base parameters, which means that some parameters can only be identified as linear combinations [11]. However, the results are still sensitive to noise. Vandanjon et al. already observed this problem in 1995 and tried to overcome those disadvantages by planning dedicated

trajectories to capture gravity, inertial, centrifugal and inertial coupling forces separately [13]. Rackl et al. attempted to solve the problem of poor excitation by parameterizing B-spline trajectories with a constraint optimization algorithm [9].

However, the presented approach of drawing the equations from the frequency response of each axis intrinsically eliminates the problem of noisy measurements and friction. Noise and zero-mean disturbances do not have an impact on the amplitude within areas of lower frequencies for measurements over several excitation periods of constant frequency. In addition, friction causes damping, which can be observed in the phase response and subtracted from the amplitude response (see Subsect. 3.2).

Current processes require complex derivation of the exciting trajectories. The presented approach offers a less complex way to get the ideal measurements to increase the rank of Eq. (1). The observation matrix here only depends on the axes' positions leading to a simpler optimization of the measurement poses.

The frequency response offers a wide range of information regarding the dynamical system, which is utilized by the presented measurement methods. The current results can be characterized as a first step to develop a new frequency-based method.

3 Theoretical Background

In the following chapters the derivation of equations is explained. The equations form the basis for the LS identification.

3.1 Modeling of Serial Kinematics

Without external forces acting on the end effector, the dynamic model for axis k of an open loop IR with n axes can be expressed as:

$$J_{AXk}\left(q_{k+1,\ldots,q_n}\right)\ddot{q}_\mathrm{k} + \tau_{ck}(\mathrm{q}_k, \dot{\mathrm{q}}_\mathrm{k}) + \tau_{\mathrm{gk}}(\mathrm{q}_k) + \tau_{\mathrm{fk}}(\mathrm{q}_k, \dot{\mathrm{q}}_\mathrm{k}) = \tau_k \tag{2}$$

Where $q_k, \dot{q}_k$ and $\ddot{q}_\mathrm{k}$ are the angular position, velocity and acceleration of joint k, $J_{AXk}\left(q_{k+1,\ldots,q_n}\right)$ is the total inertia effecting axis k, $\tau_{ck}(\mathrm{q}, \dot{\mathrm{q}})$ is the torque due to Coriolis and centripetal forces, $\tau_{gk}(\mathrm{q})$ is the gravity torques and $\tau_{\mathrm{fk}}(\dot{\mathrm{q}}_\mathrm{k})$ is the friction-related torque which together add up to the drive torque τ_k. For the presented experiments, just one axis is moved at a time by a maximum of 4°. Therefore, Coriolis effects do not occur and centripetal forces are absorbed by the bearings. The short trajectory and therefore small change of gravity forces leads to the assumption that gravity effects correspond to frequencies close to zero and therefore can be neglected here.

For an undamped system, the angular velocity would follow the drive torque with a phase shift of −90°. Since the analysis of the phase response reveals a deviation from this assumption, frictional damping effects cannot be neglected. There are various ways to describe friction with all its complex aspects [1]. An adequate and often used model is the combination of Coulomb and viscous friction [13]. Due to the constant velocity offset during the experiments, which causes a force excitation in a constant direction the discrete Fourier transformation of the Coulomb friction

would generate a spectral line at 0 Hz, so that only the viscous friction, represented by μ_{vk}, has to be included. The described assumptions lead to the following differential equation:

$$J_{AXk}\left(q_{k+1,\ldots,q_n}\right)\ddot{q}_{\mathrm{k}} + \mu_{vk}\dot{q}_k = \tau_k \tag{3}$$

The load-side total inertia of an axis is composed of the sum of the following axes' inertia tensors $J^k_{I,i}$, transformed into the corresponding joint coordinate system, the inertia due to the parallel displacement of the rotational axes described by the Huygens-Steiner theorem $J^k_{HS,i}$ and the load-side inertia of the motor $J_{mls,k}$. The latter can be expressed by the inertia of the motor on the drive side multiplied by square of the gear ratio u_k leading to the following equation:

$$J_{AXk}(q_{k+1},\ldots,q_n) = \sum_{i=k}^{n} z_k^{'}\left(R^k_{ki} I^i_i R^{k^{'}}_{ki} - m_i \tilde{v}^k_i \tilde{v}^k_i\right) z_k + u_k^2 J_{mms,k} \tag{4}$$

Where z_k is the vector of the rotational axis, R^k_{ki} is the rotation matrix of coordinate frame k to i and I^i_i is the inertia tensor described in coordinate frame i.

3.2 Systems Theory

To simplify the following equations, the index k is left out in the following explanations. The application of the Laplace transform to (3) leads to the transfer function:

$$\frac{\Omega}{\Gamma} = \frac{\mu}{\mu^2 + \omega^2 J^2} - i \cdot \frac{\omega J}{\mu^2 + \omega^2 J^2} \tag{5}$$

Decomposing (5) into amplitude and phase results in two equations:

$$\left|\frac{\Omega}{\Gamma}\right| = \frac{1}{\sqrt{\mu^2 + \omega^2 J^2}} \tag{6}$$

$$\varphi = \arctan\left(\frac{\Omega}{\Gamma}\right) = \arctan\left(-\frac{\omega \mathrm{J}}{\mu}\right) \iff \mu = -\frac{\omega \mathrm{J}}{\tan(\varphi)} \tag{7}$$

Inserting (7) into (6) leads to the final equation of the total inertia:

$$J = \frac{1}{\omega} \cdot \left|\frac{\Gamma}{\Omega}\right| \cdot \frac{1}{\sqrt{1 + \frac{1}{\tan(\varphi)^2}}} \tag{8}$$

Here $\left|\frac{\Gamma}{\Omega}\right|$ is the reciprocal of the amplitude and φ the phase of the measured frequency response of the sweep measurements.

4 Experiments and Validation

After introducing the theory to derive the necessary equations for the LS identification, the approach has to be validated. In order to reduce the complexity the following experiments focus on the first three axes (as seen from the base) while the last three axes are locked.

4.1 Measurement Setup

The used IR system consists of a MAX 100 of the company Mabi Robotics AG with a Sinumerik 840D sl control. The IR has six degrees of freedom and is equipped with direct and indirect encoders at each axis. The Computerized Numerical Control (CNC) is traditionally used for machine tools and therefore offers a wide range of testing methods for the dynamic behavior of the system, e.g.:

- Frequency response measurements of the RSCS for broader bandwidths
- And sweep measurements at discrete frequencies

The analysis of the first measurements provides the optimal frequency for the second experiment. The sweep measurements lead to more accurate amplitude responses for lower frequencies with shorter motions to perform.

As explained in Chap. 3, the amplitude of the RSCS is related to the total inertia that is acting at the observed axis. The Sinumerik enables the user to measure the frequency response of the RSCS for each axis individually. In order to reduce the influence of the control cascades, the control parameters are softened manually. During the experiment, the observed axis rotates with a speed offset while an interfering pseudo-random noise torque with specified frequency bandwidth is applied. The CNC simultaneously measures the torque via torque-related motor currents and the rotational speed using the encoders and then automatically calculates the frequency response of the RSCS (see Fig. 1). Numerous experiments with various test parameters have shown that a setting with bandwidths below 40 Hz, speed offset of 0.1 rpm and an amplitude of the excitation signal of 1% of the nominal torque generated sufficient results.

In theory the sweep measurement can be set up for arbitrary excitation frequencies. After examining the RSCS via the experiment explained above, these frequencies are selected as 2, 3, 4 and 5 Hz, below the first resonant frequency and thus lower than the frequency where the load decouples from the drivetrain. The measurement superimposes sine waves of the mentioned frequencies with a rotational speed offset, which is previously optimized by considering the periodic disturbances of the drive train (e.g. cogging forces). Therefore, a rotational speed profile is determined which leads to a target path of the axis position via integration. Hence, the control parameters are not softened for this experiment because the exciting frequencies are part of the target trajectory and the control does not treat them as interfering signals. During the measurement, the time courses of the drive torque and the rotational speed are measured and transformed into the frequency domain via the discrete Fourier transformation.

4.2 Initial Analysis of the Frequency Response

Figure 1a shows the amplitude responses for axis 1 for different poses of axis 2 and 3. An important landmark is the first locked-rotor frequency, which decreases from about 12 Hz to 5 Hz as the inertia increases. For the decade immediately below the resonance the system behaves like a first order pole. Elsewhere, this assumption is not valid. In any case, the effect of changing the robots pose and thus increasing its moment of inertia about axis 1, results in a decreased the first locked-rotor frequency.

If the torque and angular velocities are interpreted in terms of inertia the frequency response of the RSCS can be transformed leading to Fig. 1b. The graph illustrates how an ideal sweep frequency can be identified. The ideal frequency lies in between the resonant frequency and an area of lower frequencies where friction is affecting the amplitude for $f < 0.7Hz$). For a given pose the resonant frequency and the frequency at the intersection of the asymptotes of inertia- and friction-dominated areas (3 dB point) are calculated. The ideal frequency is the geometrical mean of those frequencies. In order to circumvent drive train harmonics while also maintaining a specified number of sine wave periods during the trajectory, the frequency is rounded to the next integer number of 2, 3, 4 or 5 Hz. This procedure is performed with different robot poses leading to ideal excitation frequencies for the sweep measurements described below.

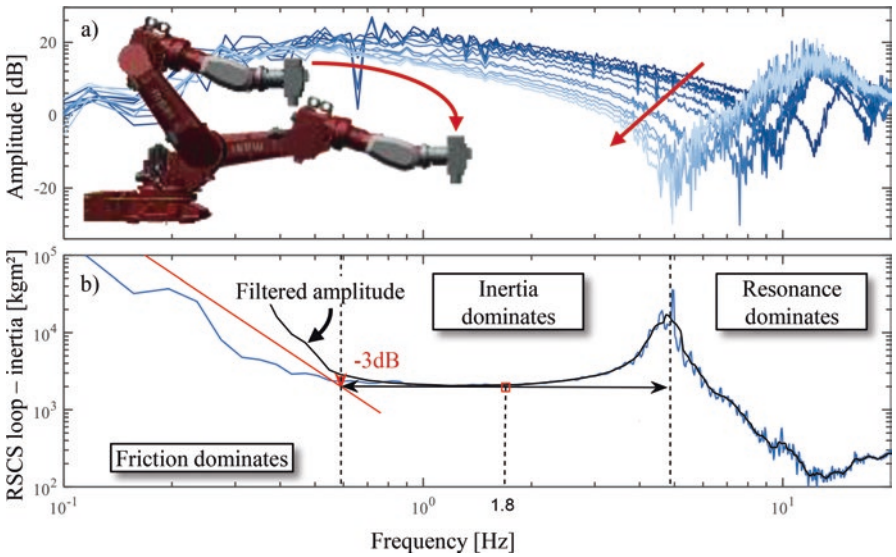


Fig. 1 Amplitude responses of axis 1 for various poses of axes 2 and 3 a) and the geometrical identification of the ideal sweep frequency b)

4.3 Sweep Measurements

The amplitude at a discrete frequency for Eq. (8) is provided by the sweep measurements. The measured total inertia and the corresponding position values are the inputs that generate new equations for the LS procedure. In order to add as many not linearly dependent equations as possible the pose of the IR has to vary. When attempting to select poses, a criterion for decision making is required. The standard procedure, which is also employed in this work, is to calculate the condition number of the identification matrix. The condition number allows the quantification of the maximum error of the identified parameters due to disturbances. To optimize the condition number an iterative procedure is used: Initially a list of all collision-free poses in a 5° grid that fit into the robot cell is generated. Then for each step, the pose, which results in the lowest condition number is added to the list.

Figure 2 illustrates the relationship between the positions of axes 2 and 3 and the total inertia measured at axis 1 for the measured data and the data generated with the reference model. The shown graph validates the presented approach as the measurements clearly match the expectable values regarding the inertial parameters provided by the manufacturer. In general, the logical relationship is that the measured inertia is higher the wider the distance between the following links of the IR and the center of rotation of the observed axis is.

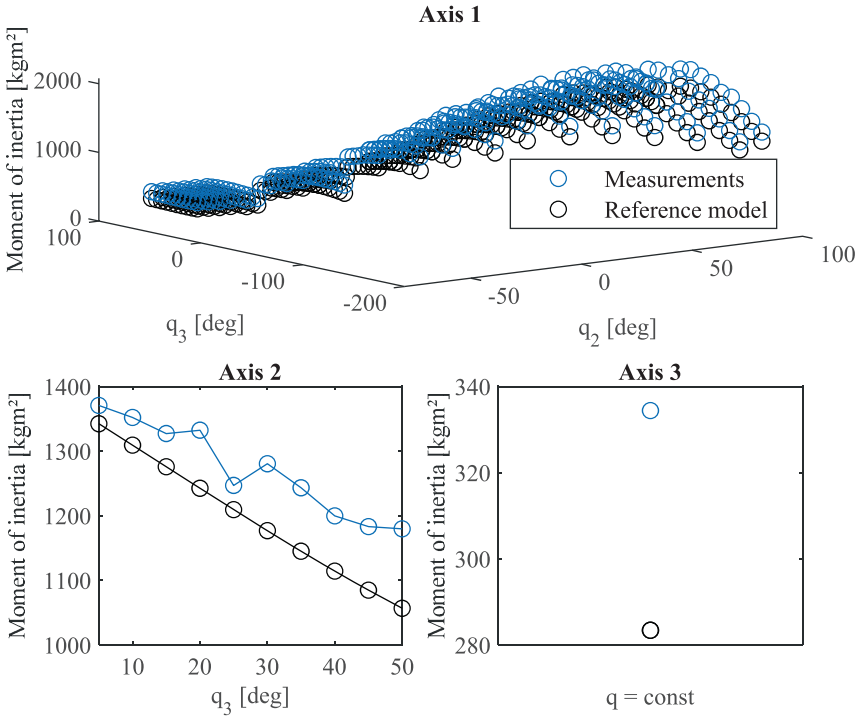


Fig. 2 Correlation of measured and reference data

5 Evaluation

In general, the whole approach of calculating new poses for the measurements proved to be comprehensible and efficient. The optimization leads to over 300 different poses for the identification. After performing the measurements the system of equations for the LS procedure was constructed. As described before, the last three axes were locked, which allows the combination of the last four links and the end effector of the robot into one body. The resulting 30 parameters were transformed into a set of eight base parameters. After carrying out the standard LS the identified set of base parameters is inserted into the model equations and compared to the model based on the reference parameters of the manufacturer. A frequently used measure for the differences between model values and a measurement is the root-mean-square error (RMSE).

In Table 1 the RMSEs between the LS fitted model, the reference model and the measurements at the various poses are listed. The high deviations between the reference model and the measurement can probably be traced back to the CAD calculation of inertial parameters, which only considers the large casting components of the IR. In addition, the reference model does not take the additional structures like the assembled ducts and hose packages into account, which results in a significantly larger RMSE for all axes. Summing up, the identified LS fitted model can map the measured inertias for the different poses adequately. The significantly lower RMSE values indicate an improved model accuracy. The last line in Table 1 illustrates the models ability to depict the physical behavior of the IR. In contrast to axes 1 and 3, the increase in RMSE of axis 2 indicates a poorer conditioning of the observation matrix which could be optimized by further poses in the future.

Table 1 Comparison of the RMSEs for the LS fitted model, the reference model and the measurements

	Axis 1	Axis 2	Axis 3
RMSE [%]: reference model – measurement	19.75	5.54	15.52
RMSE [%]: LS identified model – measurement	1.61	1.33	0.10
RMSE [%]: reference model – LS identified model[a]	20.05	12.38	15.52

[a]For this calculation of the RMSE between the reference model and the LS identified model additional poses were added to the measurement poses in order to demonstrate the improvements of the identified model for the entire workspace

Similarly good results are expected when extending the approach to the other axes of the robot. Nevertheless, further equations generated by additional exciting trajectories would provide information that could be used to identify more parameters. However, the presented approach does lead to a reduced model, which accurately matches the dynamics of the IR.

6 Conclusion

Current methods to identify masses and inertias of IR links are performed with a model consisting of differential equations, exciting trajectories and linear regression algorithms. However, those methods fail to identify the parameters properly because the algorithms are sensitive to noise and effects like friction are difficult to reproduce in a model. The presented approach of a frequency response based identification of inertial parameters delivers promising results. Due to the analysis in the frequency domain and the possibility of frequency selective measurements, the procedure offers the decisive advantage of independency to noise and a simplified process for the generation of exciting trajectories. In conclusion, the identified set of base parameters properly matches the reference and measurement data for the first three axes.

Currently the procedure is extended to all six axes. In addition, the authors plan to add more information through further exciting trajectories. The combination of the presented approach with additional equations for the LS procedure leads to an identification routine for six-axis IRs.

References

1. Bona, B., Indri, M.: Friction compensation in robotics: an overview. In: Proceedings of the 44th IEEE Conference on Decision and Control. 44th IEEE Conference on Decision and Control, Seville, Spain, 12–15 Dec. 2005, pp. 4360–4367. IEEE (2005). 10.1109/CDC.2005.1582848
2. Brunot, M., Janot, A., Carrillo, F., Cheong, J., Noël, J.P.: Output error methods for robot identification. J. Dyn. Syst. Meas. Contr. **1**(3), 428 (2020). 10.1115/1.4045430
3. Featherstone, R.: Rigid Body Dynamics Algorithms. Springer, New York (2008)
4. Gautier, M.: Dynamic identification of robots with power model, 1922–1927 (1997). 10.1109/ROBOT.1997.619069
5. Gründel, L., Lienenlüke, L., Storms, S., Brecher, C.: Compensation of process forces with a model-based feed-forward control for robot machining. In: 4th International Conference on Control and Robotics Engineering (ICCRE), 115–121 (2019). 10.1109/ICCRE.2019.8724312
6. Janot, A., Olivier Vandanjon, P., Gautier, M.: An instrumental variable approach for rigid industrial robots identification. Control Eng. Pract. **25**, 85–101 (2014). 10.1016/j.conengprac.2013.12.009
7. Mesmer, P., Neubauer, M., Lechler, A., Verl, A.: Drive-based vibration damping control for robot machining. IEEE Robot. Autom. Lett. **5**(2), 564–571 (2020). 10.1109/LRA.2019.2960723
8. Presse, C., Gautier, M.: New criteria of exciting trajectories for robot identification. In: Proceedings IEEE International Conference on Robotics and Automation. IEEE International Conference on Robotics and Automation, Atlanta, GA, USA, 2–6 May 1993, pp. 907–912. IEEE Comput. Soc. Press (1993). 10.1109/ROBOT.1993.292259
9. Rackl, W., Lampariello, R., Hirzinger, G.: Robot excitation trajectories for dynamic parameter estimation using optimized B-splines. In: 2012 IEEE International Conference on Robotics and Automation. 2012 IEEE International Conference on Robotics and

Automation (ICRA), St Paul, MN, USA, 14.05.2012–18.05.2012, pp. 2042–2047. IEEE (2012). 10.1109/ICRA.2012.6225279
10. Schnoes, F., Zaeh, M.F.: Model-based planning of machining operations for industrial robots. Procedia CIRP **82**, 497–502 (2019). 10.1016/j.procir.2019.04.331
11. Siciliano, B., Khatib, O.: Springer Handbook of Robotics. Springer International Publishing, Cham (2016)
12. Swevers, J., Verdonck, W., de Schutter, J.: Dynamic model identification for industrial robots. IEEE Control Syst. **2**(5), 58–71 (2007). 10.1109/MCS.2007.904659
13. Vandanjon, P.O., Gautier, M., Desbats, P.: Identification of robots inertial parameters by means of spectrum analysis. In: Proceedings of 1995 IEEE International Conference on Robotics and Automation. 1995 IEEE International Conference on Robotics and Automation, Nagoya, Japan, 21–27 May 1995, pp. 3033–3038. IEEE (1995). 10.1109/ROBOT.1995.525715
14. Verl, A., Valente, A., Melkote, S., Brecher, C., Ozturk, E., Tunc, L.T.: Robots in machining. CIRP Ann. **6**(2), 799–822 (2019). 10.1016/j.cirp.2019.05.009
15. Yoshida, K., Khalil, W.: Verification of the positive definiteness of the inertial matrix of manipulators using base inertial parameters. Int. J. Robot. Res. **19**(5), 498–510 (2000). 10.1177/02783640022066996

Increasing Efficiency in Maintenance Processes Through Modular Service Bundles

J. Fuchs[1(✉)], H. Herrmann[1], S. J. Oks[2], M. Sjarov[1], and J. Franke[1]

[1] Institute for Factory Automation and Production Systems, Friedrich-Alexander University, Erlangen-Nuremberg, Egerlandstraße 7-9, 91058 Erlangen, Germany
jonathan.fuchs@faps.fau.de

[2] Chair of Information Systems, Innovation and Value Creation, Friedrich-Alexander-Univer-Sity Erlangen-Nuremberg, Lange Gasse 20, 90403 Nuremberg, Germany

Abstract. Increasing complexity in production and factory automation represents a significant challenge in context of maintenance. One approach to address complexity is the implementation of automated industrial service bundles, which resemble complex business processes. To enable efficient usage of service bundles, their individual components are designed as functional modules in the form of independent micro services and are compatible with the paradigm of service-oriented architectures. For ubiquitous communication and data exchange between service entities, message-oriented middleware provides an adequate solution. This paper presents an approach for increasing efficiency in maintenance using a service bundle. The use case demonstrates the automated creation of service tickets enriched with necessary information from various sources.

Keywords: Smart services · Service bundle · Smart maintenance · Process automation

1 Introduction

The system-inherent complexity in cyber-physical production systems (CPPS) requires consequent and holistic automation integrating physical assets as well as processes from technical and value creation domains. This process is supported by efficient networks combining both information technology (IT) and operational technology (OT). These mainly software-based networks facilitate the implementation of services and their combination in service bundles. The concept of service bundles is used to provide individual and purpose-oriented services which meet the needs of specific application scenarios by combining fine granular building blocks of different services independent of their providers [1]. These combined processes resemble complex value creation processes. However, the existence of numerous monolithic

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 439–447, 2021.
https://doi.org/10.1007/978-3-662-62138-7_44

legacy systems is a major challenge for extensive automation as their proprietary elements are usually not developed according to certain aspects, e.g., modularization and microservice patterns, which hinders the reuse of software modules for multiple objectives [2].

For purposeful interaction between software-based services and components within production plants, access to asset information is essential. However, CPPS and their components are usually not fully linked to an associated information model or even a comprehensive virtual representation [3]. Therefore, asset information is difficult to access and must be obtained and stored by workers in effortful and mostly manual procedures. As a result, asset data is often aggregated incompletely and cannot be used easily in automated service environments. Furthermore, solely IP-based communication at shop floor level often misses requirements for Industrie 4.0 (I4.0)-compliant asset integration. A significant barrier is the lack of semantics in industrial communication. This requires information models that machines can both comprehend and process [4]. Established IT communication patterns have to be transferred to the OT world to enable semantically integrated communication and thereby ensure fully automated and complete data transfer as well as autonomous interaction between assets based on semantics. To enable holistic data exchange for services in production networks, viable I4.0-compliant infrastructures based on industrial platform solutions are first to be established. Virtual representations of assets are integrated there, allowing service bundles to develop their potential.

This paper presents an approach for implementing an automated maintenance process utilizing such service bundles. It is structured as follows: Sect. 2 briefly explains prerequisites for and the utilization of service bundles in digital industrial service systems (DISS). Therefore, digital industrial services are introduced (Sect. 2.1), application integration styles are presented (Sect. 2.2) and potentials for value creation by service bundle applications in maintenance (Sect. 2.3) are discussed. Section 3 presents an approach for the automation of maintenance aspects utilizing service bundles. The segmentation of value creation processes to service bundles (Sect. 3.1), the design of an automated maintenance service bundle (Sect. 3.2) and the application case-driven implementation (Sect. 3.3) are demonstrated. Section 4 summarizes the proposed approach.

2 Enabling Service Bundles in Digital Industrial Service Systems

Process optimization in the industrial context is mostly enabled by software-based services. Several services can be joined into service bundles. Such bundles handle complex business processes by efficiently combining individual service capabilities [5]. According to the paradigm of service-oriented computing (SOC), service composition can be achieved in two main ways: service orchestration and service choreography. Service orchestration includes a central controlling unit which orchestrates all services involved to achieve a desired business objective. In service choreography, services collaborate without a central unit controlling the overall procedure [6]. To

enable service bundles in industrial environments, DISS may serve as a framework. DISS describe non-hierarchical systems based on information and communication technology (ICT) and are paving the way for digitally modified and novel digital business models [7, 8]. In the following, ICT-related concepts are described in more detail, which are necessary for the implementation of independent services, their combination in service bundles and integration in I4.0-compliant production networks.

2.1 Digital Industrial Services

Modern production networks must meet a wide range of requirements in terms of technical performance and communication capabilities to generate added value in conjunction with software-based services. In such environments, digital services unfold their potential, especially if they have been developed according to aspects that emphasize modularization. Central elements are service-oriented architectures (SOA) and microservice patterns. The SOA paradigm aims at the transformation of monolithic IT systems into networks of distributed services [8]. These services can be designed following the microservice pattern [9]. Microservices are an architectural pattern that can be used for the development of complex software-based applications by combining a set of independent, small service entities which perform individual tasks. These entities communicate with each other based on different integration styles (see Sect. 2.2). The services are independently deployable and thus enable a modular structure of application software [10].

In addition to technical issues, efficient integration of human operators in DISS is a key challenge. Even if automated decision algorithms are already far advanced, various, often strategic choices can only be taken by humans, because of their implicit expert knowledge in their field. The task of the DISS is then to provide the user group with a clear and helpful basis for decision-making. In addition, the way of the user integration must be clearly defined.

2.2 Application Integration Styles

Continuous data exchange between service entities is an essential prerequisite for proper service interaction. Communication developed from mostly file-based exchange in the early stages of industrial automation towards solutions using message-oriented middleware (MOM) [11]. MOM coordinates the data transfer to various recipients and serves as a mediating instance. When application integration is required, different integration styles can be distinguished. According to Hohpe and Woolf these comprise file transfer, shared databases, remote procedure invocation and messaging, as depicted in Fig. 1. The stated order of these integration styles implies an order of increasing complexity when it comes to the realization, from file transfer being the simplest to messaging as a more sophisticated method [12].

Despite messaging being the most complex of the four stated integration approaches, it carries with it a set of advantageous properties: The data exchange between the sending and receiving applications is asynchronous, avoiding the blocking that might occur when remote procedure calls would be utilized. This

decouples sender and receiver, which is an essential property for distributed systems. Furthermore, the messaging mechanism can transform the message to comply with specific data formats [13]. The MOM further queues and orders messages, ensures their integrity and traces the success of the delivery [14]. In addition to the application integration style, semantics of provided data is a crucial success factor for automated inter-machine interaction. One enabler for this is the industrial communication standard Open Platform Communication Unified Architecture (OPC UA). With OPC UA, information can be modelled, coded, and interpreted by computers. For example, not only data values of a sensor can be represented, but also context information about the sensor type or the entire asset can be mapped [15]. The selection of a specific data format, e. g. Protocol Buffers, JSON or XML, is no longer of importance and can be done according to a best practice principle, since different formats are suitable for different tasks.

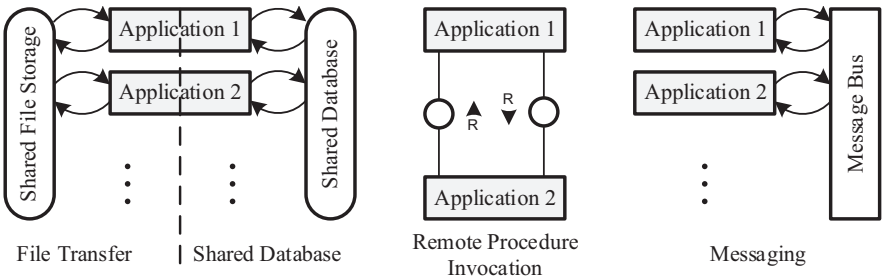


Fig. 1. Four different means of application integration according to Hohpe and Woolf [12].

2.3 Potentials for Value Creation by Service Bundle Application in Maintenance

The introduction of digital technologies and concepts, such as CPPS, brings great potential for optimizing and increasing the output of value creation activities including maintenance [16]. The gain of additional information based on continuously acquired and monitored values of the extensive sensor infrastructure and their evaluation by big data analytics-supported pattern recognition provides a fine-grained and detailed status overview of the machines, production systems and manufacturing parts. This information can be used to identify critical conditions and trends in the context of predictive maintenance and to prevent malfunctions through preventive maintenance measures. Moreover, other systems such as manufacturing execution systems (MES) or enterprise resource planning (ERP) can be integrated into CPPS-based maintenance systems. In this way, further information for the coordination of the overall value creation process is available, such as the priorities of orders and the accessibility of resources like personnel, spare parts and tools. This enables a holistic prioritization of maintenance tasks [17]. To support this, machine manufacturers and service providers are increasingly offering corresponding services and solutions for their systems. However, since in most cases equipment and software from a variety of suppliers are used, these encapsulated services must be broken up to be reconfigured

and combined according to the company-specific application scenario. This can be achieved through the utilization of service bundles since the underlying concept enables providing individual and purpose-oriented services by combining fine granular building blocks of different services independent of their providers [1].

3 Automation of Maintenance Aspects Utilizing Service Bundles

3.1 Segmentation of Value Creation Processes into Service Bundles

Value creation processes can be optimized, especially regarding time and cost factors. To achieve this, structure and dependencies of business processes must be thoroughly analyzed. Business and value creation process segmentation is depicted exemplary in Fig. 2. The systematic breakdown is used to identify separate operational units characterized by individual components which can be represented by linked service entities. Therefore, it is essential to examine which communication patterns and interfaces are necessary, suitable, and available for the digital transformation, modularization, and automation of such processes.

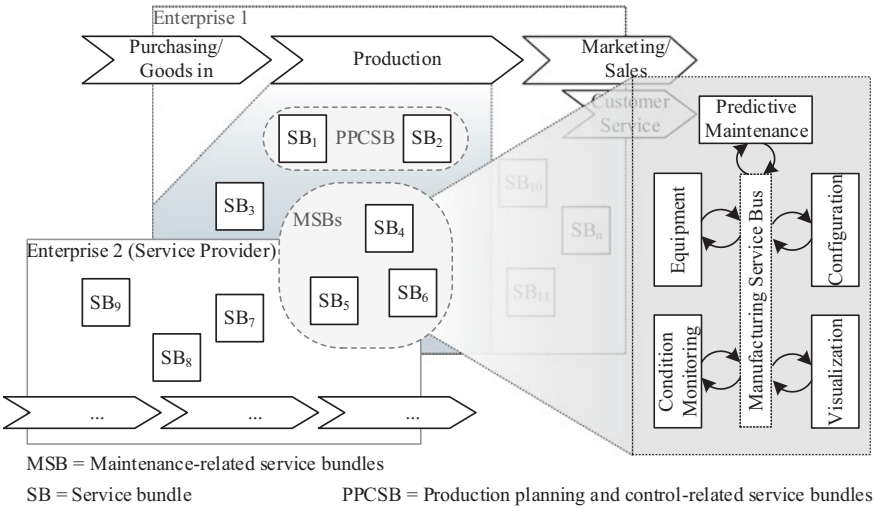


Fig. 2. Decomposition of value creation processes to service bundles.

The use case considered in the following is located within the area of maintenance. There, the possibilities of semantic machine communication developed in the context of I4.0 offer the potential to increase the degree of automation and at the same time reduce the inherent complexity. As CPPS and their components are usually not entirely linked to an information model, asset information is generally challenging to access and has to be obtained and stored by workers in mostly manual processes.

As a result, asset data is often aggregated incompletely. In contrast, automated information transfer between services ensures consistency and accuracy of required data, while the once-only manual effort to create a valid information model for virtual representation ensures completeness. To demonstrate the implementation and advantages of industrial service bundles, we use the automated creation of service tickets in the event of equipment failure as a use case. The creation of service tickets in maintenance management systems (MMS) to coordinate maintenance tasks is common. However, the collection of necessary information required by the ticket agent is often time-consuming, as manual data aggregation is predominant. These disadvantages can be compensated by service bundles. The following chapters explain the process mapping and the integration of maintenance personnel.

3.2 Design of an Automated Maintenance Service Bundle

The automated service ticket creation process involves several software components depicted in Fig. 3. Based on the mentioned use case, a system architecture in which services as well as the user can be integrated and linked in terms of ICT is introduced. The system explanation is inspired by the *4+1 view model* by Kruchten. This model has been developed to describe software-based system architectures based on several complementing views, which allow the overall system to be viewed from the perspective of different stakeholders. Scenarios can be included to better represent the architecture, which is what the "+1" is supposed to express. Regarding the presented application, the scenario view i.e. the use case-driven implementation is elaborated in Sect. 3.3 [18].

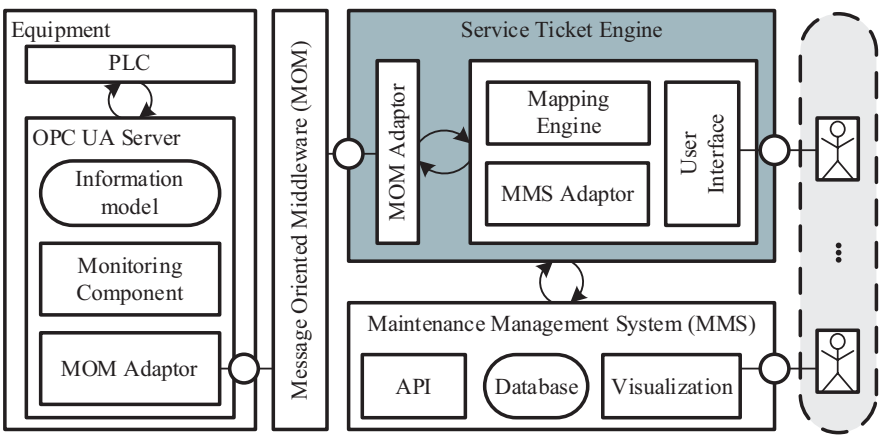


Fig. 3. Components of the maintenance service bundle in an automated process.

The **Logical View** deals with end user-related system functionalities. The task of this service bundle is to automatically detect fault conditions in equipment, to aggregate all necessary data and context-related information as quickly as possible and to provide it as a defined fault data set. This set is forwarded to downstream services

via MOM. In this specific example, the data object is further processed by the *service ticket engine* module and imported into a maintenance management system as a service order. Here, the information can be visualized and further processed by maintenance personnel.

The **Physical View** focuses on architectural elements, e.g., the distribution of components on physical level and the communication between these components. We favor container-based deployment of software. This allows greater independence from specific hard- or software requirements. The data aggregation and provisioning take place directly at machine level using the monitoring component of the integrated OPC UA server. All other software elements are not bound to defined locations and can be deployed in the sense of a distributed system.

The **Development View** describes the system from developer's perspective and deals with software management. Due to strong modularization and application of the microservice development pattern, individual components can be replaced quickly, and new ones integrated easily.

The **Process View** describes non-functional elements and dynamic aspects. The corresponding process is depicted in Fig. 4. During production, faults repeatedly occur on equipment whereby the service bundle is activated. The condition monitoring component of the OPC UA server continuously monitors the production process. Once the machine reaches an error state, the corresponding error information is gathered and mapped via an error data set and forwarded to the service ticket engine. From here, a distinction can be made between two process variants: In one case, the error data is assigned to data fields of an MMS automatically. In the other case, maintenance personnel is integrated as decision makers. Then, data forwarding to the MMS is only executed after human consideration. If the process is not terminated by the operator, the service ticket engine develops a data set suitable for the MMS. Based on this, a request for the creation of a service order is then automatically submitted to the internal or external service provider.

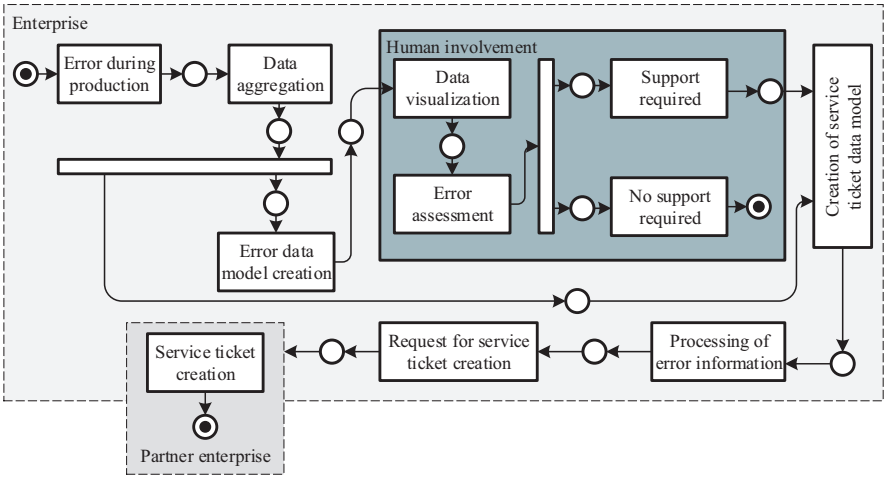


Fig. 4. Process view of the proposed service bundle.

3.3 Use Case-Driven Implementation

The prototypical implementation considers a machine for hot crimping of cable lugs on high-voltage cables [19]. Typical error conditions are: (a) temperature derivations, (b) deviation of the clamping force or (c) fluctuations in energy consumption. The static and dynamic, i.e. process-related parameters are represented by an OPC UA information model.

Automated process: In case of an error occurrence during production, the OPC UA server generates a data set based on the information model which includes all necessary asset and error-related data. This set is mapped and published as a JSON object via web application messaging protocol (WAMP). The object is then received by the MOM adaptor of the service ticket engine. During the following steps, the elements of the JSON object are assigned to the ticket request variables. The MMS adaptor then initiates the creation of the service ticket by sending a HTTPS request to the API of the MMS. The MMS stores the ticket in a database.

Human involvement: Human users can easily be integrated as decision makers: when an error occurs during production, it is visualized as an alarm in the frontend. There, the user can confirm whether the creation of a service ticket is necessary or whether the problem can be solved without an external service request. In this mode, the OPC UA server creates an extended alarm data set which contains all data necessary to make the decision. This data set is transferred to the service ticket engine. There, the information is visualized in the frontend, where maintenance personnel then confirms or declines the ticket request.

4 Conclusion and Outlook

This contribution demonstrates how process optimization in the context of maintenance can be achieved by the implementation of an automated industrial service bundle. The services in this bundle are designed as modules in the form of independent micro services which use MOM for data exchange. The presented use case demonstrates the automated creation of a service ticket within a maintenance management system in the event of an error occurring during production. Information modeling in OPC UA is applied for the digital representation of assets and processes. The OPC UA server receives occurring error and alarm messages. This triggers the automated error-relevant data transfer via MOM to a service ticket engine, which processes the data and creates the actual service ticket. In this way, cost and time savings can be realized. Furthermore, the automated information transfer ensures the completeness of required data. The service bundle can be used in cross-company networks. It is compatible with any web-based ticket management system that provides a compliant API. The degree of automation can be adapted according to individual needs.

Acknowledgements. The research presented in this paper was funded by the German Federal Ministry of Education and Research in the course of the project PRODISYS (FKZ 02K16C056).

References

1. Herterich, M.M., Uebernickel, F., Brenner, W.: The impact of cyber-physical systems on industrial services in manufacturing. Procedia CIRP **30**, 323–328 (2015)
2. Serrano, N., Hernantes, J., Gallardo, G.: Service-oriented architecture and legacy systems. IEEE Softw. **31**(5), 15–19 (2014)
3. Uhlemann, T.H.J., Lehmann, C., Steinhilper, R.: The digital twin: realizing the cyber-physical production system for industry 4.0. Procedia CIRP **61**, 335–340 (2017)
4. Fuchs, J., Schmidt, J., Franke, J. et al.: I4.0-compliant integration of assets utilizing the asset administration shell. In: Proceedings, 2019 24th IEEE International Conference on Emerging Technologies And Factory Automation (ETFA): Paraninfo Building, University of Zaragoza, Zaragoza, Spain, 10–13 September, 2019, pp. 1243–1247. IEEE, Piscataway, NJ (2019)
5. Heutschi, R.: Serviceorientierte Architektur. Springer, Heidelberg (2007)
6. Josuttis, N.: SOA in der Praxis, 1. Aufl., korr. Nachdr. Dpunkt-Verl., Heidelberg (2009)
7. Bruhn, M., Hadwich, K.: Dienstleistungen 4.0, 1. Aufl. Springer, Wiesbaden (2017)
8. Fuchs, J., Oks, S.J., Franke, J.: Platform-based service composition for manufacturing: a conceptualization. Procedia CIRP **81**, 541–546 (2019)
9. Newman, S.: Building Microservices, 1st edn. O'Reilly Media, Sebastopol (2015)
10. Fowler M., Lewis J.: Microservices: a definition of this new architectural term (2014). https://martinfowler.com/articles/microservices.html. Accessed 6 Apr 2020
11. Sommer, P., Schellroth, F., Fischer, M. et al.: Message-oriented middleware for industrial production systems. In: Vogel-Heuser, B. (Hrsg.). 2018 IEEE 14th International Conference on Automation Science And Engineering (CASE), 20–24 August 2018, pp. 1217–1223. IEEE, Piscataway (2018)
12. Hohpe, G., Woolf, B.: Enterprise Integration Patterns, 9ª reimp. Addison-Wesley signature series, Boston (2015)
13. Razzaque, M.A., Milojevic-Jevric, M., Palade, A., et al.: Middleware for internet of things: a survey. IEEE Internet Things J. **3**(1), 70–95 (2016)
14. Yongguo, J., Qiang, L., Changshuai, Q. et al.: Message-oriented middleware: a review. In: 2019 5th International Conference on Big Data Computing and Communications (BIGCOM), pp. 88–97. IEEE (2019)
15. Graube, M., Hensel, S., Iatrou, C. et al.: Information models in OPC UA and their advantages and disadvantages. In: 22nd IEEE International Conference on Emerging Technologies and Factory Automation (ETFA), pp. 1–8. IEEE (2017)
16. Oks, S.J., Fritzsche, A., Möslein, K.M.: An application map for industrial cyber-physical systems. In: Jeschke, S., Brecher, C., Song, H. (eds.) Industrial Internet of Things, 21–47. Springer, Cham (2017)
17. Jantunen, E., Zurutuza, U., Ferreira, L. L. et al.: Optimising maintenance. In: 3rd International Workshop on Emerging Ideas and Trends in Engineering of Cyber-Physical Systems (EITEC), pp. 53–58. IEEE, Piscataway (2016)
18. Kruchten, P.B.: The 4+1 view model of architecture. IEEE Softw. **12**(6), 42–50 (1995)
19. Kuehl, A., Zitzelsberger, M., Seefried, J. et al.: Hot crimping through innovative inductive heating in the production of electric motors. In: 2019 IEEE International Electric Machines & Drives Conference (IEMDC), May 12–15, Westin San Diego, San Diego, CA, pp. 1404–1409. IEEE, Piscataway (2019)

Domain-Specific Language for Sensors in the Internet of Production

M. Bodenbenner(✉), M. P. Sanders, B. Montavon, and R. H. Schmitt

Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, 52074 Aachen, Germany
m.bodenbenner@wzl.rwth-aachen.de

Abstract. A main prerequisite to applications in the Internet of Production is the integration of sensor data into an interconnected infrastructure, which in turn requires expert knowledge of sensor implementation as well as of network architecture design and communication protocols. To reduce complexity in this concern, the authors propose the *SensOr Interfacing Language (SOIL)*, a domain-specific programming language for sensor interface definition and exchange of metrological data. Based on a meta-model, the functional interface can be designed without prior knowledge of the underlying communication details. It is composed of instances of *components, parameters, functions* and *measurements* as core elements of SOIL. Subsequently, the interface is automatically defined on protocol level and its software implementation is generated, leaving only the hardware-specific implementation to the developer. The domain-specific language is prototyped and evaluated by implementing and integrating interfaces for a virtual laser tracker and a distributed temperature measurement system, confirming the envisaged benefits.

Keywords: Sensor · Internet of production · Domain-specific language · Data analytics

1 Introduction and Related Work

The *Internet of Production (IoP)* is a novel framework aimed at increasing multiple aspects of productivity, e.g. efficiency, degree of automation and resilience by means of improved data management, communication, analysis and exploitation for decision-making [1]. As for cyber-physical production systems in general, the integration of sensors and their metrological data as source of real-world information is indispensable. The increasing complexity of applications designed within the IoP and the subsequently emerging need for interoperability have led to the proposal of the *sensor information as a service* concept [2]. However, the latter is affected by the plethora of sensor hardware respectively devices with heterogeneous interfaces and communication protocols typical to production networks.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 448–456, 2021.
https://doi.org/10.1007/978-3-662-62138-7_45

Applications in the IoP are furthermore characterised by relying on knowledge contributions from different domains, i.e. the integration of sensor data concerns the domains of metrology for physical acquisition, software engineering for data transmission and storage, production engineering for context-specific exploitation as well as communication sciences for human-centric visualization. This requirement is reinforced by the tendency of merging the physically sensing device with the communicating network node to form cyber-physical components [3]. Their complexity results in an elevated development effort relying on domain knowledge in metrology and software engineering to integrate new sensor into an existing infrastructure. This contradicts the paradigm of *separation-of-concerns* [4], indicating that decoupling development and thus domain knowledge regarding the core sensor device and communication interface could reduce time and effort in software development and maintenance. The same intent can be attributed to the semantic *Middleware+* [5] layer of the IoP framework accommodating abstracted data provisioning and to the implementation opaqueness characteristic to service-oriented approaches.

The authors address this issue specific to sensors respectively the domain of production metrology by proposing a novel domain specific programming language (DSL). Such languages are characterised by high expressiveness and concise applicability. Furthermore, these languages have a high level of abstraction and reduced complexity compared to general purpose languages, such as C++or Java, which results in increased productivity and reduced maintenance costs for relying applications [6]. Most DSLs have their own syntax but are not directly executable and instead translated to a general-purpose language relying on existing compilers and software ecosystems.

To address experts from domains other than software engineering, it is advantageous to design a graphical representation, which reduces perceived complexity and simplifies usage [7] as demonstrated by the Midgar platform [8]. DSLs for interfacing and interconnecting devices in industrial context have been published under the names IoTDSL [9] and GIMLE [10]. OPC UA provides a modelling framework coupled to its communication protocol specification, and has been evaluation in conjunction with a DSL approach in literature [11]. While the physical capabilities of the considered devices can be modelled in higher detail, metrological concerns are not covered. Generally, most approaches, such as SEAL [12], GIMLE or Midgar have a strong focus on usability and reduction of complexity, but missing central concepts for ensuring data integrity and traceability which are essential for metrological applications [13].

2 Sensor DSL Requirements

From the domain perspective of production metrology, the primary intents of an interconnected sensor device are to offer a set of required interactions and unambiguously communicate a measurement result, i.e. representing physical meaning, unit and uncertainty. Providing interoperability is regarded as a requirement, but not as primary development focus or expertise to the domain of production metrology itself.

Communication protocols, used databases and available network technologies are likely to vary depending on the target environment as well as due to ongoing standardization and technological advances, such that their interchangeability, up to the extent of protocol routing, is desired [14]. This demand coincides with the need for abstraction in the sense that protocol- and communication-specific parts of the sensor service implementation are hidden to the DSL's users to achieve separation of concerns and reduced cost-of-learning [15]. Moreover, the need for a domain- respectively device-specific resource model in addition to data formats and communication technologies prevails to fully address interoperability [16]. Thus, one objective is to constitute a protocol-agnostic modelling language for the aforementioned resource model of a sensor. The generation of protocol-specific code is expected to be inherent part of the automated DSL translation process.

Previous studies on interfacing Large-Scale Metrology systems have revealed that a technologically heterogeneous set of sensors with a common measurand, e.g. spatial coordinates, share a majority of characteristics [17]. Consequently, standardization and reusability can be amplified by modelling the sensor's capabilities as a set of resources with defined actions encapsulated as service embracing a functional rather than a physical point of view. Another finding is the requirement to extend or reduce parts of the sensor service during runtime. For example, units be added to or removed from a device with multiple distributed sensing units during operation. Access patterns considered typical to sensor data comprise on demand retrieval as well as processing a continuous data stream. Moreover, the unambiguous representation of measurement results allows to leverage cryptographic and distributed ledger techniques for traceability.

Finally, the amalgamation of the communication interface generated on the basis of a model expressed in the DSL and programmatic sensor hardware access must be permitted. As this is only possible via sensor- and manufacturer-dependent application interfaces, generalization is not viable. Those interfaces are mainly available for general-purpose languages (GPLs), e.g. C, C++ and Python. A DSL for the considered domain and purpose must offer the possibility to manually implement parts of the application in a general-purpose language and inject these into generated source code. Consequently, the DSL must be translated into those GPLs, which are most commonly used in the considered domain.

3 Sensor DSL Concept and Prototype

To address the problem of separation-of-concerns across domains and respect the aforementioned requirements to a DSL in the domain of production metrology, the authors propose the *SensOr Interfacing Language (SOIL)* as novel DSL. It is centred around a functional model for sensor services consisting of four basic building *elements*: *components*, *functions*, *parameters* and *measurements*. A defined model forms the static, syntactical description of the sensor service and has tree-like shape. The underlying meta-model is illustrated in Fig. 1.

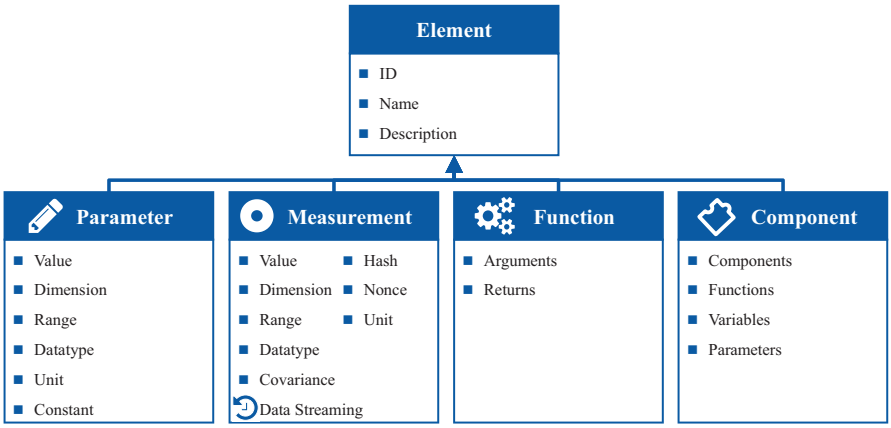


Fig. 1. The four basic elements of the meta-model of SOIL: parameter, measurement, function and component. For each element, the fields specified are required to design an unambiguous interface description.

Element is the base class of all tree nodes. Each element has a unique identifier, a human readable name and a short description of its meaning. Identifiers are locally unique and become globally unique by prepending the parents' identifiers successively.

Measurements are used as representative for all information physically sensed by the sensor. This includes the intended purpose, like temperature or distances, but also information like battery level or signal strength. The inherited value cannot be altered externally in contrast to parameters. Following Evertz et al. [16], physical meaning of the values is persisted by defining the data type, the allowed range, the mathematical dimension and the physical unit. Currently, each measurement is of one of the following primitive datatypes: *Boolean*, *Integer*, *Double*, *String*, *Enum* or *Time*. Measurements also allow for the additional field *covariance*, applicable for consistent uncertainty expression. Data integrity and traceability are promoted by specifying a *nonce* and generating a *hash* of the measurement data.

Parameters cover all data not directly measured by the device but required for operation or user interaction. Examples are calibration data, addresses for internal communication or meta-information. It has the same fields as measurement, except covariance, hash and nonce. The Boolean field *constant* indicates unmodifiable parameters.

Functions are used for triggering complex tasks or changing the inner state of the sensor. A function generally accepts an ordered set of *argument parameters* and returns an ordered set of *return parameters*. Both, arguments and returns, are constant parameters.

Components are structural elements of a sensor model. Each component contains an arbitrary number of children elements such that the overall model has tree-like shape.

To meet the requirement of interface modifications during runtime, SOIL allows for specification of *dynamic components*. Components marked as dynamic can be considered as type definitions, which can be initialised multiple times with varying name, identifier and description of the component, as well as varying values of its parameters and measurements.

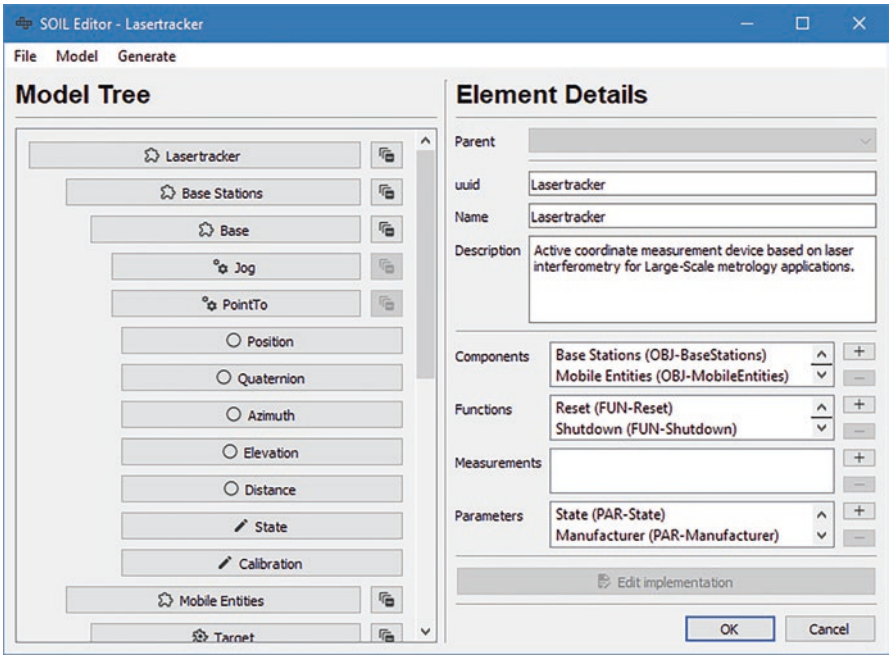


Fig. 2. The *SOIL-Editor*. On the left, the model can be browsed and on the right, the fields of the currently selected element can be edited.

SOIL is a declarative, graphical modelling language. Sensor interfaces are defined using the *SOIL-Editor* depicted in Fig. 2. The left area of the editor shows the model tree presented as indented list. This depiction is both compact and easy to explore. The detail view on the right is used for editing the fields and characteristics of the selected element. Using SOIL, the development of a sensor service is a three-step procedure. In the first step, the user models the interface by adding elements and specifying the required fields, serving as basis for a static service template. The elements of the model are translated into classes and/or functions of the target GPL. The generated source code serves as second template, offering dedicated places to ingest device specific code. Components with dynamic child components are generated with fully implemented methods for addition and deletion of these components during runtime.

The final step in the development process is completed by generating the communication endpoints. The implementation of the communication endpoints is inherent part of SOIL, such that the developer of the sensor practically does not have to deal with communication specific code. The implementation of the relevant communication endpoints can be generated on demand without any changes to the sensor model

or hardware specific code required. By integration of multiple different communication protocols into SOIL the sensor can be deployed into different network systems without additional effort or need for protocol routers. Besides, a designed SOIL model can be exported and imported into and from the JSON-format, serving as an offline, human-readable and device-independent service model respective description.

Fig. 3. IoP testbed infrastructure used at the laboratory of the authors comprising sensors, data bases and analysis tools and responsive user interfaces in different network tiers. The copyright of reused logos and device photos remains among their authors.

4 Proof-of-Application

In order to prove applicability, the authors implemented a prototype for translation of a SOIL model to Python. The places for ingestion of the hardware specific source code are provided by generating an object-oriented implementation of the sensor model. Components are translated into classes having all children elements as protected attributes. For each measurement and parameter, a method is generated for the parent component class, which gives access to the value of the element and returns the corresponding attribute by default. Additionally, a function for setting non-constant parameters is generated. For each function, a method forwarding specified arguments to the device-specific implementation and returning respective return values is generated. The generated implementation of the designed model is already executable.

The integration of a sensor interface developed with SOIL has been tested with the testbed network infrastructure at the authors' laboratory shown in Fig. 3. Within this testbed each device must implement a REST-like interface via HTTP. Via the REST interface the resource model of the sensor service can be determined, devices can be controlled and data can be actively queried using GET-requests. The identifiers of the elements are concatenated to form the URL at which the data of this element is available. All representation use serialization in JSON. Because all information necessary for syntactical, physical interpretation of the (measurement) data is transmitted,

requirements for the client implementation can be automatically derived. Other methods used are POST for triggering functions and PATCH for updating parameters. The continuous data access requirement is respected by generating a schedule for an MQTT publisher that calls the methods giving access to the measurements at fixed, user-defined intervals and publishing the measured values to the broker. The SOIL structure is also represented in the MQTT topics.

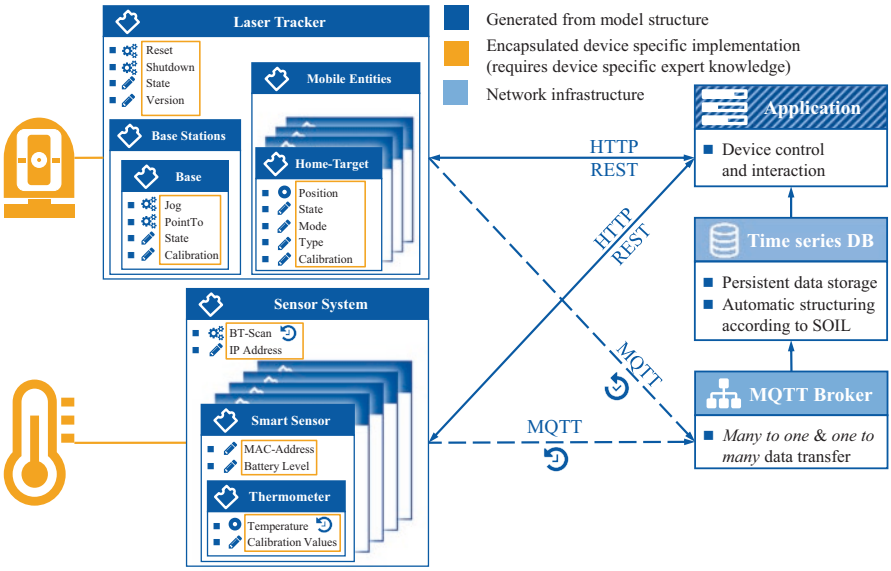


Fig. 4. Exemplary system implemented using SOIL. The main benefit is visible here, as the device specific implementation (yellow) is encapsulated using SOIL (dark blue), requiring minimal knowledge regarding network infrastructure implementation (light blue).

Due to the self-documenting interface protocol, the use of SOIL allows for automatically creating databases and user interfaces for arbitrary devices. In conjunction with visualization tools (e.g. Grafana™), data can be monitored immediately. Applications working with SOIL devices can also access historic, streamed measurement data from a database according to the SOIL device structure.

The authors tested SOIL implementing a virtual laser tracker and a sensor node for a self-developed distributed smart sensor network [18]. Practically, the encapsulation of device specific code consists of two main steps. The model structure for SOIL is intuitively predetermined by the nature of the device (simplified in the following). Thus, the functional structure of the device in hardware and corresponding software structure is simply to be implemented as a device model in SOIL (see Fig. 4) using the SOIL-Editor. For Large-Scale Metrology devices, the SOIL structure can generally resemble the device structure depicted in [17].

With the SOIL model reproducing the functional device structure, the device specific code segments have to be implemented in SOIL-Editor, oftentimes being a one-to-one transfer of already existing device interfacing implementation.

5 Outlook/Conclusion

This paper proposes *SOIL,* a new graphical domain-specific programming language for defining the interfaces for sensor services. It enables production engineers and sensor developers with little-to-none knowledge about network architectures and communication protocols to implement a ready-to-use server implementation for a sensor. The tree-like sensor model is constructed from a simple meta-model consisting of four elements only. From that model, the communication specific and the static definition of the sensor's implementation is generated in a general-purpose language. Access to the sensors hardware is injected into the empty method bodies of the sensor implementation. The development process is finished by generating the server code mapping the sensor implementation with the communication endpoints.

The authors implemented a standalone and easy-to-use graphical editor in Python. The created model is translated into executable Python code. The concept has been proven by implementation of a virtual laser tracker and a distributed temperature measurement system.

Next steps include defining a formalised language description and implementing a web-editor for SOIL. To extend the number of applicable use-cases it is envisaged to add more target languages, such as C++, and communication protocols. It is intended to provide a set of predefined standard-elements, e.g. a 3D position measurement, IP-address parameter or template initialisation functions, to increase reusability and provide examples for reduced learning effort. For developers with profound programming experience, it is envisaged to develop a textual representation. Last, the authors plan to implement modules for applications using SOIL devices, simplifying development by abstracting remote device interfacing.

Acknowledgements. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

1. Pennekamp, J., Glebke, R., Henze, M., Meisen, T., Quix, C., Hai, R., Gleim, L., Niemietz, P., Rudack, M., Knape, S., Epple, A., Trauth, D., Vroomen, U., Bergs, T., Brecher, C., Buhrig-Polaczek, A., Jarke, M., Wehrle, K.: Towards an infrastructure enabling the internet of production. In: 2019 IEEE International Conference on Industrial Cyber Physical Systems (ICPS). IEEE, pp. 31–37 (2019)
2. Schmitt, R.H., Voigtmann, C.: Sensor information as a service – component of networked production. J. Sens. Sens. Syst. **7**(1), 389–402 (2018). 10.5194/jsss-7-389-2018
3. Thramboulidis, K., Christoulakis, F.: UML4IoT-A UML-based approach to exploit IoT in cyber-physical manufacturing systems. Comput. Ind. *82*, 259–272 (2016). https://doi.org/10.1016/j.compind.2016.05.010
4. Hürsch, W. L., Videira Lopes, C.: Separation of Concerns, Technical Report, Bosten (1995)
5. Brecher C, Klocke F, Schmitt R et al. (eds.): Internet of production für agile Unternehmen: AWK Aachener Werkzeugmaschinen-Kolloquium 2017, 18. bis 19. Mai, 1. Aufl. Apprimus Verlag, Aachen (2017)

6. Mernik, M., Heering, J., Sloane, A.M.: When and how to develop domain-specific languages. ACM Comput. Surv. **37**(4), 316–344 (2005). 10.1145/1118890.1118892
7. Rodriguez Corral, J.M., Ruiz-Rube, I., Civit Balcells, A., Mota-Macias, J.M., Morgado-Estevez, A., Dodero, J.M.: A study on the suitability of visual languages for non-expert robot programmers. IEEE Access **7**, 17535–17550 (2019). 10.1109/ACCESS.2019.2895913
8. González García, C., Pelayo G-Bustelo, B.C., Pascual Espada, J., Cueva-Fernandez, G.: Midgar: generation of heterogeneous objects interconnecting applications. A domain specific language proposal for internet of things scenarios. Comput. Netw. **64**, 143–158 (2014). 10.1016/j.comnet.2014.02.010
9. Amrani, M., Gilson, F., Debieche, A., Englebert, V.: Towards user-centric DSLs to manage IoT systems. In: Pires, L. F., Hammoudi, S., Selic, B. (eds.) MODELSWARD 2017: Proceedings of the 5th International Conference on Model-Driven Engineering and Software Development, Porto, Portugal, February 19–21. SCITEPRESS - Science and Technology Publications Lda, Setúbal, pp. 569–576 (2017)
10. Tomlein, M., Grønbæk, K.: A visual programming approach based on domain ontologies for configuring industrial IoT installations. In: Proceedings of the Seventh International Conference on the Internet of Things, ACM, New York, pp. 1–9 (2017)
11. Goldschmidt, T., Mahnke, W.: Evaluating domain-specific languages for the development of OPC UA based applications. IFAC Proceedings Volumes **45**(2), 860–865 (2012). 10.3182/20120215-3-AT-3016.00152
12. Elsts, A., Judvaitis, J., Selavo, L.: SEAL: A domain-specific language for novice wireless sensor network programmers. In: Demirors, O. (ed.) 2013 39th EUROMICRO Conference on Software Engineering and Advanced Applications (SEAA 2013), Santander, Spain, 4–6 September 2013. IEEE, Piscataway, pp. 220–227 (2013)
13. Hackel, S.G., Härtig, F., Hornig, J., Wiedenhöfer, T.: The digital calibration certificate. PTB-Mitteilungen **127**, 57–81 (2018). 10.7795/310.20170403
14. Pfrommer, J., Grüner, S., Goldschmidt, T., Schulz, D.: A common core for information modeling in the Industrial Internet of Things. at - Automatisierungstechnik **64**(9) (2016). doi: 10.1515/auto-2016-0071
15. Fischer, G., Giaccardi, E., Ye, Y., Sutcliffe, A.G., Mehandjiev, N.: Meta-design. Commun. ACM **47**(9), 33 (2004). 10.1145/1015864.1015884
16. Evertz, L., Epple, U.: Semi-automatic development of service adaptors from property-based service descriptions. In: 2015 IEEE 20th Conference on Emerging Technologies & Factory Automation (ETFA), 8–11 Sept. 2015, City of Luxembourg, Luxembourg. IEEE, Piscataway, pp. 1–5 (2015)
17. Montavon, B., Peterek, M., Schmitt, R. H.: Model-based interfacing of large-scale metrology instruments. In: Stella, E. (ed.) Multimodal Sensing: Technologies and Applications, 26–27 June 2019, Munich, Germany. SPIE, Bellingham, Washington, p. 11 (2019)
18. Ohlenforst, M., Jantzen, M., Schmitt, R.H.: Verfahren und System zur in-process-Berechnung einer dreidimensionalen Temperaturverteilung. Rheinisch-Westfälische Technische Hochschule (RWTH), Aachen (2018)

An Economic Solution for Localization of Autonomous Tow Trucks in a Mixed Indoor and Outdoor Environment Using a Node Based Approach

M. Zwingel[1](✉), M. Herbert[1], M. Lieret[1], P. Schuderer[2], and J. Franke[1]

[1] Institute for Factory Automation and Production Systems (FAPS), Friedrich-Alexander-University Erlangen-Nuremberg (FAU), Egerlandstr. 7-9, 91058 Erlangen, Germany
maximilian.zwingel@faps.fau.de
[2] University of Applied Sciences Ingolstadt, Esplanade 10, 85019 Ingolstadt, Germany

Abstract. Flexible intralogistics solutions are basic components to enable classic manufacturing companies to compete in dynamic and global production networks. Developments in the field of autonomous road traffic technology and the increasing availability of cost-efficient sensors allow the economic use of autonomous systems in various industrial sectors. One of the subareas of autonomous systems with economic and technological potential is the intralogistics material supply with driverless transport systems. While the usage between buildings of such systems is not yet economically feasible, development of cost-efficient sensors from other markets are opening up new approaches in this area of operation. Localization and navigation, which must work seamlessly in mixed indoor and outdoor usage scenarios, are particularly critical to the success of autonomous systems.

In this paper, a solution for the localization and navigation of autonomous driverless transport systems in a mixed indoor and outdoor scenario is proposed. It focuses on the use of cost-efficient sensors to enable a mostly infrastructure-independent localization of each system. The proposed solution is validated and a first approach for a dynamic fusion of localization data is done.

Keywords: AGVs · Autonomous mobility · Cost-efficient sensors

1 Automation in Intralogistics to Meet Flexibility Challenges

A growing number of variants and shorter manufacturing cycles in the serial production of industrial goods require increasing flexibility in transport and supply chains, both between and within companies [1]. In conjunction with the increasing scarcity of available and appropriately trained personnel, especially in high-wage countries,

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 457–466, 2021.
https://doi.org/10.1007/978-3-662-62138-7_46

there is a growing demand for automated technical solutions that reflect the flexibility and productivity of existing human driven systems [2]. In the intralogistics sector, an increasing number of automated guided vehicle systems and autonomous transport robots are used to meet the requirements of driverless, dynamic transport. On closer inspection, however, it becomes clear that these systems cover individual applications and do not include a holistic view of the requirements of cross-infrastructure use, especially when used in a shared environment with people and other road users [3].

In general, available solutions show a strong focus on material transport in indoor environments, for which a variety of commercial solutions and prior research exist, both for isolated use and for use in a shared environment [4]. For material transport in outdoor areas, however, there are only individual solutions that can rarely be used economically due to the high acquisition costs and the need for additional infrastructure [3].

Research and development in other fields of mobile robotics, especially in the automotive [5, 6] and human transport sector [7], offer an increasing number of cost-efficient sensors and robust algorithms for use in the intralogistics sector. Particularly in the areas of localization and navigation of autonomous systems, serious progress has been made in recent years, driven by the huge market potential in those sectors. Based on an analysis of the relevant sensors for the environmental perception of autonomous transport systems indoors and outdoors, this paper presents cost-efficient sensors for all relevant measurement principles. In combination with a modified approach for the definition of the working environment and the thus possible adaption and weighting of individual sensor data, a solution for localization and navigation is made possible.

The presented solution is tested by adapting a commercially available tow truck system for use as an autonomous transport vehicle, validating the basic functionalities of the localization and navigation system based on cost-efficient sensors.

2 State of the Art

Autonomous material transport systems in intralogistics applications benefit highly from the availability of cost-efficient sensors, which primarily originate from high volume markets, e.g. the automotive supplier sector. Modern road vehicles across all price classes already provide a wide range of assistance functions that are implemented by means of appropriate environmental perception. The size of the market ensures that the corresponding sensors, regardless of the measuring principle, are becoming more cost-efficient as competition rises. This includes, for example, radio detection and ranging (RADAR) sensors used for adaptive cruise control (ACC), camera systems for traffic sign recognition, ultrasonic systems for parking assistants or light detection and ranging (LiDAR) systems in the field of autonomous driving. In addition, the growing tendency of individual manufacturers to use free interface communication allows more and more applications other than the pure traffic vehicle [8].

In the field of driverless transport systems, research has been conducted particularly in the area of indoor material transport, where numerous publications and

developments focus on the localization and navigation of driverless transport vehicles. Ultrasonic sensors, optical sensors or LiDAR systems are used to achieve basic navigation functionalities for autonomous transport vehicles using different mapping and localization algorithms. Due to the high information content and numerous research in the field of efficient algorithms, the use of LiDAR systems in the interior area has become particularly established [9–11].

In the area of outdoor personal transports, other means of localization and environmental perception are used. The focus is put primarily on robust and weather-resistant sensor technologies based on measuring principles such as RADAR and high-performance LiDAR systems. For localization, global navigation satellite systems (GNSS) are used to enable a robust localization independent of known environmental maps [12, 13]. Individual preliminary work expands the horizon beyond a single defined operational environment, but is usually limited to one functional aspect of autonomous transport robots or relies on redundant and costly sensors for environment perception [14, 15].

In the field of navigation for autonomous transport vehicles in intralogistics, proprietary system solutions are used. These mostly use on-site mapping and offline teach-in of relevant routes within the framework of general commissioning. When adapting the necessary routes or making changes in the infrastructure, a new recording of the current environment is necessary. Fully autonomous systems that can compensate for such changes do not exist in intralogistics [14].

3 Cost-Efficient Sensors for Reliable Environment Perception

In order to meet the demands of rapidly changing production environments while also being economically viable, it is therefore necessary to investigate a fully autonomous solution based on cost-efficient sensors for a mixed indoor and outdoor environment. The analysis of the state of the art shows that a large number of relevant measuring principles must be covered by the sensor setup of a fully autonomous transport vehicle. In particular, the following sensors need to be taken into account and have to be investigated in a comprehensive approach.

- RADAR-Sensors for collision avoidance
- LiDAR-Sensors for localization and navigation
- Optical Sensors for localization and navigation
- Ultrasonic sensors for collision avoidance
- GNSS Sensors for absolute localization
- Ultra-wideband (UWB) Sensors for absolute localization
- Odometer for relative localization

For each of these measuring principles cost-efficient sensors can be found in the automotive and the consumer sector. The classification of cost-efficiency depends on the information content of the sensor and its value in the context of localization and navigation of autonomous transport systems in indoor and outdoor intralogistics.

The sensors functionality can be evaluated according to the characteristics specified in [16]. When comparing the sensor principles with the evaluation criteria of cost-efficiency, optical systems and LiDAR systems in particular show their strengths in use with autonomous systems and thus justify higher prices. The high information density makes the sensor data best suitable for localization and navigation tasks. Optical systems such as the Realsense T265 for tracking and the Realsense D435i for depth and image perception show excellent performance in regard to their retail price. A combination of those sensors provides a visual localization based on optical data as well as a colour and depth image of the surroundings suitable to use in SLAM Algorithms and navigation tasks, enabling relative and absolute localization. In the field of LiDAR sensors, the RPLiDAR A3 from SLAMTech qualifies as a cost-efficient sensor for localization and mapping. By using different algorithms, the sensor data can be used for relative localization, absolute localization as well as collision avoidance in navigation tasks. To assist the functionality of absolute localization, sensors using a global reference system are used, such as GNSS and UWB receivers. RADAR and ultrasonic sensors are used according to their functionality as means of collision avoidance. Proposed sensors, which qualify as cost-efficient in relation to their functionality in an autonomous transport system, are given in Table 1.

Table 1. Cost-efficient sensors for environmental perception

Sensor:	Details and purpose:	Est. Price:
Bosch 5Q0907561	RADAR 77 GHz Purpose: Collision Avoidance	180 €
RPLiDAR A3	LiDAR (785 nm), Purpose: Collision Avoidance, Relative localization, Absolute localization	600 €
NEO-M8N GNSS	GPS L1 & GLONASS L1, Purpose: Absolute localization	30 €
DWM1001	UWB (6.5 GHz) Purpose: Absolute localization	300 €
Realsense T265	Optical features Purpose: Relative localization	200 €
Realsense D435i	RGB-D, IMU Purpose: Collision Avoidance, Absolute localization	200 €
HC-SR04	Ultrasonic (40 kHz), Purpose: Collision Avoidance	6 €

Together, the cost-efficient sensors cover the common measuring principles of driverless transport systems in intralogistics. In order to achieve the required autonomy, the perception of the environment must be implemented in a holistic approach, which combines the individual sensor data and thus enables the system's ability to localize and navigate. The solution must enable the autonomous system to determine the reliability of the individual sensor's data and generate a combined position estimate with increased accuracy and reduced variance.

4 Proposed Approach for Seamless Navigation and Localization

To achieve such a feature, a global model of the working environment is required. For use with cost-efficient sensors and without the need to teach environments, we propose the use of geodata following the convention of the.osm standard, which is composed of four data primitives and relies on referencing their position in geographic coordinates [17]. Nodes and ways are used for the definition of contours and traffic routes, transit areas or stations of material transport. They thus define the layout of the operational area as well as possible routes before the use of an autonomous transport vehicle system, as shown in Fig. 1.

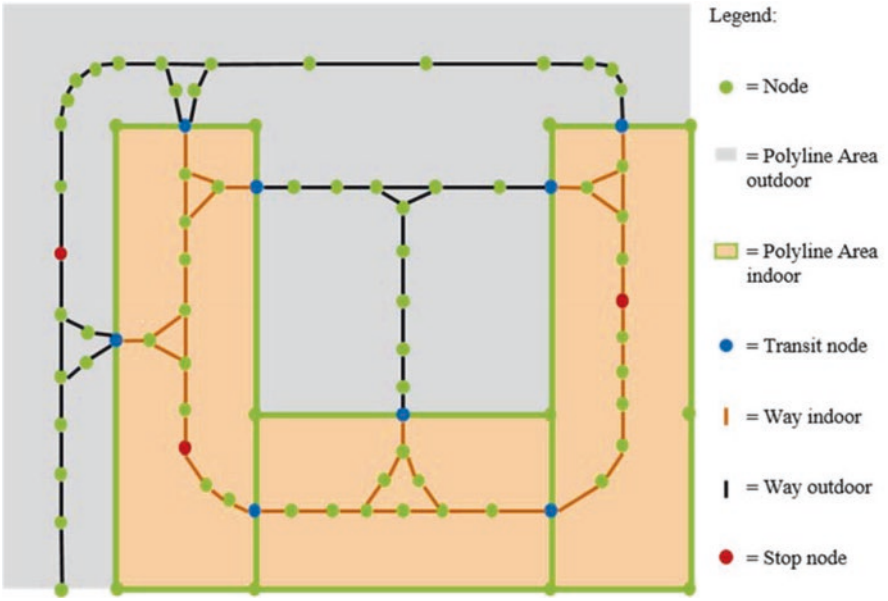


Fig. 1. Example layout for a mixed indoor-outdoor environment using a node based approach

The third component, the tags, allows the assignment of the external environmental conditions and, using an automated information parser, a fully autonomous behaviour of the driverless transport vehicle. They are designed as key-value pairs and can be assigned as needed. For the use of an autonomous transport vehicle in an intralogistics context, the following key-value pairs are defined and assigned to the corresponding nodes and ways as given in Table 2.

The geodata according to the.osm standard is stored as an xml file, which makes the corresponding information readable by computer systems. This enables the use of the generated map inside a global solution for autonomous transport systems by referencing local maps, navigation goals and possible environmental influences on the cost-efficient sensors based on their GNNS coordinates and linking them to specific keys and values.

Table 2. Key Value Pairs used to define custom behavior suited to driverless transports

Key:	Value:	Info:
Environment	Indoor	Defines a way as indoors
Environment	Outdoor	Defines a way as outdoors
Transit	YES	Defines a node as transit gate
Tow_truck_stop	1,2,3, ,n	Defines a node as n-th tow truck stop

5 Combination of Cost-Efficient Sensors with a Node-Based Approach for Environment Dependent Behaviour

The combination of geodata, its automated evaluation and the ability to define a specific behaviour based on every possible position of the autonomous system enables the use of cost-efficient sensor technology in a dynamic concept. The node-based approach providing knowledge about the working environment enables the autonomous configuration and evaluation of the different cost-efficient sensors as well as the adaption of localization and navigation algorithms. To test the fusion of cost-efficient sensors with the proposed approach and to evaluate the corresponding effects on localization and navigation, a commercially available tow truck, adapted to be driven autonomously and shown in Fig. 2, is equipped with the sensors stated in Chap. 3 and provided with a node-based representation of its working environment.

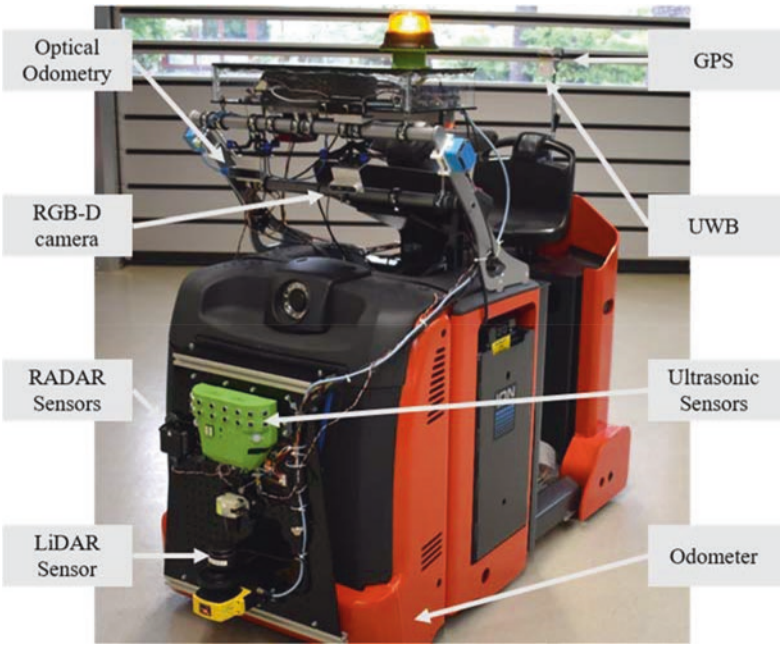


Fig. 2. Tow Truck system equipped with cost-efficient sensors for environment perception and localization

Based on global positioning sensors (UWB or GPS) the system can initially state its current position and link it to the node representing the closest match. Based on the information provided by the node's tag, the cost-efficient sensors for relative localization and navigation are configured and the corresponding algorithms loaded. Following the initial localization, the position is calculated by using all sensors available [18]. Taking into account the reliability of the sensors used, which is extracted from the covariance matrix of the individual sensor's data, and taking into account the environmental conditions, the individual data is combined to form a global position estimate. The available data is used to calculate the absolute and relative pose from the corresponding data sets according to the provided covariance. To increase the significance of the localization results, the covariances of the sensor information are adapted by individual functions before the fusion itself is done based on environmental information. The necessary data about the environmental parameters to adequately adapt the covariance matrix is provided by assistance sensors, which record all relevant environmental parameters. This may include type and amount of precipitation, light intensity, outside temperature, air pressure and wind speed. The same adaption of sensor data is done in regards to environment perception, especially for collision avoidance using the cost-efficient sensors.

6 Reference Course and Validation of Cost-Efficient Sensors for Localization of Autonomous Systems

To test the capabilities of the cost-efficient sensors and the proposed approach for the dynamic weighting of single sensor information for localization and navigation purposes, the environment of the Institute of Factory Automation and Production Systems at the University of Erlangen-Nuremberg is used. Here, a mixed indoor and outdoor operation with transit through a hall gate can be realized.

To evaluate the overall system the described sensors were placed on a mobile system and merged according to the proposed fusion based on sensor covariance and autonomous environment detection. Figure 3 shows the results of a representative drive with the individual localizations and the calculated total trajectory. The according offset after the closed loop is detailed in Table 3. Despite the limitations of individual data sources, the approach shows the potential for a suitable localization and environment perception, thus enabling navigation, of driverless systems using cost-efficient sensors. Due to the proposed approach with its implied knowledge an accuracy for final localization of less than 2 m after a closed loop is possible in a use case scenario in-between buildings with dynamic environmental conditions. Sensor sources affected by drift, such as the vehicle's own wheel odometry, are taken into account as well as location-based localization solutions, for example UWB and GPS.

Fig. 3. Test environment and calculated trajectories based on cost-efficient sensor data (Background © Stadt Erlangen| © Hansa Luftbild AG, CC BY-4.0, [19])

Table 3. Resulting geometrical offset for the test run shown in Fig. 3 in relation to ground truth (positive x is right, positive y is up, positive yaw is counter-clockwise)

Type:	Offset x:	Offset y:	Offset yaw:	Abs. distance:
Optical odometry	+3.482 m	−2.533 m	+1.08°	+4.306 m
Wheel odometry	−2.046 m	+58.202 m	+38.36°	+58.238 m
Fused odometry	+4.414 m	+3.834 m	+3.41°	+5.847 m
GPS localization	+7.797 m	+2.238 m	Not applicable	+8.112 m
Optical localization	+0.343 m	−0.150 m	−0.69°	+0.374 m
UWB localization	−0.381 m	−0.038 m	Not applicable	+0.374 m
Fused localization	+0.331 m	−0.132 m	−1.41°	+0.356 m

7 Conclusion and Further Research

By using a superordinate methodology based on a node approach for the definition of environmental conditions in a global context in connection with cost-efficient sensors that cover all relevant measurement parameters for the autonomous operation of driverless transport vehicles, an approach for the economically reasonable use of driverless transport systems has been proposed. Experiments in the environment of the Chair of Factory Automation and Production Systems show promising results for the localization, navigation and thus the autonomous use of such a system.

In the context of further research, the limits of cost-efficient sensors in the area of driverless systems are to be investigated. Based on the results obtained, the navigation of driverless transport systems will be adapted and the path to a fully autonomous system will be pursued. The dynamic adaptation of sensor data for an optimal fusion of autonomous transport vehicles will also be further investigated.

Acknowledgements. The Project "E|SynchroBot" (20184 N, Forschungsgemeinschaft Intralogistik/Fördertechnik und Logistiksysteme e. V.) is supported by the German ministry for Economy and Energy by decision of the German Parliament.

References

1. Schönsleben, P.: Integrales Logistikmanagement: Operations und Supply Chain Management innerhalb des Unternehmens und unternehmensübergreifend, 8th edn. Springer Vieweg, Berlin (2020)
2. Scholz, M., Zhang, X., Kreitlein, S., Franke, J.: Decentralized intelligence: the key for an energy efficient and sustainable intralogistics. Procedia Manuf. **21**, 679–685 (2018)
3. Sabattini, L., Digani, V., Secchi, C., Cotena, G., Ronzoni, D., Foppoli, M., Oleari, F.: Technological roadmap to boost the introduction of AGVs in industrial applications. In: 2013 IEEE 9th International Conference on Intelligent Computer Communication and Processing (ICCP), pp. 203–208 (2013)
4. Scholz, M., Zwingel, M., Schuderer, P., Franke, J.: Sustainable intralogistics due to uniform software and modular transport entities. Procedia CIRP **80**, 239–244 (2019)
5. Vivacqua, R., Vassallo, R., Martins, F.: A low cost sensors approach for accurate vehicle localization and autonomous driving application. Sensors (Basel, Switzerland) **17**(10), 2359 (2017)
6. Wall, R. W., Bennett, J., Eis, G.: Creating a low-cost autonomous vehicle. In: IEEE 2002 28th Annual Conference of the Industrial Electronics Society. IECON 02, pp. 3112–3116 (2002)
7. Azad, M., Hoseinzadeh, N., Brakewood, C., Cherry, C.R., Han, L.D.: Fully autonomous buses: a literature review and future research directions. J. Adv. Transport. **2019**, 1–16 (2019)
8. Fleming, W.J.: New automotive sensors—a review. IEEE Sens. J. **8**, 1900–1921 (2008)
9. Wang, Y.T., Peng, C.C., Ravankar, A.A., Ravankar, A.: A single LiDAR-based feature fusion indoor localization algorithm. Sensors (Basel, Switzerland) **18**(4), 1294 (2018)
10. Menegatti, E., Maeda, T., Ishiguro, H.: Image-based memory for robot navigation using properties of omnidirectional images. Robot. Auton. Syst. **47**, 251–267 (2004)

11. Ramer, C., Sessner, J., Scholz, M., Zhang, X., Franke, J.: Fusing low-cost sensor data for localization and mapping of automated guided vehicle fleets in indoor applications. In: 2015 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI), pp. 65–70 (2015)
12. Kümmerle, R., Ruhnke, M., Steder, B., Stachniss, C., Burgard, W.: Autonomous robot navigation in highly populated pedestrian zones. J. Field Robot. **32**, 565–589 (2015)
13. Kentley, T.D., Levinson, J.S., Gamara, R.Y., Sibley, G.T.: Autonomous vehicle fleet service and system. Anmeldenr. US14/932,959, US. Veröffentlichungsnr. US9606539B1
14. Easy Mile: Press kit: delivering driverless mobility solutions. https://easymile.com/wp-content/uploads/2019/03/EasyMile-Press-Kit.pdf. Überprüfungsdatum 2020-05-15
15. Lingemann, K., Surmann, H., Nuchter, A., Hertzberg, J.: Indoor and outdoor localization for fast mobile robots. In: 2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566), pp. 2185–2190, 28 Sept.-2 Oct. (2004)
16. VDI-Richtlinie 2510. Fahrerlose Transportsysteme (FTS). Beuth, Berlin (2005)
17. Haklay, M., Weber, P.: OpenStreetMap: user-generated street maps. IEEE Pervas. Comput. **7**, 12–18 (2008)
18. Moore, T., Stouch, D.: A Generalized Extended Kalman Filter Implementation for the Robot Operating System. Springer, Cham (2016)
19. Hansa Luftbild AG, Stadt Erlangen: Luftbild der Stadt Erlangen. https://osm.rrze.fau.de/map-ll-osm?zoom=20&lat=49.5743054&lon=11.025798&layer=Luftbild%20Erlangen%202018

Automated Assembly of Thermoplastic Fuselage Structures for Future Aircrafts

S. Kothe[1(✉)], B. Diehl[2], D. Niermann[2], L. Chen[2], M. Wolf[1], and W. Hintze[2]

[1] Fraunhofer IFAM, Wiener Str. 12, 28359 Bremen, Germany
simon.kothe@ifam.fraunhofer.de
[2] Fraunhofer IFAM, Ottenbecker Damm 12, 21684 Stade, Germany

Abstract. In order to meet the high demand for aircrafts and the strict requirements regarding CO_2 emissions in the future, new designs and technologies are necessary. Existing aircraft layouts and production processes are not designed for production rates of more than 75 aircrafts per month. One promising approach to reach significant improvements regarding weight and process cycle time is the combination of thermoplastic carbon fiber reinforced plastic aircraft structural elements, lining parts and cabin system elements to one integrated structure module. By this, improvements in productivity and environmental sustainability are accomplished likewise. This paper provides an overview of the Clean Sky 2 Multifunctional Fuselage Demonstrator, which comprises these topics. The focus of the scientific discourse lies on the concept of the full-scale assembly demonstrator especially discussing handling technology for position and shape adjustment of fuselage shells. Moreover, a systematic evaluation of thermoplastic welding technologies and a discussion regarding welding joint design is presented.

Keywords: Automated aircraft assembly · Thermoplastic CFRP · Thermoplastic welding · Lightweight design

1 Introduction

Today one of the most important challenges of the aviation industry is to make aircrafts more climate friendly. This particularly includes the reduction of CO_2 emissions, which is one of the primary goals of the vision "Flightpath 2050" [1] and "The European Green Deal" [2] of the European Commission. To achieve this, lightweight structures based on carbon fiber reinforced plastics (CFRP) are a promising solution as they offer a very high strength-to-weight-ratio [3]. Whilst current short-range aircraft programs (Airbus A320 and Boeing 737) still rely on a a metal fuselage – mostly aluminum alloys –, the current medium and long-range aircraft programs (Airbus A350 XWB and Boeing 787) already have a fuselage made of CFRP [4]. Nevertheless, based on currently mainly used thermoset CFRP materials lightweight

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 467–475, 2021.
https://doi.org/10.1007/978-3-662-62138-7_47

potential of aircraft fuselages is not exhausted. Reason for this is that aircraft structures based on thermosets are typically assembled with rivets, which need pre-drilled holes in the structure, leading to thicker CFRP material to ensure adequate safety.

An approach to overcome this disadvantage would be the application of thermoplastic CFRP materials for fuselage structures in combination with thermoplastic welding as joining method. Since no holes are necessary for this joining method, thinner aircraft fuselage structures, in particular thinner fuselage shells, can be realized still ensuring adequate safety. Furthermore, this strategy would also allow highly integrated fuselage sub-structures, e.g. pre-equipped with floor-grids, overhead storage compartments or lining parts, as thermoplastic welding is a dust- and chip-less joining method [5]. Based on riveting methods, that have to be used for thermoset fuselage structures, this pre-equipment is not possible as chips and dust generated during drilling of rivet holes might remain in hollow spaces between the structures.

Unfortunately, high performance thermoplastic CFRP materials suitable for aircraft fuselages are typically more than 30% more expensive than thermosets [6–8]. Due to this, the additional material costs should be compensated by significantly lower manufacturing costs to the extent possible. For this, a higher degree of automation in combination with a significant reduction of lead-time is essential [5, 9]. Fortunately, the opportunity of a higher integration of the fuselage sub-structures at an early stage enabled by the application of thermoplastic welding boosts this objective as the up to now very sequential assembly of fuselage structures can be more parallelized. Even an upstream outsourcing of the pre-integration is possible.

In this paper, the Clean Sky 2 project MultiFAL is presented with focus on the concept and design of a full-scale joining and assembly demonstrator for thermoplastic fuselage structures. In addition to a comparison of different thermoplastic welding technologies and joint designs, the elaborated robot-based flexible fuselage handling technology, which allows a position and shape adjustment of the fuselage structures, is discussed. Furthermore, the planned Major Component Assembly (MCA) process with a fully automated flexible aircraft assembly facility will be presented. As the project is currently in the concept phase, no final evaluation of the assembly process is shown in this paper. Rather, it is an overview of the current state giving the scientific community the possibility to discuss the work done so far and influence further steps.

2 The Multifunctional Fuselage Demonstrator and the MultiFAL Project

Thermoplastic CFRPs have the potential to develop new fuselage designs and system installation architectures with significant weight and cost savings for future aircraft programs. To demonstrate the feasibility and develop technologies needed for such an innovative approach, a full-scale Multifunctional Fuselage Demonstrator (MFFD) is worked out in the Clean Sky 2 research program. The key goals of the MFFD are to enable production rates of 75–100 aircrafts per month, to reduce the total fuselage weight by 1,000 kg and to lower the recurring costs by 20% [10].

To achieve the ambitious weight and cost savings, completely new system and system installation architectures optimized for an assembly process of highly

integrated thermoplastic fuselage segments are developed. By this, a completely optimized fuselage concept is realized, which is superior in all respects to the sum of the traditional, functionally isolated optimization processes [5]. The full-scale cylindrical MFFD (Fig. 1) will contain these developments, comprising the thermoplastic CFRP fuselage skin, frames, stringers as well as all fasteners relevant for the structure. Passenger and cargo floor, cabin and cargo systems, overhead storage compartments and lining parts are also pre-installed.

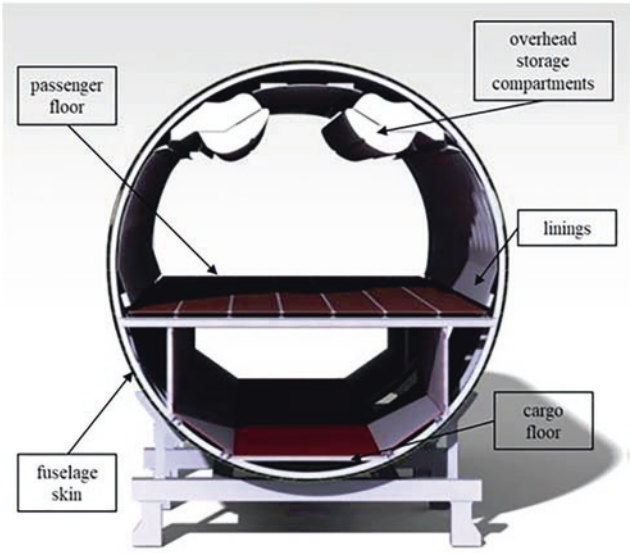


Fig. 1. Multifunctional Fuselage Demonstrator (MFFD)

Since technologies for many different scenarios are combined, the demonstrator value is significantly increased in order to validate the best possible technological and architectural concepts. Even the fuselage itself will not be limited to a single scenario since elements of different designs will be considered. The demonstrator has a length of 8 m and a diameter of 4 m representing a typical single aisle fuselage section. This enables the validation of a number of important technological concepts in full scale.

To demonstrate the economic advantage of highly integrated thermoplastic CFRP structures, the pre-equipped upper and lower fuselage segments will be joined by means of thermoplastic welding. As shown in Fig. 2, the fuselage demonstrator contains two longitudinal joints. To be able to compare different joint designs the demonstrator's left-hand side joint (LH: left-hand side in flight direction) will be a butt strap joint whilst the demonstrator's right-hand side joint (RH) will be an overlap joint. The more complex butt strap joint geometry of the left-hand side is necessary since the door surround in this area leads to a skin thickness variation that can be handled by this special joint design. On the right-hand side the skin thickness is constant (roughly 3 mm), allowing a less complex overlap joint design [11]. Furthermore, frame

couplings will be integrated by means of thermoplastic welding to connect the upper and lower frames of the upper and lower fuselage segments.

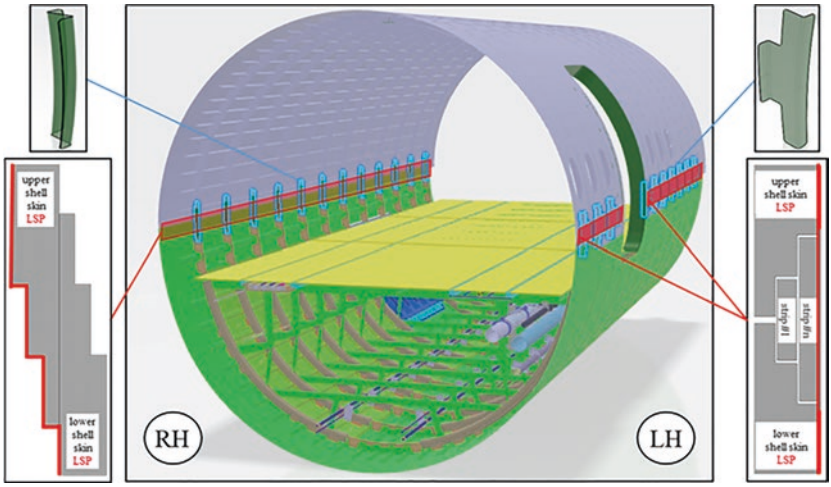


Fig. 2. Longitudinal joints of the MFFD, frame couplings and joint designs [11]

Within the project MultiFAL [9] (project partners: Fraunhofer IFAM, FFT Produktionssysteme GmbH, CT engineering group and AIMEN Centro technologico) the assembly system and process for the MFFD is developed. The work of the project is supported by further sub-projects, e.g. STUNNING (manufacturing of the fully equipped lower segment of the MFFD) [5] or other still to come CfP11 projects [9] [11] focusing on basic technologies for the assembly demonstrated by MultiFAL.

As stated before, thermoplastic welding as joining technology will be used. In this case, the assembly of the upper and lower pre-equipped thermoplastic half segments by longitudinal joints and frame couplings will be in focus. In order to gain as much knowledge as possible, different welding methods will be used.

3 Joining Thermoplastic Fuselage Structures

As stated above, the Multifunctional Fuselage Demonstrator (MFFD) comprises different design alternatives to be able to evaluate different technologies. This is also the case regarding the joining of the two thermoplastic fuselage segments, comprising two types of longitudinal joints, namely a butt strap joint (LH) and an overlap joint (RH), and frame couplings to connect the frames of the upper and lower fuselage half shells, see Fig. 2. To be able to achieve a dustless, clean and highly automatable joining process – these were the central requirements to be able to join highly integrated fuselage segments – the focus is on thermoplastic welding as promising technological solution. Nevertheless, there are several types of thermoplastic welding, namely conduction, ultrasonic, induction, resistance and laser welding, that could be used. Table 1

provides information regarding the functional principle, advantages and disadvantages that helped to make an appropriate choice for the different use-cases.

Table 1. Thermoplastic welding technologies

	Functional principle	Advantages/disadvantages
Conduction	heat energy at surface of one adherent ("electric iron") [5]	+ well-known process [5] + no additional element in the weld [5] + joining of complex geometries [5] – complete heating of adherent [5] – slow process [15]
Resistance	electric current passes resistive element in interface of adherents [12]	+ heat only at welding interface [12] + time independent of welding length [12] – resistive element stays in the weld [13]
Induction	induction of eddy current in conductive element in interface of adherents [17]	+ carbon fibers, especially woven cloth, as conductive element [17] – challenging with UD material [9]
Ultrasonic	friction in interface of adherents [3]	+ good for spot welds [14] + very energy-efficient [16] + fast and automatable [9, 16] – continuous welding to be developed [9]
Laser	laser radiation generates heat on surface of laser absorbing adherent [12, 14, 19]	+ high speed [14] – special weld design necessary [18] – safety issues [14] – high invest cost [15]

Resistance welding and induction welding were not chosen for the demonstrator's longitudinal joint to avoid additional metallic resistive elements or any additional inductive elements, e.g. carbon fiber woven cloth or metallic mesh, in the weld seam. Laser welding was left out due to high invest cost and possible safety issues with respect to the large demonstrator.

Due to joint complexity, conduction welding using a heated pressure plate was chosen for the butt strap joint (LH) (see Fig. 2) though this process requires a complete heating of one adherent. Anyway, this circumstance of a typically slower process was accepted since conduction welding is a well-known welding technique resulting in good welding quality even for complex welding designs.

In contrast to this, ultrasonic welding was selected for the less complex overlap joint (RH) (see Fig. 2) to develop and demonstrate a fast, highly automated welding process for high production rates. Especially the development of a reliable continuous ultrasonic welding technology is seen as one of the primary development goals of the project.

Furthermore, resistance welding was the technological choice for the joining of the frame couplings (see Fig. 2), giving the possibility to weld different frame coupling geometries on the left-hand (LH) and on the right-hand (RH) side with the same technology. Additionally, the fact that the heat is generated directly in the welding

interface resulting in a very fast and energy-efficient welding process is a great advantage. Due to this, a sequential frame coupling by frame coupling welding is possible still allowing an acceptable total process time.

Of course, for the real production of a fuselage one would probably focus on fewer welding technologies or even one technology to reduce costs. Nevertheless, the project consortium decided to choose several technologies to be able to demonstrate, mature and compare different welding technologies at full scale and make a differentiated selection for a possible future production of thermoplastic aircraft fuselages.

To analyze the welding results, an inline monitoring system based on thermography and electromechanical impedance (EMI) measurement, developed with the partners Fraunhofer LBF and Fraunhofer ENAS, will be used. It will allow a continuous condition monitoring of the structures during the process. By this, a systematic validation of the welding process also during future production will be possible.

4 Flexible Aircraft Assembly Facility

The MultiFAL project develops technologies and hardware to realize the Major Component Assembly (MCA) process of a full-scale aircraft fuselage section with welded longitudinal joints. To enable a production rate of more than 75 aircrafts per month on the one hand and to maintain the ability to react to upcoming technologies and trends on the other hand, MultiFAL developments focus on the implementation of a modular, flexible assembly system. For the demonstration of this, the PARAMONT facility, a unique robotic set-up to develop highly efficient solutions for aircraft assembly processes, will be used, see Fig. 3.

Fig. 3. The PARAMONT facility

The PARAMONT facility allows highly automated handling of full-scale fuselage segments. The facility possesses industrial robots on a linear rail equipped with

end-effectors for assembly processes and several parallel robot kinematics (hexapods) for the handling of fuselage structures [20]. In the setup for MultiFAL a rigid structure – the so-called inner deck (see Fig. 4) – will be used to absorb high loads during the thermoplastic welding process and thus relieve the fuselage segments. In addition, systems necessary for the assembly process, such as linear axes for guiding lightweight robots that will perform automated assembly inside the fuselage section, can be added in a modular manner. The MCA process includes the measurement of the components, the precise positioning of the fuselage segments and the provision of the necessary welding pressure in the area of the 8 m long weld lines.

Using the support structure, the lower fuselage segment is either transported on a linear axis or moved by crane under the rigid structure. The upper fuselage segment is transported into the system by crane and then adjusted by hexapods – parallel robot kinematics with six degrees of freedom –, equipped with vacuum-gripper-end-effectors, see Fig. 4. The so-called inner positioners are installed on the inner deck. They perform the fine positioning of the segments for the welding process. The welding pressure is generated during the welding process from the outside by the welding end-effectors itself or, if necessary, by additional elements that press against the inner positioners. The welding end-effectors are guided along the 8 m long seam by linear axes installed on both sides of the fuselage (Fig. 4). The entire system is structured flexibly so that different welding processes and/or weld seam designs can be used.

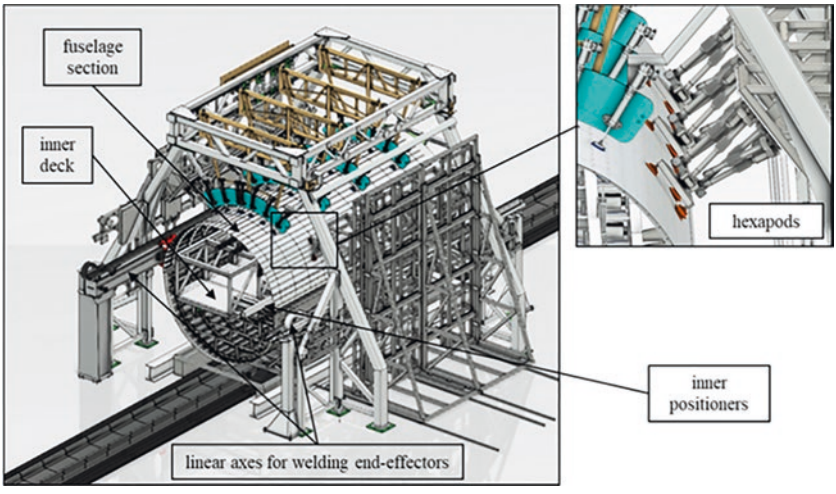


Fig. 4. Flexible aircraft assembly facility

The interfaces of all installed end-effectors and automated systems are designed in a manner that modified tools can be installed if required. Aircraft fuselage sections of different diameters and/or different materials could thus be processed on this flexible, modular assembly system, which enables long-term, eco-efficient and economical operation.

5 Summary and Outlook

Being part of the Clean Sky 2 Multifunctional Fuselage Demonstrator project, the project MultiFAL will investigate the technological and economic feasibility of welding based assembly of highly integrated thermoplastic aircraft fuselage structures. For this, a full-scale test setup comprising different joint designs and welding methods will be developed. Besides the welding process itself, automated metrology-based manipulation and assembly of the CFRP structures to ensure tolerance requirements are in focus. The work performed will be an important step towards implementation of thermoplastic carbon fiber materials for future aircrafts.

Acknowledgement. The work was conducted with financial support from the European Union's Clean Sky 2 program under Grant Agreement LPA-IADP CS2-LPA-GAM-2018-2019-01.

References

1. European Commission – Directorate General for Research and Innovation: Flightpath 2050 – Europe's vision for aviation. ISBN: 978-92-79-19724-6 (2011)
2. European Commission: The European green deal – striving to be the first climate-neutral continent (2019)
3. AVK – Industrievereinigung Verstärkte Kunststoffe e. V.: Handbuch Faserverbundkunststoffe/Composites – Grundlagen, Verarbeitung, Anwendungen. 4th ed., Springer, Wiesbaden (2013)
4. Kumar, K., Zindani, D., Davim, J.: Sustainable engineering products and manufacturing technologies, pp. 207–209. ISBN 978-0-1281-6641-3 (2019)
5. Veldmann, S. L., Kortbeek, P. J., Wölcken, P. C., Herrmann, R., Kos, J., Fernandez Villegas, I.: Development of a multifunctional fuselage demonstrator. In: Proceedings of the Aerospace Europe Conference 2020 (AEC 2020), Feb 25–28, Bordeaux (2020)
6. Composites World: Thermoplastic composites: poised to step forward. https://www.compositesworld.com (2019). Accessed 10 Mai 2020
7. Guo, Q.: Thermosets. Structure, Properties, and Applications, 2nd ed. Elsevier, Amsterdam (2018)
8. Favaloro, M.: A comparison of the environmental attributes of thermoplastic vs. thermoset composites. https://www.thomasnet.com/pdf.php?prid=101809 (2012)
9. Composites World: Proving out LM PAEK welding for Multifunctional Fuselage Demonstrator. https://www.compositesworld.com (2020). Accessed 10 Mai 2020
10. Clean Sky 2 Joint Undertaking: The next generation Multifunctional Fuselage Demonstrator - leveraging thermoplastics for cleaner skies. https://www.cleansky.eu. Accessed: 10 Mai 2020
11. Clean Sky 2 Joint Undertaking: Annex of 11th Call for Proposals (CFP11) – List and Full Description of Topics Call Text R1 [V1]. Brussels (2019)
12. Costa, A.P., et al.: A review of welding technologies for thermoplastic composites in aerospace applications. J. Aeros. Technol. Manage. São José dos Campos **4**(3), 255–265 (2012)
13. Campbell, F. C.: Manufacturing technology for aerospace structural materials, pp. 273–368. ISBN: 978-1-85617-495-4 (2006)
14. Troughton, M. J.: Handbook of plastic joining: a practical guide, 4th. edn., ISBN: 978-0-8155-1581-4 (2008)

15. Neitzel, M.: Handbuch Verbundwerkstoffe – Werkstoffe, Verarbeitung, Anwendung. Werkstoffe, Verarbeitung, Anwendung, 2nd edn. Hanser, München (2014)
16. Villegas, I., Moser, L., Yousefpour, A., Mitschang, P., Bersee, H.: Process and performance evaluation of ultrasonic, induction and resistance welding of advanced thermoplastic composites. J. Thermoplast. Compos. Mater. **26**(8), 1007–1024 (2013). 10.1177/0892705712456031
17. van Ingen, J. W.; Buitenhuis, A.; Wijngaarden, M.; Simmons, F.: Development of the Gulfstream G650 induction welded thermoplastic elevators and rudder. In: SAMPE 2010 – Seattle WA – May 17–20 2010 (2010)
18. Jaeschke, P.; Wippo, V.; Suttmann, O.; Overmeyer, L.: Advanced laser welding of high-performance thermoplastic composites. J. Laser Appl. **27** (2), S29004. 10.2351/1.4906379 (2015)
19. Wippo, V.; Jaeschke, P.; Suttmann, O.; Kaierle, S.; Overmeyer, L.: Laser heat conduction welding of CFRP with modified matrix material. In: High-Power Laser Materials Processing: Applications, Diagnostics, and Systems VIII. 5–7 February 2019, San Francisco, SPIE (Proceedings of SPIE. 5200-, volume 10911), p. 27 (2019)
20. Schmick, F., Borrmann, C., Niermann, D., Wollnack, J.: Tolerance adjusted bonding of large CFRP structures: automated gap analysis in an offline programmed robot assembly cell. J. Adhes. Soc. Jn. **51**, 342–346 (2015)

Towards Adaptive System Behavior and Learning Processes for Active Exoskeletons

B. Otten[1(✉)], N. Hoffmann[1,2], and R. Weidner[1,2]

[1] Laboratory of Manufacturing Technology, Helmut-Schmidt-University/ University of Federal Armed Forces Hamburg, Hamburg, Germany
ben.otten@hsu-hh.de
[2] Chair of Production Technology, Institute of Mechatronic, University of Innsbruck, Innsbruck, Austria

Abstract. Industrial workers still face work-related musculoskeletal disorders daily and therefore physical support systems like exoskeletons are being developed. Making these wearable robots adaptable to different tasks and users in terms of its support characteristics is expected to generate greater performance and broader acceptance. By analyzing relevant elements of joint tasks in groups of humans and the environment exoskeletons are typically being used in, this paper derives the need for a framework allowing for adaption of the exoskeleton to the task, but also predictability for the user of the exoskeleton. A situation aware gain-scheduling controller with internal state feedback to the user is proposed as a means for adaption and predictability.

Keywords: Exoskeleton · Adaptive behavior · Human-machine-interaction · Support systems

1 Introduction to Exoskeleton

Industrially applicable exoskeletons are attributed with a profound influence on the future workspace, aiming to facilitate the workload, prevent occupational injuries, or improve the productivity of the workforce [1–3]. For a successful industrial implementation, it is necessary that the selected exoskeleton fits to the conditions, generally determined by the users, the performed activities, and contextual issues like safety standards or working environments [4, 5]. However, exoskeletons are usually tailored to specific use cases [2, 6, 7], because it is not possible to develop universal applicable solutions considering upfront all possible circumstances due to variations in task (e.g. weight, posture, tools) or user groups (e.g. physical working capacity due to health and/or age, movement patterns) [4]. Thus, the systems' adaptability to individual use cases (tasks, environment, and user) is important [8]. Passive mechanical exoskeletons are therefore especially suitable for monotonous (static or repeated) tasks [9, 10], but hamper versatile tasks and often evoke discomfort [3, 11, 12]. Alternatively, active

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 476–484, 2021.
https://doi.org/10.1007/978-3-662-62138-7_48

exoskeletons can take a variety of input data into account to potentially "identify" and consequently adapt to different use cases, making them more suitable for tasks with greatly changing support demands, dynamics, or postures [7, 13].

2 Challenges in the Interaction Between Exoskeleton and User

Some of the exoskeletons used for industrial purposes conform to the "third arm" type [14], which directly transfer the tools or the handled objects weight by a structure entirely parallel to the human body (see left image in Fig. 1). However, the majority of the systems (e.g., Lucy [7], Levitate [10], Laevo [11]) support individual joints or a group of human joints (see right image in Fig. 1).

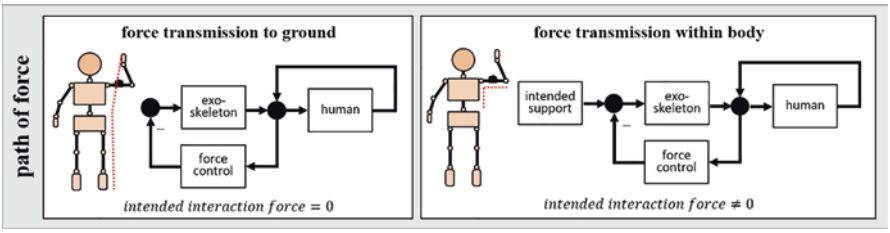


Fig. 1. Basic exoskeleton types in contrast from control perspective. "Get-out-of-the-way" scheme compared to individual joint assist.

Systems which transfer a load entirely parallel to the body must ensure, that interaction forces to the user are low to avoid unwanted contact with the user. This type of control scheme is not applicable to systems designed to assist individual joints, as reducing interaction force would render them useless. Instead these exoskeletons need some way to estimate an optimal interaction torque

$$\tau_{opt} \sim opt(biomechanic\ eval., user\ comfort, task\ performance, \ldots) \tag{1}$$

for their active joints, which is designed such that the exoskeleton provides the best performance from a biomechanical perspective while ideally not compromising on task performance and user comfort. The estimated optimal torque is then the basis for the lower level exoskeletons control and actuation to generate the actual interaction torque

$$\tau_{actual} = f_{drive_{system}}(\tau_{opt})! = \tau_{opt}, \tag{2}$$

which is a function of the drive systems characteristics. Assuming one can measure or predict the actual interaction forces one can calculate the deviation

$$e = \tau_{opt} - \tau_{actual} \tag{3}$$

as the difference of these. Without loss of generality one can assume that the low-level force control loop will be designed such, that the deviation between targeted optimal torque τ_{opt} and actual torque τ_{actual} is reduced to zero over time. Consequently, measuring or estimating the actual interaction force will only inform about deviation of targeted and estimated interaction torques, as it mainly dependents on the limitations of the control and drive system but will hardly give us meaningful insight into the user's movement intention.

Analyzing current passive exoskeletons one can see that all these systems comprise a nonlinear transmission between their energy storage devices (mechanical springs, gas springs, elastomer springs) and the physical user interfaces. The nonlinear transmissions come in a variety of forms like simple lever arm mechanisms, rope and pulley assemblies and cam shaft systems. All these systems have in common that they seek to lower the generated torque in a certain "resting joint range" and increase the torque in certain other "work" ranges. Therefore, one can say that by means of a nonlinear transmission, all passive exoskeletons derive the estimated optimal interaction torque

$$\tau_{actual} = \tau_{\widehat{opt}} = f_{passive}(q_{active\_joint}) \tag{4}$$

during operation from the position of the active joint q_{active_joint}. Some passive exoskeletons allow the user to adjust the function $f_{passive}(q_{active\_joint})$ by either modifying the elastic element or by altering the characteristics of the nonlinear transmission, (e.g. Laevo [11], suitX [15], OttoBock Paexo [9]).

For active exoskeletons to perform significantly better than passive exoskeletons these must provide a more differentiated estimation of desired interaction torques, than passive ones can with their combination of nonlinear transmission and mechanical adjustments. In principle, active exoskeletons can calculate the desired force at the physical interface with a control unit considering various sensor data or direct user inputs.

Since the implementation of an adaptive system behavior can be complex by considering numerous aspects [3, 16, 17], this paper will discuss the environment the exoskeleton is deployed to. From this we will derive potential solutions for improving interaction by looking at other types of human interaction for inspiration.

3 Fundamental Elements and Concepts in a Generalized Exoskeleton Interaction Concept

Machines generally excel human performance in the tasks, they have been designed for, while not being able to adapt to different tasks [18]. Research in robotics has been seeking to provide a more universal system able to perform large varieties of tasks in different environments. However, decades of research have proven that outperforming a human in flexible and changing environments remains challenging. Especially manipulation is still a hard task for systems, while humans can manipulate objects with ease in a variety of industrial settings ranging from assembly to logistics [18].

Comparing a typical machine with an exoskeleton will show that the former is made for a certain task while the latter is made for a certain body region and or posture. While this generally holds true and makes exoskeletons an appealing solution, for tasks that cannot be automated or otherwise improved, it seems that all current exoskeleton have some form of specialization and work well with some tasks, but not others [6]. Consequently, one must consider that task, exoskeleton, and human influence each other, having respective properties that should be accounted for. The quality of their interaction will be the basis for the overall acceptance of an exoskeleton in a specific use case [4].

3.1 Task Coordination in Groups of Humans as Well as Between Human and Robot as a Model for Interaction with Exoskeletons

Assuming one helps a friend to lift or hold something, timing and other parameters must be agreed upon and/or during interaction in a similar manner as between exoskeleton and wearer.

Research in joint action has found that the interaction in groups of humans can be characterized by the "what", "when" and "where" of the joint action [19]. To achieve task performance, one can distinguish between adaption-based strategies, where human co-actors iteratively reduce the timing differences in their actions and anticipation-based approaches, where they use internal models to predict another's persons action [20, 21]. While adaption strategies are helpful in cyclic tasks and have inspired the use of adaptive oscillatory type control schemes in exoskeletons [22], most exoskeletons are deployed in complex tasks, since simple repetitive tasks are easier to automate. Consequently, focusing on strategies, that allow both the user and the exoskeleton for anticipatory action is necessary.

Anticipatory action relies on internal models of the human co-actors about their partners actions. It has been shown that co-actors try to make themselves predictable by reducing variability in task execution [20]. Furthermore, shared knowledge in task execution is known to increase task performance. In experiments where a task, was either shared among a group or performed by an individual, in the shared setting, feedback the participants received or did not receive about the groups actions increased or decreased task performance, while it did not in same way in the individual task execution [23]. Similarly, additional knowledge about which object in a joint task is being looked at by a person interacting with a robot has been known to improve the interaction performance [24]. Furthermore, it has been recognized in the same experiment, that feedback about the robot's intention would have helped the user interacting with it.

Transferred to the exoskeleton to human interaction, this leads to believe, that active two-way communication alongside perception of intention is of great importance. By allowing the user as well as the exoskeleton to create or improve their model about the joint task, they can anticipate the next steps of the other.

Anticipation based on a shared model is crucial as it is reasonable to assume, that similar to common dynamic tasks, like playing soccer, the interaction between exoskeleton and user will not be of great performance unless the user as well as the exoskeleton are able to make predictions about the next steps of the other.

Assuming the user is the more intelligent and flexible of the two then providing feedback to the user and allowing him to adapt is of greater importance, than making the exoskeleton adapt to the user and the task. While for slow and continuous adaption processes the user can adapt to changing exoskeleton behavior, the same is not true for rapid changes in response to external triggers. Therefore, one shall explore if additional feedback allowing for anticipatory action by the user is possible to generally allow the exoskeleton for rapid changes in characteristics. This could take the form of making the exoskeleton signal to the user its potentially limited understanding of the world and its intended future reactions (like a notice just before switching assist mode on/off).

3.2 What is the Goal of Learning and How Can we Formalize This Goal?

Assuming we can't explicitly define the relationship between environmental perception and action or want to optimize parameters in an otherwise known model, one must define a goal for the optimization of torque estimation in terms of an objective function.

In the problem statement (1) we have defined, that we need to calculate the target support torque τ_opt, which is to be optimal in regards to variety of criteria such as biomechanical parameters, subjective parameters as well as task performance criteria.

Evaluating exoskeletons and comparing them to another over a variety of scenarios has been shown to be challenging even with a lot of expert knowledge [4]. Consequently, the same will hold true for any attempts to formalize the performance criteria in a way, that they can be measured and used for automatic optimization of actions. Nonetheless, single parameters have been used to formalize the goal and change the dynamic behavior during runtime [22].

4 Framework for Adaptive Exoskeletons

We propose a framework inspired by the interaction in groups of humans and human–robot-collaboration. The framework is presented in the form of a sense-plan-act paradigm and will establish physical and conceptual elements of the exoskeleton and its environment (see Fig. 2), while we discuss the requirements for these.

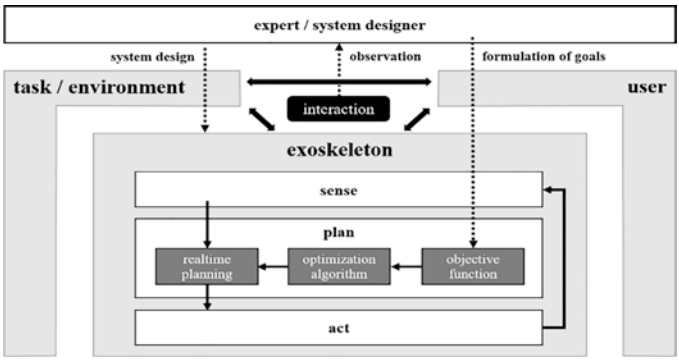


Fig. 2. Sense-plan-act paradigm is embedded in the interaction of user, exoskeleton, and task.

4.1 Observation and Feedback Methods as a Prerequisite to Optimize the System

When thinking about feedback one can classify feedback methods by their temporal resolution. Some feedback methods, such as questionnaires or observation by experts will give a performance metric, that is directly usable, but it will be difficult to obtain a lot of data and get information about the behavior at certain albeit important parts of the process (i.e. lifting starts). On the other hand, methods such as EMG will give a good temporal resolution. Evaluation methods tied to global physiological parameters like heart rate or VO2 will give a good estimate of the performance, but only limited temporal resolution as well.

Though being simple in nature, one shall not forget deliberate feedback from the user, like intentional direct (often spoken) communication is used in group interactions. It has been shown that a button mounted to a power tool, which the user could use to switch between different levels of assist, can improve the perceived interaction performance [7]. While this type of real time feedback is clearly not applicable in every scenario, generalizing the concept towards learning the user's interaction with an HMI and automating the interaction or assisting the user with it can be promising.

To overcome the limitations of individual methods one will have to combine them to multi-modal schemes. This can allow for both high temporal resolution and high relevance to actual and perceived reduction in effort. It is important to note, that while some parameters such as EMG, can be used as part of estimating system performance, other environmental data types (such as weight of the tool) can be crucial in the sensing part to improve the performance by inferring their relationship to adequate action, but cannot serve to evaluate the performance.

Necessary next steps will include ways to integrate expert knowledge in an objective function, that represents the different stakeholders.

4.2 Sensing Methods for Exoskeletons

Estimating the optimal torque has been linked in the past to detecting the user's intention. Especially for rehabilitation exoskeletons a variety of bio-signal based control methods have been studied [25]. While there are examples of using physiological phenomena (most notably EMG) in literature [16], for industrial exoskeletons most of the sensors measuring physiological phenomena are not feasible for permanent field deployment but rather assessment in a lab [13]. Easier to use in industrial settings however are sensors tracking the human pose. Active exoskeletons can measure relative orientation of actuated joints and sometimes non-actuated joints. These relative positions and their respective time derivatives alongside orientation estimated from additional body mounted inertial measurement units (IMUs) [17] can be used to detect relevant events [13]. Apart from user centered information one can potentially look at task related information. The weight, type, etc. of the handled tool as well as other parameters such as relative orientation to the work area can be sensed or received in an industry 4.0 environment [26]. In human robot collaboration a large variety of environmental and/or task related information has been explored for meaningful cues [27]. Using this groundwork will clearly be beneficial in sensing the user and the environment, as the underlying problem is similar.

4.3 Planning Methods for Exoskeletons

Of course, the planning will always be tightly linked to the perception method. However, some general classifications can be made. Firstly one can decide between methods that provide insight (white-box) models by having a limited set of rules, that a system designer can understand, while on the other hand there are black box models, that only give input–output relations and don't give you meaningful insight to their inner workings.

To achieve a working interaction, we must ensure, that the behavior of the exoskeleton is sufficiently predictable for the operator, such that he can create an internal model of the system and allow him to perform anticipatory action. Wherever this is not feasible, because of the underlying complexity of the task, it is important, that the planning algorithm provides sufficient insight, such that the operator can be given feedback about the systems state and know what behavior to expect.

To give the user insight into the internal states of the exoskeletons torque estimation, we propose to structure it as a gain scheduling controller [see Fig. 3]. Assuming in each situation where the tasks general requirements do not change, one can primarily derive the target interaction force from the joint angles. However, any changes in these parameters by the gain scheduler shall be communicated to the user either by dedicated feedback through an HMI or by dedicated force feedback, which the user can decode.

4.4 Acting Methods for Exoskeletons

The prime acting method of an exoskeleton is applying force to its actuators. This serves both the purpose of meeting the primary design goal and allows for distinctive haptic feedback. This can be used to feedback internal states by modulating the target force (see Fig. 5) Other forms of acting might comprise visual or audio feedback.

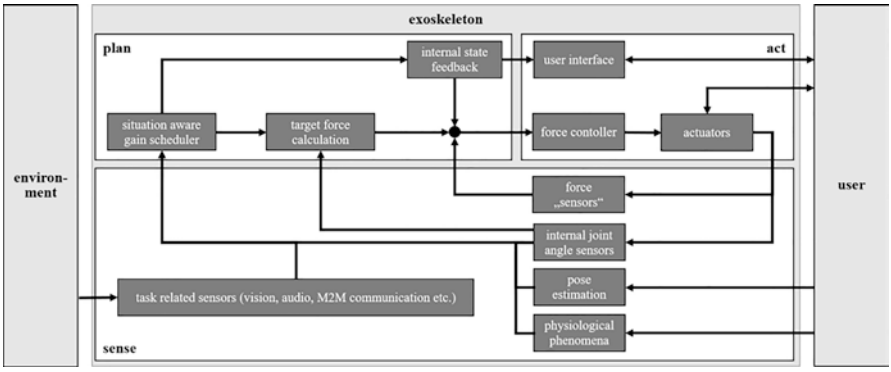


Fig. 3. Sense-plan-act applied in a control loop allowing for environmental awareness and feedback of internal states to the user as force feedback or via a user interface

5 Conclusion

By looking at other types of interaction in joint tasks, we have established the need for feedback and predictability. While previous work in adaptive exoskeletons has mainly focused on sensing, planning, and learning of the exoskeleton, we have highlighted that feedback to user is of at least the same importance as the perception of the exoskeleton. Furthermore, we have discussed the role of the system designers in formalization of goals for an iterative optimization algorithm.

To allow the exoskeleton to be predictable, we proposed a situation aware gain scheduling controller, that allows the slow adaption within a task set and signals the user any rapid change between operation modes via multi-modal feedback.

Acknowledgements. Parts of this research are funded by the Federal Ministry of Education and Research (BMBF) in the project "smart ASSIST – Smart, Adjustable, Soft and Intelligent Support Technologies" (funding number 16SV71114) and "Exo@Work - Influences of Exoskeletons on the workplace" (funded by the German employers' liability insurance association (BGHW)). The authors are solely responsible for the manuscript content.

References

1. Bogue, R.: Exoskeletons – a review of industrial applications. Industrial Robot: An Intern. J. **45**(5), 585–590 (2018)
2. Fox, S., Aranko, O.: Exoskeletons: Comprehensive, comparative and critical manufacturing performance. J. Manuf. Technol. Manag. (2019)
3. de Looze, M.P., Bosch, T., Krause, F., Stadler, K.S., O'Sullivan, L.W.: Exoskeletons for industrial application and their potential effects on physical work load. Ergo **59**, 1–11 (2016)
4. Hoffmann, N., Argubi-Wollesen, A., Linnenberg, C., Weidner, R.: Towards a Framework for Evaluating Exoskeletons. In: Production at the Leading Edge of Technology, pp. 441–450. Springer, Berlin (2019)
5. Weidner, R., Kong, N., Wulfsberg, J.P.: Human hybrid robot: a new concept for supporting manual assembly tasks. Prod. Eng. **7**(6), 675–684 (2013)
6. Weidner, R., Karafillidis, A.: Distinguishing support technologies. A general scheme and its application to exoskeletons. In: Karafillidis, A., Weidner, R. (eds.) Developing Support Technologies, pp. 85–100. Springer, Cham (2018)
7. Otten, B., Weidner, R., Argubi-Wollesen, A.: Evaluation of a novel active exoskeleton for tasks at or above head level. IEEE Robot. Autom. Lett. **3**(3), 2408–2415 (2018)
8. Argubi-Wollesen, A., Weidner, R.: Biomechanical analysis: adapting to users' physiological preconditions and demands by the use of biomechanical analysis. In: Karafillidis, A., Weidner, R. (eds.) Developing Support Technologies, pp. 47–62. Springer, Cham (2018)
9. Maurice, P., Ivaldi, S., Babic, J., Camernik, J., Gorjan, D., Schirrmeister, B., Bornmann, J., Tagliapietra, L., Latella, C., Pucci, D., Fritzsche, L., Ivaldi, S., Babic, J.: Objective and subjective effects of a passive exoskeleton on overhead work. IEEE Trans. Neural. Syst. Rehabil. Eng. **28**(1), 152–164 (2020)
10. Spada, S., Ghibaudo, L., Gilotta, S., Gastaldi, L., Cavatorta, M.P.: Investigation into the applicability of a passive upper-limb exoskeleton in automotive industry. Procedia Manuf. **11**, 1255–1262 (2017)

11. Amandels, S., Op het Eyndt, H.O., Daenen, L., Hermans, V.: Introduction and testing of a passive exoskeleton in an industrial working environment. Adv. Intell. Syst. Comput. **820**, 387–392 (2019)
12. Baltrusch, S.J., van Dieën, J.H., van Bennekom, C.A.M., Houdijk, H.: The effect of a passive trunk exoskeleton on functional performance in healthy individuals. Appl. Ergon. **72**, 94–106 (2018)
13. Chen, B., Grazi, L., Lanotte, F., Vitiello, N., Crea, S.: A real-time lift detection strategy for a hip exoskeleton. Front. Neurorobot. **12**, 1–11 (2018)
14. Lockheed Martin. Homepage https://www.lockheedmartin.com/en-us/products/exoskeleton-technologies/industrial.html. Acessed 28 March 2020
15. Van Engelhoven, L., Kazerooni, H.: Design and intended use of a passive actuation strategy for a shoulder supporting exoskeleton. In: WearRAcon, pp. 7–12 (2019)
16. Koopman, A.S., Toxiri, S., Power, V., Kingma, I., van Dieën, J.H., Ortiz, J., de Looze, M.P.: The effect of control strategies for an active back-support exoskeleton on spine loading and kinematics during lifting. J. Biomech. **91**, 14–22 (2019)
17. Yao, Z., Linnenberg, C., Weidner, R., Wulfsberg, J.P.: Development of a soft power suit for lower back assistance. Proceedings of the IEEE International Conference on Robotics and Automation, pp. 5103–5109 (2019) (2019)
18. Mason, M.T.: Toward robotic manipulation. Annu. Rev. Cont. Robot. Auton. Syst. **1**(19), 1–28 (2018)
19. Sebanz, N., Knoblich, G.: Prediction in joint action: what, when, and where. Top. Cogn. Sci. **1**, 353–367 (2009)
20. Vesper, C., van der Wel, R.P., Knoblich, G., Sebanz, N.: Are you ready to jump? Predictive mechanisms in interpersonal coordination. J. Exp. Psychol. Hum. Percept. Perform. **39**(1), 48–61 (2013)
21. van der Steen, M.C., Keller, P.E.: The ADaptation and Anticipation Model (ADAM) of sensorimotor synchronization. Front. Hum. Neurosci. **7**, 253 (2013)
22. Peternel, L., Noda, T., Petrič, T., Ude, A., Morimoto, J., Babič, J.: Adaptive control of exoskeleton robots for periodic assistive behaviours based on EMG feedback minimisation. PLoS One **11**(2), 1–26 (2016)
23. Knoblich, G., Jordan, J.S.: Action coordination in groups and individuals: learning anticipatory control. J. Exp. Psychol. Learn. Mem. Cogn. **29**(5), 1006–1016 (2003)
24. Huang, C.M., Mutlu, B.: Anticipatory Robot Control for Efficient Human-Robot Collaboration. In: ACM/IEEE International Conference on Human-Robot Interaction, pp. 83–90 (2016)
25. Lobo-Prat, J., Kooren, P.N., Stienen, A.H., Herder, J.L., Koopman, B.F.J.M., Veltink, P.H.: Non-invasive control interfaces for intention detection in active movement-assistive devices. J. Neuroeng. Rehabil. **11**(168), 1–22 (2014)
26. Bances, E., Schneider, U., Siegert, J.T.B.: Exoskeletons towards industrie 4.0: benefits and challenges of the IoT communication architecture. Procedia Manuf. **42**, 49–56 (2020)
27. Ajoudani, A., Zanchettin, A.M., Ivaldi, S., Albu-Schäffer, A., Kosuge, K., Khatib, O.: Progress and prospects of the human–robot collaboration. Auton. Robots **42**, 957–975 (2018)

Machine Learning

Research on Preprocessing Methods for Time Series Classification Using Machine Learning Models in the Domain of Radial-Axial Ring Rolling

S. Fahle[1(✉)], A. Kneißler[2], T. Glaser[1], and B. Kuhlenkötter[1]

[1] Chair of Production Systems, Ruhr-University of Bochum, Universitätsstraße 150, 44801 Bochum, Germany
fahle@lps.rub.de

[2] Thyssenkrupp Rothe Erde Germany, Tremoniastraße 5-11, 44137 Dortmund, Germany

Abstract. Machine learning models trained to predict certain outcomes bear great potential in a variety of applications. This research takes a step to elevate the hot forming technology of radial-axial ring rolling towards a fully digitalized and even more efficient forming technology. For successful machine learning the preprocessing step is essential. This paper presents current research regarding the most promising preprocessing approaches of time series data for the specific use case of classifying form errors of the radial-axial ring rolling process. By predicting form errors (in-situ), scrap and rework rates can be lowered due to an alert by the model for form errors in advance of a potential error, thus contributing to a more efficient industry. The data used exists in form of time series from log-data of an industrial used, single ring rolling machine. Concluding, the proposed preprocessing approaches are evaluated by comparing different model performances, trained on actual production data.

Keywords: Machine learning · Radial-axial ring rolling · Preprocessing

1 Introduction

Machine learning models achieved great results in a variety of research areas but have yet to proof their explicit usage in many production environments and use cases [1]. By merging the highly complex and multidimensional hot forming of radial-axial ring rolling (RARR) with machine learning models to predict a quality outcome of the process, a step is taken to push data driven models into the area of ring rolling. The complexity of the process is high, mainly to temperature effects in the hot forming of the ring at about 1200 °C, the two opposing rolling gaps (radial and axial) as well as the influence of improper stock materials and variations in the process chain. In the area of radial-axial ring rolling, the use of data driven methods such as machine

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 487–496, 2021.
https://doi.org/10.1007/978-3-662-62138-7_49

learning models has not been popular, due to the fact that this sector lacks a broadly applied data analysis strategy [2]. This is why the proposed research carried out on a dataset of real production data (of thyssenkrupp rothe erde) is a big step forward to proving the usability of machine learning models for the production of radial-axial rolled rings. The presented approach of preprocessing the ring rolling dataset consists of a comparison of the three standardization methods: min–max-scaling, z-score-normalization and no preprocessing at all. In addition to those normalization/standardization methods, the models at hand are used to compare different approaches to align all time series to an equal length. The data at hand is real-world production data and every sample is unique as geometric shapes differ, thus do process times resulting in different length time series. A general approach for dealing with varying length time series is given by TAN ET AL. [3]. The overall structure of the paper ranging from data-acquisition on the process-level to validation of the proposed methods is depicted in Fig. 1.

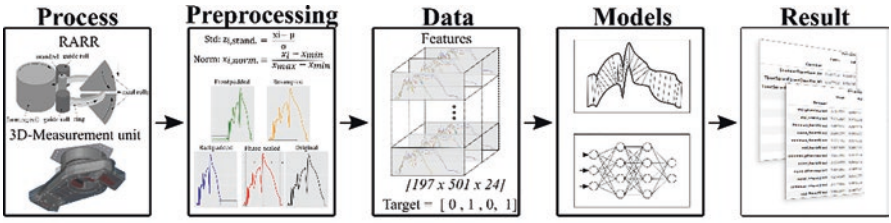


Fig. 1. Concept figure of conducted research using machine learning in RARR

2 Related Work

Artificial intelligence (AI) and especially methods of machine learning (ML) as a subcategory of AI have become the focus of research and spotlight within the industry in recent years. A lot of euphoria is certainly expressed due to the success in the field of image recognition, yet recent domains in the factory environment using ML are scheduling problems [4, 5], predictive quality [6, 7] and process control and optimization [8, 9]. The field of time-series analysis used for process control and optimization by HWANG ET AL. is an subtopic of ML with great potential for manufacturing processes [9] as modern machines provide a lot of data in form temporally related sensor readings of the manufacturing process. This success of machine learning in general and time series analysis in particular are the baseline motivation for the presented approach.

As for general research in the subdomain of time series classification (TSC), new algorithms have been developed in the last years such as the COTE-Ensemble by BAGNAL ET AL. [10]. Later, additional developments were carried out by means of a hierarchical voting system resulting in HIVE-COTE by LINES ET AL., which has become one of the state-of-the-art algorithms/ensembles for TSC [11]. A comparison of the performance of different algorithms was performed on different time-series benchmark data sets of the UCR Time Series Archive [12, 13]. Due to the success of neural networks, especially in the field of deep learning (neural networks with several hidden layers) in many areas, they have been adapted to the field of TSC. FAWAZ

ET AL. have investigated the applicability of neural networks (especially multi layer perceptron; convolutional neural networks and echo state networks) for the solution of TSC related problems [14]. The presented success of deep learning models is the reason for the implementation of different deep learning models into the conducted experiments that are described in Sect. 3.5.

Within TSC, there is the early time series classification, meaning a classifier issues a classification of the state before the end of the actual time series [15]. This form of TSC is particularly suitable for the research project under consideration here, in which a production process is predictive or even preventive on-line, i.e. already during the production process, in order to obtain a statement about the result with regard to a quality criterion. The research in this area ranges from different approaches to generate the temporal inputs (feature) by MORI ET AL., to an analysis of multivariate instead of merely univariate time series by HE ET AL. and LIN ET AL. [16–18]. A different approach has been taken by RUßWURM ET AL. by proposing a complete end-to-end approach using neural networks. The advantage of such an approach is the reduced manual effort of extracting temporal inputs of the time series [19]. The presented transition from a time series classification task to an early time series classification task could be a viable option for the field of RARR in the future and needs to be analyzed further.

However, the application of TSC in the hot forming sector of RARR has yet to be further researched as shown by a study in 2019 [2].

3 Preprocessing of RARR Data for TSC

Before going into further detail about the methods and models that are being used, it needs to be stated that the aim of this paper is not to find the one and only, best working algorithm with already tuned hyperparameters. The aim of this paper is to investigate the best approach of working with RARR data regarding an alignment of lengths and normalization/standardization approaches. Neither is the evaluation of best input parameters, nor finding the best model the aim of this paper, as this will be addressed in future work once a bigger dataset becomes available. Due to this fact, the models explained and later trained are not tuned at all so they give a general answer to the question which of the chosen methods works better in this specific domain of time series data in RARR.

3.1 Definition of the TSC-Problem

Before starting with the description of the dataset and preprocessing steps that are being used, a definition of the used TSC-problem for RARR in reference to FAWAZ ET AL. and their definition of TSC must be given [14]. The complementary definitions are the following:

Definition 1: A univariate time series $X = [x_1, x_2, \ldots, x_T]$ is a timestamp ordered set of real values with a sampling rate of 10 Hz. The length of X is equal to the real values T depending on the individual rolling time of each ring in milliseconds. Each univariate time series represents a logged channel of the RARR process.

Definition 2: A dataset $D = \{(X_1, Y_1), (X_2, Y_2), \ldots, (X_N, Y_N)\}$ is a collection of corresponding pairs (X_i, Y_i) with X_i either being a time series and Y_i its target variable vector. The classes for now are defined as a binary classification task of either "no ovality" or "ovality".

In the following, Definition 1 and Definition 2 are used to form a baseline dataset containing time series (of different length) in their raw format with their corresponding target variable. This baseline dataset and the conducted preprocessing methods to form twelve different datasets of different preprocessing combinations are discussed further.

3.2 Dataset

The dataset consists of 197 representative log-samples from industrially rolled rings at thyssenkrupp rothe erde. Every rolled ring was measured using a laser-based 3D-measurement unit. From this quality inspection a target regarding ovality was derived for every sample. The target was calculated by using an internally defined threshold and the measurement data for each rolled ring. The dataset consists of representative samples forming an evenly distributed dataset, resulting in 99 samples with no ovality and 98 samples with ovality. Furthermore, all samples of the dataset were initially validated regarding their usability as comparable conditions have to be guaranteed. Such conditions can either be a defect in the production chain resulting in a not representatively rolled ring or a machine error due to which too much time passes before the already cooled ring is measured resulting in different geometric results.

Due to the (up till now) small size of the dataset a randomly distributed train-test-split ration of 80/20 was chosen. The 197 representative samples are taken from one production day and represent eight different geometrical shapes in total with differences in diameters, wall thickness and height.

Experiments are based on the ring-growth-rate input feature resulting in a dataset of univariate time series of different lengths. The dataset is subject to a confidentiality agreement and must therefore not be made available to the public.

3.3 Feature Scaling

For the aspect of feature scaling within data preprocessing there are three different approaches used. All of the later described are validated in the experiment Sect. 4.

No Scaling at all was used to keep raw data and see whether there is an improvement or decrease in accuracy as compared to this baseline approach.

Standardization or Z-score normalization rescales all values within one univariate time series so that this univariate time series has zero mean and a standard deviation of one. This scaling helps to compare different measurements of different units (especially for a multivariate approach). This is done using the following equation:

$$z_{i,standardized} = \frac{X_i - \mu}{\sigma} \tag{1}$$

This results in time series that all share the characteristic of having zero mean and a standard deviation of one, yet their ranges are not fixed between specific limits.

Min–max scaling or normalization rescales all values within one univariate time series to a fixed range. The range used is [0, 1] and is done by the following equation:

$$x_{i,normalized} = \frac{x_i - x_{min}}{x_{max} - x_{min}} \tag{2}$$

A combination of standardization and normalization (i.e. rescaling the range of an already standardized dataset) is not conducted as the individual feature scaling influence in conjunction with the following resampling methods is the focus of the conducted research.

3.4 Resampling

In the following, three common techniques for aligning time series data to an equal length are presented, as well as one proposed technique for the use case of RARR.

Backpadding is a technique were each individual time series is filled with zeros at the end of the time series to a maximum length depending on the longest time series available in the dataset.

Frontpadding is a technique were each individual time series is filled with zeros at the front of the time series to a maximum length depending on the longest time series available in the dataset.

Uniform scaling is a technique were each individual time series is interpolated to a given step size. To apply this to all time series a TimeSeriesResampler by Tavenard et al. was implemented [20].

Phase-Scaling is an individually developed technique to stretch each time series to a total length that equals the longest time series available in the dataset. This technique was developed with regard to domain expertise, utilizing the rolling state bit value inside each log-dataset. This rolling state bit value tracks the present rolling phase which can be divided into the four phases of: roll-up phase, main phase with constant ring-growth rate, main phase with reduced ring growth-rate and rounding phase [21]. The phases were scaled according to statistical characteristics of the entire data set. Those characteristics are mean and standard deviation of the phase length relative to each individual sample length. The characteristics were extracted using the whole dataset and have formed the basis for the phase-scaling approach on each individual sample.

Figure 2 visualizes all proposed techniques using the same sample from the dataset and transforming it using the described approaches.

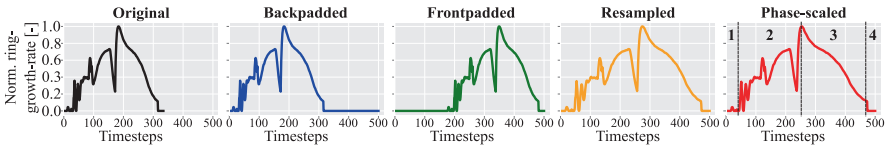


Fig. 2. Comparison of all techniques used to align RARR time series

This results in time series with comparable rolling phases and thus resulting in comparable time series of equal lengths. The advantages in mind while developing this technique are: for distance based algorithms the addition of zeros at the end or front of the time series might influence their performance as time series can be equal in length thus needing same padding but being completely different regarding their rolling process and quality outcome. Moreover, the process must consist of the regarded phases and therefore all equally long phases throughout all time series should increase comparability between different geometries and thus time series lengths. The biggest noticeable effect is the alignment in the rounding phase as it can be seen in Fig. 3 by all time series dropping at the exact same timestamp at about 480.

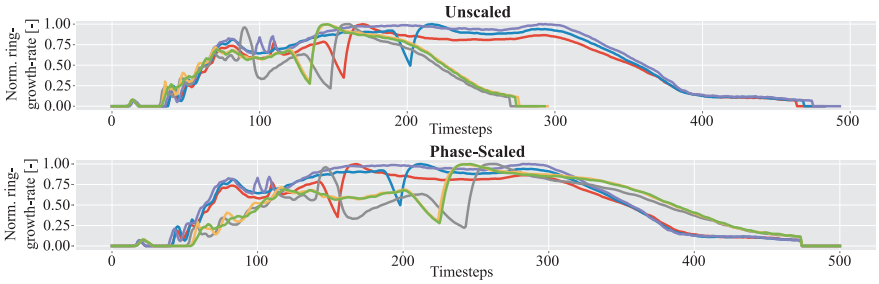


Fig. 3. Phase-scaled time series in contrast to original time series

3.5 Used Models

Regarding the models used in the later experiments it was ensured that models of different types were implemented and used for the validation of the preprocessing steps. All implementations of algorithms were implemented in Python using different libraries especially sktime by LÖNING ET AL. [22].

Distance-based models are algorithms where there is some kind of distance metric used to compare time series. The classifier then uses the distance measurement as the input for its classification task. As for distance-based models a k-Nearest-Neighbor with dynamic time warping as its distance metric and a proximity forest have been implemented [23].

Dictionary-based represented by BOSS in our case uses windows to slide over time series and represent those in the form of a word. In contrast to other models, BOSS uses a discrete fourier transform on each window and subsequences are discretized using multiple coefficient binning [24].

Interval-based models use summary statistics (mean, standard deviation and slope) of intervals as features for a random forest approach. This paper used two different time series forests with 50 and 100 numbers of trees implemented [25].

Shapelet-based models transform time series into small shapelets. This shapelets then serve as local, phase-invariant subseries. Further, these shapelets are then analyzed regarding similarities. The implemented shapelet model consists of a shapelet transform in conjunction with a random forest [26].

Deep-learning based models represent neural network architectures with several "hidden" layers. The implemented architectures of a multi layer perceptron (MLP),

time convolutional neural network (CNN), fully convolutional neural network (FCN) and resnet architecture were originally implemented by FAWAZ ET AL. [14].

4 Experiments

To validate the earlier described approaches of preprocessing steps for the task of classification within RARR, several models were trained. To further support the statistics of the conducted experiments the shown results represent the aggregated mean and standard deviation with regard to all models' accuracies of five separate model trainings each. The analysis will be separated with special regard to the best working dataset for all algorithms as well as a grouped result by each model showing its best working dataset.

The conducted experiments and their results are depicted in Table 1. The results are ordered descending and include all conducted model-trainings. As there are twelve datasets and eleven classifiers used with five model trainings each (12 datasets, 11 classifiers, 5 runs = 660 individual model trainings), Table 1 shows the accuracy of all classifiers and all model trainings grouped together by datasets for the test set. The accuracies of all runs and classifiers were aggregated into the columns "Mean" and "Std" as the intent is to find the best preprocessing technique for a dataset in RARR regarding classification accuracy in general and not the best working classifier, yet. Only the "MAX" column shows a non-aggregated value displaying the highest "single run accuracy" for each dataset by any classifier on the test set.

It can be seen that the proposed technique of phase-scaling works best of all techniques in conjunction with a standardization, regarding not only mean accuracy but also scoring the highest over-all single run accuracy with 0.825. The standard deviation for the proposed technique in conjunction with standardization is mediocre and by far neither the best nor worst. Notably, the lowest standard deviation of all datasets is reached by the raw features in conjunction with phase-scaling with a value of 0.046.

Table 1. Experimental results grouped by dataset

Dataset	Mean	**Accuracy** Std	Max (single run)
Standardization + Phase-scaled	**0.703**	0.059	**0.825**
Standardization + Resampled	0.676	0.079	0.775
Normalization + Backfilled	0.667	0.085	0.800
Raw + Frontfilled	0.648	0.068	0.800
Normalization + Frontfilled	0.633	0.051	0.750
Standardization + Backfilled	0.630	0.060	0.750
Normalization + Phase-scaled	0.625	0.062	0.725
Raw + Backfilled	0.624	0.069	0.750
Raw + Phase-scaled	0.620	**0.046**	0.700
Raw + Resampled	0.615	0.058	0.725
Normalization + Resampled	0.614	0.091	0.800
Standardization + Frontfilled	0.590	0.049	0.675

By looking at the results it can be seen that all alignment methods as well as all feature scaling approaches have a broad spread of accuracy. This leads to the proposal of the utilization of phase-scaling in combination with standardization as a baseline when starting a classification task in RARR regarding quality issues by scratch. This shows that the domain expertise in form of using a process specific feature as input for an alignment method results in better or at least similar results than generally proposed methods. Yet, this should only serve as a guideline and individual experiments should be conducted.

To illustrate the performance of all classifiers a critical difference diagram and its implementation by FAWAZ ET AL. is used with a Holm's alpha of 5% [14]. All classifiers have not been hyperparameter tuned and thus their hyperparameters remain the default values of their implementation in *tslearn* and *sktime(-dl)*, An exception has to be made regarding time series forest classifiers (TSF_50/TSF_100) and k-Nearest-Neighbor (1-NN/3-NN) classifiers as the number in their name indicates the number of neighbors (NN) or estimators (TSF) used.

The critical difference diagram of all 660 experiments shows the average computed ranks in terms of accuracy for the used classifiers on the top scale with the best classifier having the lowest rank (most right position). A thick horizontal line indicates a clique of classifiers that are not-significantly different with regard to accuracy.

As can be seen in Fig. 4, the overall best working classifier is the time series forest with 50 estimators, yet there is no critical difference between all classifiers. The highest individual classification accuracy of 0.825 (cf. Table 1 column "Max") was reached twice. Once by the shapelet transformation model and once by the time series forest model with 100 estimators, both on the standardized and phase-scaled dataset. The lack of a critical difference using the the wilcoxon-holm pairwise test can occur due to the small size of used datasets as well as the (intentionally) left out hyperparameter tuning of the classifiers.

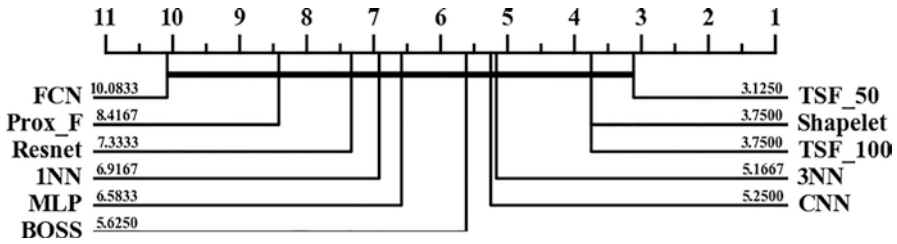


Fig. 4. Critical difference diagram for all classifiers

5 Conclusion

We have conducted 660 different model trainings to investigate the best preprocessing steps for a time series classification task in the domain of RARR on a real world dataset from industry. The dataset consists of (as for now) 197 samples of thyssenkrupp rothe erde with variable lengths. We propose a domain expertise based alignment method for time series in RARR consisting of a ring rolling phase-scaling. The results

of the conducted experiments show that our proposed alignment method in conjunction with a standardization works very well, reaching the highest average accuracy score for all implemented models of 0.703 and highest single run maximum accuracy score of 0.825 as well. Yet, the difference to other combinations was not drastic enough, underlining the relevance of domain expertise in the domain of preprocessing in machine-learning tasks. A final recommendation cannot be given at this moment as the results do not show significant differences and are therefore in need of further analysis in the future.

Nevertheless, this work shows a clear trend of the proposed phase-scaling preprocessing method works very well for the used data of the RARR process and can therefore be used as a good baseline preprocessing method when starting a time series classification in RARR from scratch.

6 Future Work

Future work will be driven towards a bigger dataset by integrating more production data to investigate the problem of quality classification in RARR further. A focus will be on the capabilities of deep-learning models as dataset size grows. Moreover, further analysis will be conducted whether a multivariate approach results in better accuracies and thus to what extend different input features show higher significance.

Acknowledgements. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – 404517758

References

1. Fahle, S., Prinz, C., Kuhlenkötter, B.: Systematic review on machine learning (ML) methods for manufacturing processes – Identifying artificial intelligence (AI) methods for field application. Procedia CIRP (2020, in press)
2. Fahle, S., Kuhlenkötter, B.: A framework for data integration and analysis in radial-axial ring rolling. In: 1st Conference on Production Systems and Logistics (2020)
3. Tan, C.W., Petitjean, F., Keogh, E., Webb, G.I.: Time series classification for varying length series. Preprint available at: https://arxiv.org/abs/1910.04341, arXiv:1910.04341v1, 1-23. (2019)
4. Kim, D., Lee, T., Kim, S., Lee, B., Youn, H.Y.: Adaptive packet scheduling in IoT environment based on Q-learning. Procedia Comput. Sci. **141**, 247–254 (2018)
5. Lubosch, M., Kunath, M., Winkler, H.: Industrial scheduling with Monte Carlo tree search and machine learning. Procedia CIRP **72**, 1283–1287 (2018)
6. Lavrik, E., Panasenko, I., Schmidt, H.R.: Advanced methods for the optical quality assurance of silicon sensors. Nucl. Instrum. Methods Phys. Res. Sect. Accel. Spect. Detect. Assoc. Equip. **922**, 336–344 (2019)
7. Ma, L., Xie, W., Zhang, Y.: Blister defect detection based on convolutional neural network for polymer lithium-ion battery. Appl. Sci. **9**(6), 1085 (2019)
8. Du Preez, A., Oosthuizen, G.A.: Machine learning in cutting processes as enabler for smart sustainable manufacturing. Procedia Manuf. **33**, 810–817 (2019)

9. Hwang, S., Jeon, G., Jeong, J., Lee, J.: A novel time series based Seq2Seq model for temperature prediction in firing furnace process. Procedia Comput. Sci. **155**, 19–26 (2019)
10. Bagnall, A., Lines, J., Hills, J., Bostrom, A.: Time-series classification with COTE: the collective of transformation-based ensembles. IEEE Trans. Knowl. Data Eng. **27**(9), 2522–2535 (2015)
11. Lines, J., Taylor, S., Bagnall, A.: Time series classification with HIVE-COTE. ACM Trans. Knowl. Discov. Data **12**(5), 1–35 (2018)
12. Bagnall, A., Lines, J., Bostrom, A., Large, J., Keogh, E.: The great time series classification bake off: a review and experimental evaluation of recent algorithmic advances. Data Min. Knowl. Disc. **31**(3), 606–660 (2017)
13. Dau, H.A., Bagnall, A., Kamgar, K., Yeh, C.-C.M., Zhu, Y., Gharghabi, S., Ratanamahatana, C.A., Keogh, E.: The UCR time series archive. Available at: https://arxiv.org/abs/1810.07758 arXiv:1810.07758v2, 1–12 (2018)
14. Ismail Fawaz, H., Forestier, G., Weber, J., Idoumghar, L., Muller, P.A.: Deep learning for time series classification: a review. Data Min. Knowl. Disc. **59**(2), 195 (2019)
15. Xing, Z., Pei, J., Keogh, E.: A brief survey on sequence classification. SIGKDD Explor. **12**(1), 40–48 (2010)
16. He, G., Duan, Y., Peng, R., Jing, X., Qian, T., Wang, L.: Early classification on multivariate time series. Neurocomputing **149**, 777–787 (2015)
17. Lin, Y.F., Chen, H.H., Tseng, V.S., Pei, J.: Reliable early classification on multivariate time series with numerical and categorical attributes. In: Cao, T., Lim, E.P., Zhou, Z.H., Ho, T.B., Cheung, D., Motoda, H. (eds.) Advances in Knowledge Discovery and Data Mining, vol. 9077, pp. 199–211. Springer, Cham (2015)
18. Mori, U., Mendiburu, A., Keogh, E., Lozano, J.A.: Reliable early classification of time series based on discriminating the classes over time. Data Min. Knowl. Disc. **31**(1), 233–263 (2017)
19. Rußwurm, M., Lefèvre, S., Courty, N., Emonet, R., Körner, M., Tavenard, R.: End-to-end learning for early classification of time series. Preprint available at: https://arxiv.org/abs/1901.10681, arXiv:1901.10681v1, 1–9 (2019)
20. Tavenard, R., Faouzi, J., Vandewiele, G., Divo, F., Androz, G., Holtz, C., Payne, Marie, Yurchak, Roman, Rußwurm, M., Kolar, K., Woods, E.: tslearn: a machine learning toolkit dedicated to time-series data. https://github.com/rtavenar/tslearn (2017)
21. Thyssen: Technologiehandbuch: Einführung in die Ringwalztechnologie. Thyssen Wagner Maschinenbau GmbH, Dortmund (1990)
22. Löning, M., Bagnall, A., Ganesh, S., Kazakov, V., Lines, J., Király, F.J.: sktime: A unified interface for machine learning with time series. Available at: https://arxiv.org/abs/1909.07872, arXiv:1909.07872v1, 1-9 (2019)
23. Lucas, B., Shifaz, A., Pelletier, C., O'Neill, L., Zaidi, N., Goethals, B., Petitjean, F., Webb, G.I.: Proximity Forest: an effective and scalable distance-based classifier for time series. Data Min. Knowl. Disc. **33**(3), 607–635 (2019)
24. Schäfer, P.: The BOSS is concerned with time series classification in the presence of noise. Data Min. Knowl. Disc. **29**(6), 1505–1530 (2015)
25. Deng, H., Runger, G., Tuv, E., Vladimir, M.: A Time series forest for classification and feature extraction. Inf. Sci. **239**, 142–153 (2013)
26. Hills, J., Lines, J., Baranauskas, E., Mapp, J., Bagnall, A.: Classification of time series by shapelet transformation. Data Min. Knowl. Disc. **28**(4), 851–881 (2014)

Process Monitoring Using Machine Learning for Semi-Automatic Drilling of Rivet Holes in the Aerospace Industry

L. Köttner(✉), J. Mehnen, D. Romanenko, S. Bender, and W. Hintze

Institute of Production Management and Technology (IPMT), Hamburg University of Technology (TUHH), Denickestraße 17, 21073 Hamburg, Germany
lars.koettner@tuhh.de

Abstract. The majority of aircraft rivet holes are drilled with semi-automatic and manually controlled, pneumatically driven machines as full automation is often unsuitable due to workspace restrictions. Lightweight materials of difficult machinability complicate drilling. This is particularly relevant when drilling stack materials, where the machining parameters are determined by the most difficult to machine material layer. To provide reliable rivet connections, drilling in multiple steps, use of minimum quantity lubrication as well as subsequent manual deburring and cleaning are indispensable. Newly developed electrically driven semi-automatic advanced drilling units (ADUs) enable intelligent process layouts and online condition monitoring by evaluating integrated sensor data. Additionally, process parameters can be adapted to suit each material in the stack.

In this paper, machine learning is applied to ADU sensor data to predict cutting forces and process conditions based on the internally measured currents of the ADU's electric motors. The application of machine learning to ADU data is beneficial as drilling in the aerospace industry shows high repeatability and many produced holes, providing a large dataset. The machine learning methods linear regression, artificial neural network and decision tree are applied to force prediction. Furthermore, the k-nearest neighbour method is used to predict material, feed rate and lubrication state. Process monitoring based on the presented results minimizes manual control and rework by the identification of process deviations, resulting in a comprehensive quality assurance as well as optimal tool life exploitation. This leads to a step change in semi-automatic drilling of aircraft structures by overcoming a major productivity limitation.

Keywords: Precision drilling · Machine learning · Condition monitoring

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 497–507, 2021.
https://doi.org/10.1007/978-3-662-62138-7_50

1 Introduction

In spite of the current situation in the aerospace industry, forecasts predict a medium-term rebound in the world aircraft demand in addition to already high backlogs of aircraft orders, requiring a future growth of production rates on the industries' existing assembly lines [1]. Here, the manufacturing of precision-boreholes for riveting in structural assembly is crucial with respect to lead time and cost. The automation level of the associated drilling processes is already high in easily accessible work environments such as shell production. Other workspaces such as closed structures (boxes) and final assembly lines only allow for medium or low degrees of automation. In these workspaces, the use of novel, electrically driven, semi-automatic Advanced Drilling Units (ADUs) is expected to contribute to an increase in productivity and reliability. In contrast to the currently used pneumatically driven ADUs, process parameters can be adapted along the feed path, offering complete control of the drilling process. This leads to a quality increase, reducing the need for rework or even part replacement. Moreover, the introduction of electric drives is supposed to replace time- and cost-intensive manual quality monitoring by in-process condition monitoring based on internal sensor data, presumably reducing manufacturing concessions. These cost Airbus approximately 400 million EUR per year, while 50% of the concessions in the production of the Airbus A350 were due to drilling errors [2].

State of Technology. Since the 1990s, several authors have shown that machine learning (ML) may be applied to machining processes. ML methods such as artificial neural networks (ANN) enable even complex relationships to be identified from evaluated external sensors' data, e.g. to predict tool wear [3]. Depending on the situation, data from internal machine sensors can also provide sufficient accuracy, e.g. to predict specific cutting forces in milling only from the utilization factor of spindle power [4]. Recent developments of ML in machining are presented by [5]. The investigations are particularly far-reaching in the fields of milling and turning [5, 6].

ML has also been applied to study drilling processes. Neto et al. used a multilayer perceptron to predict borehole diameters in Ti6Al4V/Al2024 stacks using data of cutting force, spindle motor power, acoustic emissions and process vibrations [7]. Shaban et al. estimated typical quality parameters during drilling of quasi-isotropic CFRP laminates based on feed force and cutting torque data [8]. Kim et al. used statistical quantities of the spindle motor power to classify the state of chip disposal in S45C steel with ANNs [9]. Caggiano et al. extracted fractal and statistical features of feed force and cutting torque to predict flank wear land width VB with ANNs [10]. Another study by Caggiano et al. in [11] predicts VB with ANNs after using principal component analysis on statistical features of feed force, cutting torque and acoustic emission.

Objectives. In this context, the paper focuses on the implementation of process monitoring in semi-automatic drilling using ML. In contrast to the presented state of the art, only internal sensor data, specifically internally measured motor currents, are used to determine process forces and monitor process properties (e.g. feed rate). The scope of the investigation is deliberately close to aircraft industry standards. Therefore, it provides the basis for intelligent, semi-automatic drilling processes with electrical ADUs, which are expected to significantly reduce process times while increasing drilling quality and tool life exploitation.

2 Experiments and Process Database

Drilling experiments were conducted to provide a database for ML investigations. Electric currents of internal motors for feed and rotational spindle movement of the SmartADU (a new type of ADU from *Johannes Lübbering GmbH*) and external feed force and cutting torque measurements by a dynamometer *Kistler* Type 9272 were recorded separately. The test setup in Fig. 1 features a test rig in which the ADU is clamped in a drilling mask above the test material, imitating the industry's operating conditions. Material samples are clamped on a jig which is bolted to the dynamometer. To mechanically decouple ADU and dynamometer, the machine's clamping system (concentric collet) was kept at a small distance to the material sample, allowing only the drill bit to transmit forces. The ADU's peck feed system superimposes axial oscillation on the conventional tool rotation at a frequency equal to 1.5 times the spindle speed. This improves chip removal and reduces damage to the bore surface. Laser sensors *Micro-Epsilon* ILD2300-10 were used to guarantee stable peck feed and cutting conditions.

In contrast to other industrial electric ADUs, SmartADU has two DC motors, a 180 W motor for the feed movement and a spindle motor with a rating of 2 kW, predominantly used for turning the spindle. Via a supply box, the ADU is connected to power supply, compressed air for internal tool cooling, minimum quantity lubrication (MQL) and the central control system which is used to set drill parameters and evaluate the two motor currents measured by internal sensors.

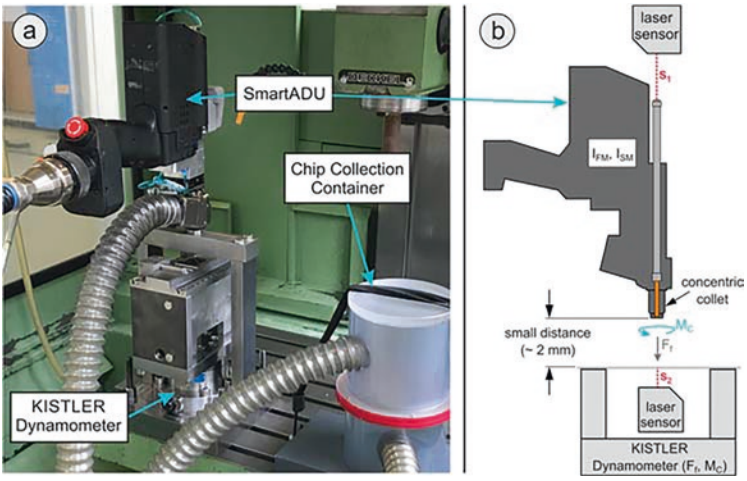


Fig. 1. System components of the experimental setup for drilling with SmartADU, see [12]

A diamond-coated cemented carbide tool with a diameter of 4.78 mm and a point angle of 120° was selected for the experiments. Test stripes from two typical aerospace materials, aluminium (Al7075, t = 6 mm) and titanium (Ti6AlV4, t = 4 mm) were drilled. Cutting parameters for the test series listed in Table 1 are based on industrial practice. Motor currents of the ADU were internally recorded at its

maximum sample rate of 100 Hz while cutting torque and feed force were externally measured at a sample rate of 5 kHz (separate recordings without time synchronization were necessary due to missing sensor data output of the ADU). MATLAB was used for data pre-processing, model design, calculations, and evaluations. To avoid information losses of dynamometer data, the ADU data was interpolated to a sampling rate of 5 kHz.

Data pre-processing of sensor signals is crucial before using them as input parameters for ML algorithms. For this purpose, the signals of ADU and dynamometer were first synchronized using signal correlation between feed motor current and feed force. Subsequently, trimming to the relevant signal range from just before tool entry to just after tool exit was performed. Low-pass filtering of the data was applied to remove noise and high-frequency (e.g. peck feed) oscillations, thus obtaining the mean measured motor currents as input parameters for the ML models. For comparison, two different filters were applied for this purpose: A moving average (window size: 3500 samples), as typical and frequently used smoothing filter and a first-order Butterworth filter (cut-off frequency: 0.5 Hz), which is already implemented in the internal data processing of the ADU. Feature scaling was applied to the data before splitting it into sets for hold-out validation, which is suitable for large datasets [13]. For training, cross-validation (CV) and test sets, bin sizes were 60%, 20% and 20%, similar to [10].

Table 1. Drilling experiments and cutting parameters

Exp.	Amount	Material	n [rpm]	f [mm/rev]	PF [%]	MQL [%]
1	30	Ti6Al4V	600	0.03, 0.05, 0.08	100	0
2	30	Ti6Al4V	800	0.03, 0.05, 0.08	100	0
3	15	Ti6Al4V	600	0.05	100	1, 50, 100
4	10	Al7075	600	0.05	100	0

3 Application of Machine Learning Methods

Based on the prepared data sets, the performance of LR, ANN and DT to predict cutting forces is compared. Afterwards, the KNN method is applied to process monitoring.

As force and process monitoring is supposed to be continuous in the future, i.e. possible at any time and in any operating situation (tool entry, full cut, tool exit), real time values of available features are used in the following approaches.

3.1 Feed Force and Cutting Torque Prediction

Feed force F_f and cutting torque M_C are the characteristic quantities of the drilling process and allow the estimation of process properties such as tool wear [10, 11]. Thus, their accurate prediction is necessary for the assessment of process quality if direct measurements (e.g. with load cells) are not applicable, as is the case with ADUs.

The prediction of F_f and M_C using currents of feed motor I_{FM} and spindle motor I_{SM} is based on the correlation of mechanical and electrical power

$$P_{mechanical} = P_{electrical} - P_{loss,motor} - P_{loss,other} = U \cdot I - R \cdot I^2 - P_{loss,other}, \quad (1)$$

assuming DC motor losses to be a quadratic function of the motor currents, yielding

$$P_{mechanical,F_f} = F_f \cdot f \cdot n = \left(U - R \cdot I_{FM} - \frac{P_{loss,other,FM}}{I_{FM}} \right) \cdot I_{FM} \quad (2)$$

as the decisive equation for feed force prediction and

$$P_{mechanical,M_C} = M_c \cdot 2 \cdot \pi \cdot n = \left(U - R \cdot I_{SM} - \frac{P_{loss,other,SM}}{I_{SM}} \right) \cdot I_{SM} \quad (3)$$

for the cutting torque, assuming that both powers are predominantly proportional to the respective motor current using an unknown transmission function. In the end, this relationship is predicted by LR, ANN and DT. Apparently, tool feed f and spindle speed n are influential and should be considered as features. Consulting Fig. 2, it is evident that F_f is dictated primarily by I_{FM} and f, whereas M_C depends mainly on I_{SM}, matching the assumption of Eq. (2) and (3). Nevertheless, the best later predictions of F_f and M_C used both I_{FM} and I_{SM}, indicating a relationship as the motors collectively act on a single planetary gear, resulting in potential mutual interactions.

Data with $n = 800\,rpm$ are not plotted in Fig. 2 and neglected in later model training due to amplitude modulation of the measured spindle current signals occurring at this speed, leading to a non-linear relationship between I_{SM} and M_C. The cause of this effect is unclear and requires further investigations but is suspected in signal digitization at the low sampling rate of the internal ADU sensors without hardware filtering.

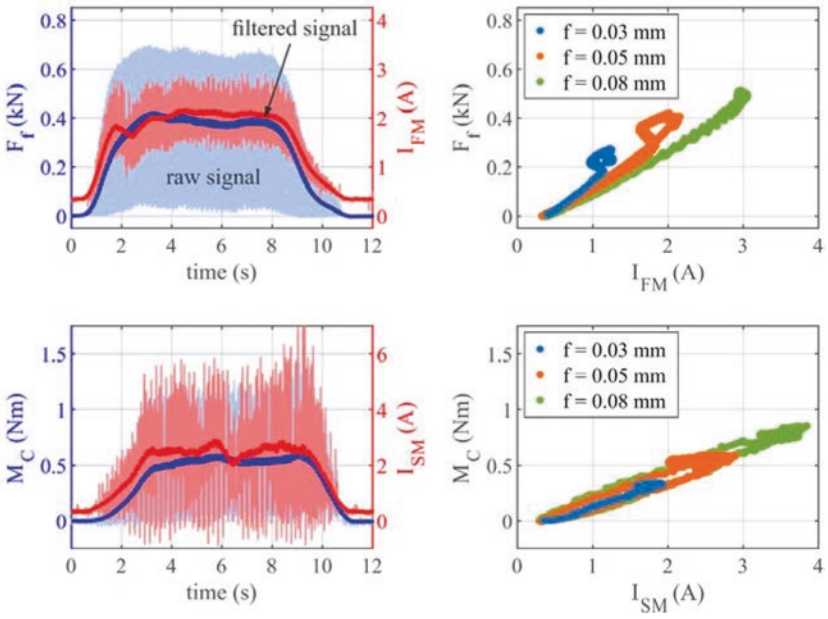


Fig. 2. F_f and M_C over time with corresponding motor currents for feed of f = 0.05 mm (left); filtered F_f and M_C over corresponding filtered motor currents at different feed rates (right)

Therefore, smoothed current signals, feed rates as well as machining time are available as features for the force prediction. Here, consideration of the normalized machining time t, reaching from 0…1 for every drill hole, enables to distinguish between different drilling situations (tool entry, full cut, tool exit). In the case of feed, the inverse value is used, corresponding to Eq. (2). Different combinations of I_{FM}, I_{SM}, $1/f$ and t as well as quadratic and interaction terms (e.g. I_{FM}^2 or $I_{FM} \cdot I_{SM}$) were tested for the ML methods of LR, ANN and DT. In addition, method specific settings were systematically varied. These were neuron and layer amount for the ANN, the type of terms in linear (and quadratic) regression and minimum leaf size of the DT. The corresponding models were implemented in MATLAB using functions of the *Statistics and Machine Learning Toolbox* (e.g. "fitlm", "fitrtree") and *Deep Learning Toolbox* (e.g. "trainNetwork").

After applying the training set to different model settings and feature sets, the minimum root-mean-square error (RMSE) of the cross-validation set was used to choose the final model configurations. Performance was then confirmed with the test set. The evaluation was carried out with data from the first experiment set in Table 1.

A performance analysis of the three ML methods trained with data from exp. 1 (titanium, $f = 0.03, 0.05, 0.08$, smoothed by moving average) is shown in Fig. 3. Solid lines show the real and predicted values of F_f and M_C for a drilling process in titanium from the test set. Dashed lines represent real and predicted values of a drilling process in aluminium, using the models trained for titanium. The application to aluminium data was done to underline the scope of different types of models depending on their training set. RMSEs were calculated after the application of the trained models to training, CV and test set data in titanium (exp. 1) and for the full data set in aluminium (exp. 4).

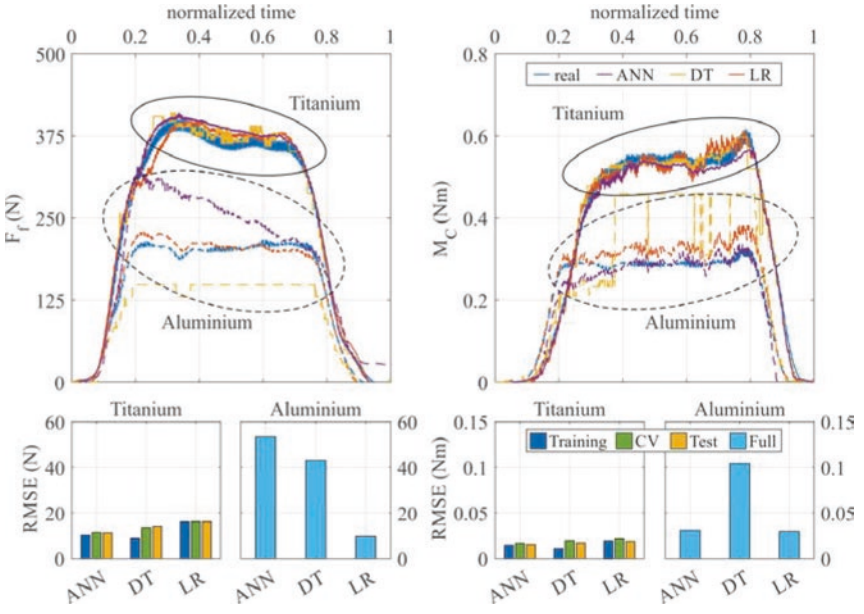


Fig. 3. Performance of ML methods in titanium and aluminium trained with exp. data set 1 (containing only titanium); top: predictions of F_f and M_C for feed rate of $f = 0.05$ mm; bottom: RMSE of the complete set of exp. 1 (titanium) and exp. 4 (aluminium)

The LR method, which is the basic ML approach, transparently shows the relationship between features and output. It revealed that the consideration of quadratic and interaction terms of I_{FM} and I_{SM} in addition to standard linear terms significantly increases the model quality. Squared motor currents have negative scaling factors θ_i for F_f and for M_C after training which corresponds to the theoretical considerations on the losses in the DC motors in Eq. (1). In contrast, quadratic and interaction terms without one of the motor currents increase model accuracy only slightly. The final LR model with features X_i chosen for the training is

$$Y_{LR} = \sum \theta_i X_i;\ X = \left\{ 1, I_{FM}, I_{SM}, I_{FM} \cdot I_{SM}, I_{FM}^2, I_{SM}^2, I_{FM} \cdot t, I_{SM} \cdot t, \frac{I_{FM}}{f}, \frac{I_{SM}}{f} \right\} \quad (4)$$

In the application of ANNs, using all possible 14 input features, thus considering all four essential features (I_{FM}, I_{SM}, $1/f$ and t) as well as their interaction and squared terms, has proven to be favourable. Different architectures with one and two hidden layers as well as varying numbers of neurons were systematically tested and evaluated by lowest cross-validation set RMSE. Here, large networks proved to be prone to overfitting. In the end, a network architecture with one hidden layer and a neuron number of [1, 5, 14] for F_f and [1, 7, 14] for M_C was chosen. This confirms statements by [14]. The properties of the ANN are shown in Table 2.

Table 2. Properties of the ANNs for prediction of F_f and M_C

Property	Value
Maximum number of epochs	3000
Activation function	leakyReLu
Initial learning rate	0.1
Learning rate drop per 10 epochs	0.99
Input features	I_{FM}, I_{SM}, $\frac{1}{f}$, t, $I_{FM} \cdot I_{SM}$, I_{FM}^2, I_{SM}^2, $I_{FM} \cdot t$, $I_{SM} \cdot t$, $\frac{I_{FM}}{f}$, $\frac{I_{SM}}{f}$, $\frac{1}{f^2}$, t^2, $\frac{t}{f}$,

While DTs provide the best representation of the training data sets, the prediction accuracy deteriorates when applied to unknown data, i.e. the test set. This indicates overfitting, a typical behaviour of DTs [15]. By changing the characteristic DT property "minimum leaf size" to 50, the tree depth could be limited, reducing overfitting. Nevertheless, LR and ANNs behave more uniformly than DTs which is reflected by similar RMSEs among training, CV and test data sets as well as continuous prediction behaviour in the upper plots of Fig. 3 compared to the DT predictions.

The good performance of LR for aluminium and the only marginally worse performance in titanium compared to ANN and DT indicates that the chosen LR configuration can describe physical effects of Eq. (2) and (3) adequately. In contrast to LR, there is no direct interpretability of the model for ANNs and DTs.

On the other hand, ANNs and DTs performed better than LRs for the experiments with $n = 800\,\text{rpm}$, where amplitude modulations appeared, leading to discontinuities in the internally measured current signals. LR, which cannot map nonlinear behaviour, shows a better performance with stronger smoothing of the distorted motor

current signals by Butterworth filtering compared to moving average smoothing. DTs and ANNs, which better capture nonlinearities, showed almost no performance differences between the two filter types.

In summary, the use of an LR model with quadratic, linear and interaction terms is recommended for this use case. For non-linear signal behaviour, the use of ANNs is recommended for a reliable prediction of the forces. This is particularly important if further characteristics or specific values for automatic process monitoring (e.g. tool breakage detection) shall be derived from the forces. DTs are less suitable since they are prone to overfitting and erratic prediction while being limited in the output to the maximum values of the training response.

3.2 Process Monitoring of Feed Rate, Lubrication State and Material

Feed rate, lubrication, and material have a major influence on drilling process, quality and tool wear. Thus, identification of these parameters is crucial in process monitoring, for which k-nearest neighbour (KNN) was applied to classify discrete states. KNN is a supervised classification method which assigns a given new data point with its features to a majority class of the nearest number (="k") of already classified data points in the neighbourhood [15]. The model was also implemented in MATLAB using functions of the *Statistics and Machine Learning Toolbox* (e.g. "fitcknn"). Both motor currents (I_{FM} and I_{SM}), machining time t and spindle speed n were used as model features. For feed prediction, I_{FM} was left out as to show prediction for ADUs without separate feed motor is possible. Spindle speed was included as the complete training set of Table 1 was used. Feature scaling in terms of mean normalization was applied to the input data in order to gain a balanced feature influence.

As model specific parameter, the Euclidean distance metric is used. The number of neighbours was varied in the range $K = 1 \ldots 20$, showing no significant impact on prediction quality. A higher number of nearest neighbours for the classification is expected to disturb the prediction through a higher probability of including outliers from distant classes. Class estimation (MATLAB "predict" function) for new data on a trained model was almost instant, demonstrating the applicability of this algorithm for real-time monitoring in a production line.

Feed Rate Prediction. The motor of industrially common single-engine ADUs acts on an exchangeable gear box which drives feed movement as well as spindle revolution. With the objective to conclude the value of acting feed of this gear box, a KNN model was established using I_{SM} and omitting I_{FM}. As most of the motor power is used for spindle revolution, the motor current of single-engine ADUs can be compared to I_{SM} of SmartADU (this is reflected by the power rating of SmartADU's two motors, see Sect. 2). Surprisingly, the prediction is precise, although the ratio of cutting and feed movement powers changes along the feed path, mainly in the entrance and exit phase, see also Fig. 2. Identification of the momentary feed offers the possibility to monitor if the correct feed gear from the work plan is in use. It was observed that the feed rate can be predicted with a maximum error rate of about 14%, see Fig. 4.

Lubrication Monitoring. Lubrication is vital for certain materials regarding tool lifetime and quality of the produced parts. Drilling processes with a machine lubrication setting of 0 and 1% (using only pressurized air) in Fig. 4 were defined as "Lubrication Off", the others in exp. 4 as "Lubrication On". The KNN modelling shows an error rate between 4 and 6%, having I_{FM} as an additional feature compared to feed prediction.

Material. By recognizing the stack composition and comparing it with e.g. the CAD-data of the current, position-tracked borehole, a reliable quality assurance is established. Furthermore, the recognition of the material can be used to adapt the cutting parameters to enable a material-dependent machining of the present stack. The material is recognized most accurately with error rates between about 4 and 5%, see Fig. 4.

In conclusion, KNN modelling offers a reliable prediction of feed, lubrication state and material composition of the underlying stack if statistically appropriate numbers of training examples are available. Clearly, prediction errors for feed in Fig. 4 do not primarily appear during the full cutting period. A more specific selection of test data would improve the classification result. To improve model performance, a study on the number of neighbours could be helpful, giving a hint for the ideal data and sampling reduction ratio. An analysis of misclassification locations could improve the model further.

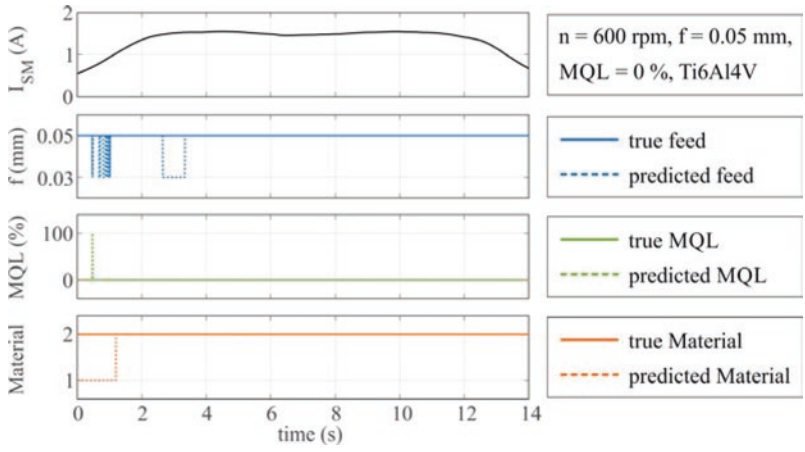


Fig. 4 True and predicted states of feed, lubrication and material determined by KNN-algorithm for an exemplary drill hole in titanium (for visualization purposes, one of the less accurately predicted drill holes was used)

4 Summary

In this paper, different machine learning methods have been applied to drilling processes with an Advanced Drilling Unit featuring two electric motors. ANN, DT and LR were used to predict feed force and cutting torque. The training was based on internally recorded motor current signals from cutting experiments in titanium. LR

modelling showed general applicability by accurate predictions even in aluminium, though it was only trained with titanium data. ANNs were able to model nonlinearities and had the best RMSEs in the titanium test data. DTs did not deliver reliable results and are therefore less suitable in this case.

Process monitoring of feed, lubrication, and material was performed using the KNN method. Accurate predictions of the process states were possible for all properties, even though the feed motor current was deliberately neglected in the case of feed assessment.

The investigations thus provide a basis for future real-time process monitoring of drilling processes with ADUs in the aerospace industry without the use of expensive external measurement technology. Applied in intelligent, semi-automatic drilling processes, this promises a reduction in process times, an increase in drilling quality and optimum tool life exploitation.

Acknowledgements. This publication resulted from a project funded by the *German Federal Ministry for Economic Affairs and Energy* under funding code 20Q1522C.

References

1. Berger, R.: COVID-19 – how we will need to rethink the aerospace industry. https://www.rolandberger.com/en/Point-of-View/COVID-19-How-we-will-need-to-rethink-the-aerospace-industry.html (2020)
2. Airbus Group: Qualität – Besser Bohren. Airbus ONE, p. 28 (2015)
3. Dornfeld, D.A., DeVries, M.F.: Neural network sensor fusion for tool condition monitoring. CIRP Ann. **39**, 101–105 (1990)
4. Arnold, F., Hänel, A., Nestler, A., Brosius, A.: New approaches for the determination of specific values for process models in machining using artificial neural networks. Procedia Manuf. **11**, 1463–1470 (2017)
5. Kim, D.H., Kim, T., Wang, X., Kim, M., Quan, Y., Oh, J., Min, S.H., Kim, H., Bhandari, B., Yang, I., Ahn, S.H.: Smart machining process using machine learning: a review and perspective on machining industry. Int. J. Precis. Eng. Manuf. Technol. **5**, 555–568 (2018)
6. Preez, A., Oosthuizen, G.A.: Machine learning in cutting processes as enabler for smart sustainable manufacturing. Procedia Manuf. **33**, 810–817 (2019)
7. Neto, F.C., Gerônimo, T.M., Cruz, C.E.D., Aguiar, P.R., Bianchi, E.E.C.: Neural models for predicting hole diameters in drilling processes. Procedia CIRP **12**, 49–54 (2013)
8. Shaban, Y., Yacout, S., Balazinski, M., Meshreki, M., Attia, H.: Diagnosis of machining outcomes based on machine learning with Logical Analysis of Data. In: 5th International Conference on Industrial Engineering and Operations Management, pp. 1–8 (2015)
9. Kim, H.Y., Ahn, J.H.: Chip disposal state monitoring in drilling using neural network based spindle motor power sensing. Int. J. Mach. Tools Manuf. **42**(10), 1113–1119 (2002)
10. Caggiano, A., Rimpault, X., Teti, R., Balazinski, M., Chatelain, J.F., Nele, L.: Machine learning approach based on fractal analysis for optimal tool life exploitation in CFRP composite drilling for aeronautical assembly. CIRP Ann. **6**(1), 483–486 (2018)
11. Caggiano, A., Angelone, R., Napolitano, F., Nele, L., Teti, R.: Dimensionality reduction of sensorial features by principal component analysis for ANN machine learning in tool condition monitoring of CFRP drilling. Procedia CIRP **78**, 307–312 (2018)

12. Hintze, W., Loedding, H., Friedewald, A., Mehnen, J., Romanenko, D., Moeller, C., Brillinger, C., Sikorra, J. N.: Digital assistance systems for smart drilling units in aircraft structural assembly. In: 7th International Workshop on Aircraft Systems Technologies, pp. 255–266 (2019)
13. Suominen, H., Pahikkala, T., Salakoski, T.: Critical points in assessing learning performance via cross-validation. In: Proceedings of the 2nd International and Interdisciplinary Conference on Adaptive Knowledge Representation and Reasoning, pp. 9–22 (2008)
14. Masters, T.: Practical Neural Network Recipes in C++. Morgan Kaufmann, San Francisco (1993)
15. James, G., Witten, D., Hastie, T., Tibshirani, R.: An Introduction to Statistical Learning, 8th edn. Springer, New York (2013)

Sustainable Interaction of Human and Artificial Intelligence in Cyber Production Management Systems

P. Burggräf(✉), J. Wagner, and T. M. Saßmannshausen

Chair of International Production Engineering and Management, University of Siegen, Paul-Bonatz-Str. 9-11, 57076 Siegen, Germany
peter.burggraef@uni-siegen.de

Abstract. AI will increasingly take over complex cognitive tasks and support human thinking and thus change the system of production management over decades to a cyber production management system. It has to be considered that AI can behave proactively, unexpectedly and incomprehensibly for humans. Here the human factor trust is essential and even becomes more relevant to determine sustainable relationship between humans and AI.

This leads to the research question at the edge of production research: What does human trust in an AI assistant depend, on in production management decisions? To answer this question this article statistically examines a set of previously identified influencing factors on human trust. From these results an explanatory model is derived, which serves as a first design guideline for a socially sustainable human-AI interaction in production management.

Keywords: Cyber production management · Production management · Trust · Artificial intelligence · Machine learning

1 Introduction

Today's manufacturing companies are confronted with a number of challenges: Mastering increasing complexity, realizing shorter delivery times and reacting to dynamic changes [1, 2]. It becomes evident that current planning and control mechanisms (e.g. mathematical optimization and heuristics) cannot efficiently deal with these complex, dynamic and unpredictable conditions [1]. Here, data-driven approaches from the field of artificial intelligence (AI) provide valuable solutions [1]. Through advances in statistical pattern recognition and machine learning algorithms, AI has proven that it is superior to humans in special domains [3]. Some of the most prominent examples are the victory of AlphaGo over the world's best Go player [4] and the victory in heads-up, no-limit Texas Hold'em poker over professional players [5]. In real-world applications AI will take over knowledge work [3], so that human work develops towards creative, problem-solving and monitoring activities [6]. AI

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 508–517, 2021.
https://doi.org/10.1007/978-3-662-62138-7_51

will support production managers to cope with the increasing complexity and uncertainty. This requires cooperative decision-making mechanisms between artificial and human intelligence [3]. For this reason, current and expected improvements in production management towards a cyber production management are presented in this next chapter. Following this, factors that positively influence a sustainable human-AI interaction are examined and combined in an explanatory model. This serves as the first step towards the design of trustworthy human-AI cooperation in cyber production management systems.

2 Cyber Production Management Systems

The first approaches of AI in production management have been showing their usefulness since 1980 [7]. In particular, these include methods of operations research, such as fuzzy logic, genetic algorithms or knowledge-based systems [7]. But as described, AI will increasingly take over complex cognitive tasks and support decision making [3, 8]. Thus, AI forms the counterpart to robotics in manufacturing companies: while robots facilitate the making of blue-collar workers, AI will support the decision making of white-collar workers [3]. The development of AI support for human production managers can be described in five stages, from sole human management via assisted and automatic management to fully autonomous decision-making by AI (see Fig. 1) [9].

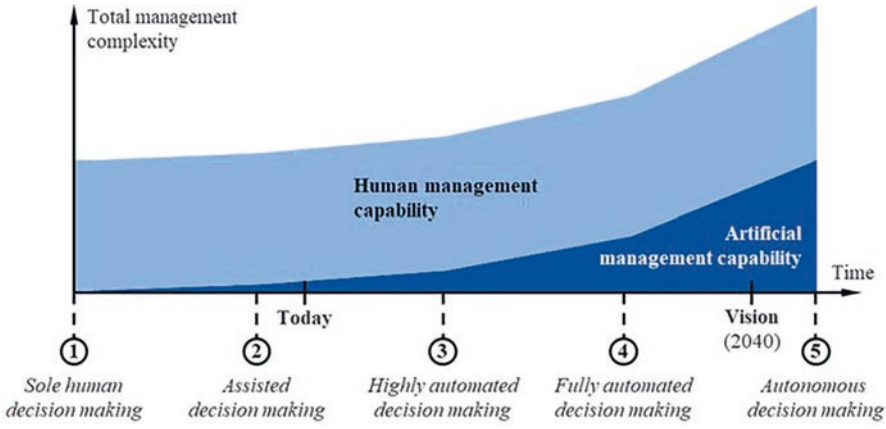


Fig. 1. Estimated development of decision making in production management [9]

Today's promising applications of AI in production management vary between assisted and automated decision making. These applications are mainly quality-based process control, predictive maintenance and order scheduling/dispatching [2, 9]. First, in *quality-based process control* AI is able to reveal and learn dependencies between parameter combinations in process control and the resulting quality of the product, which are unknown today [10]. Thus, AI can replace classic control systems. Second,

in *predictive maintenance* AI is capable of predicting machine failure based on the current machine status and the machine history and thus reduce downtimes. In the long run predictive maintenance and further developments like prescriptive maintenance will even change the business model of the machine manufacturers [11]. Last, AI algorithms increasingly are considered for *scheduling/order dispatching*. In this area of application, Reinforcement Learning algorithms receive a lot of attention. They compete against algorithms of operations research [12], against heuristics of engineering science (e.g. first in first out, earliest due date) [1], against other, modified versions of themselves [13], and against humans [14]. Special characteristics of these algorithms are that they react as fast as heuristics, achieve a better solution quality than these, but not as good as the slower optimization methods.

It turns out that AI currently does not go beyond applications for special tasks of operational production planning and control. However, AI will play an important role at higher decision levels in production management in the future [9]. This offers the opportunity of relieving human production managers of complex routine tasks and decisions. In addition, autonomous colleagues have advantage of objectifying decisions which today are too often based on subjective opinions [3]. The long-term estimation is that AI will change the system of production management over decades to a cyber production management (CPM) system where artificial and human intelligence interact and cooperate successfully [9]. While this autonomous decision making at all levels of the hierarchy is considered inevitable, no one dares to make concrete estimates of when would this happen [15]. Therefore, an estimation based on the costs per hour is to be attempted at this point: While a manager of a German medium-sized company costs on average 100 euro per hour [16], todays costs of an equivalent computing power are far higher. Following Moore's law – which is a rather conservative assumption given the current development of quantum computers with double exponential growth (Neven's law) [17] – studies expect computing power on a human cognitive level with costs of 100 euro per hour in seven to fourteen years [18, 19]. The expectation is that it will take a few more years of implementation and deployment before the first AI production managers can be found, and thus start the era of CPM systems. Similar to shop floor automation [20], CPM requires a dynamic, capability-based division of tasks. Thus, human and AI work together to combine their respective strengths and cooperate as hybrid intelligence [6]. The particular strength of the human is, as described by the cyberneticist Heinz von Foerster, to decide questions that are undecidable [21].

Based on systems theory, CPM systems are regarded as a combination of social and technical subsystems. In socio-technical systems the subsystems influence each other and lead to a synergetic and common optimization [22]. An aspect that socio-technical approaches and models of production management currently do not address is the equivalence of social (human) and technical (AI) subsystems. This is supported by expert interviews, which show that AI is currently regarded as a method in engineering [23] and that production managers refrain from leaving decisions to AI unless a human being could be made accountable [24]. Independent artificial agents, which can autonomously take over decisions in primary tasks of production management and independently improve their capabilities, are not considered, today. Here

CPM systems introduce a new paradigm by incorporating the growing dependence on information technologies and data volumes that go beyond human perceptual capabilities. CPM is a new form of production management between human and AI, which is intended to ensure future business success. To ensure this the according designing of CPM systems and thus human-AI interactions is required. For the design of CPM systems, three models have to be considered (see Fig. 2).

Fig. 2. Required models for the design of cyber production management systems

First, given the fact that intelligent systems operate better on a certain domain, but are with little to no use outside of it [3], the *Performance Model* aims to identify which production management tasks can be performed by either human or AI, and which tasks collaboratively. This model has to be dynamic over time, because an underperformance of AI in one field may change drastically with more computing power. Second, the *Trust Model* will describe and shape how production managers perceive AI-made decisions. The aim is to explain, predict and thus calibrate appropriate human trust. Trust serves as prerequisite for successful cooperation and interaction [25]. While trust in human interaction is well studied, the factors influencing human-AI trust differ and are subject of current research [26] that has to be examined in the light of production management [24]. Third, at some point autonomous decision-making will be too complex, fast-paced and impervious for humans. Here the *Control Model's* aim is to ensure decision accuracy, safety, security, ethics and human intervention. Today these issues are topic of various different initiatives for AI principles around the world [27]. These principles have to be adopted to fit the requirements of CPM systems.

In the interplay of these three models, the *Control Model* has an overarching role, while the *Performance* and *Trust Model* build opposites to each other. The *Performance Model* asks the question "What *can* AI (not) take over?" and the *Trust Model* asks "What *should* AI (not) take over?". This question of *can* versus *should* raises in times of robotics [28] and is today on top of the agenda of the European Commission, which proclaims human-centred AI [29]. Consequently, the *Trust Model* will be examined in detail below, to move a step towards sustainable human-AI interaction in CPM.

3 Research Field and Method

It is essential to underline that AI is understood as a social unit and teammate (instead of a tool) and that AI acts differently to previous systems: autonomously (model-based) instead of automatically (rule-based) [26]. This means that AI can behave proactively, unexpectedly and incomprehensibly for humans [24]. Here, the

research focus for sustainable interaction and cooperation lies on the human trust [24, 25, 29, 30]. There are various definitions of trust, depending on the disciplines. However, a common essence of trust can be identified: positive expectations and vulnerability [31]. In information systems (IS) research most studies on trust refer to the theoretical foundation of interpersonal trust [32], which is continuously adapted to IS research needs [33, 34]. The reason for that shift is that due to the rise of smart algorithms, technology is no longer seen as mediators between humans, but as trustees in a trust relationship [35]. Previous research on trust relationship, e.g. in military robots [36], in IT artefacts [33], in medical assistance systems [37], and in autonomous vehicles [26], points out that there are various antecedents which belong to the trustor, the trustee, the context and the interaction [34]. The major goal is the appropriate calibration of trust [26], to avoid over-trust which leads to misuse as well as under-trust which leads to disuse [38, 39]. This calibration is based on factors increasing and other factors decreasing trust [34].

To identify these influencing factors in the combined field of production management, AI and trust, we already conducted, first, an interview study, and second, an online survey, see [24]. In the interview study we interviewed four experts for production management from industry and science. We asked the experts for the potential factors influencing trust in AI, following our interview guide. The interviews were transcribed and then analysed using the qualitative content analysis [40]. As result of this analysis we discovered six potential factors. These six factors were divided into three groups (see Fig. 3), according to similar categories from literature [41–43]. In order to test whether these factors really influence trust, an online survey was conducted. Therefore, we developed a questionnaire where each factor is measured by different items (questions) and 5-stage Likert scales. Over a period of two weeks we gathered 130 completed questionnaires. On the basis of this collected data (online available on *Mendeley Data*, see [44]), we validated the measurement method and then tested the individual influences of the six factors one by one [24]. All individual influences were statistically significant (p-value < 5%) [24], but the combined influence and thus the joint explanation of trust remained unclear. This gap is closed by the present article. Based on the collected data of the online survey [44] a data-driven approach is pursued. For this purpose, multiple linear regression models are tested, consisting of all six main effects and all associated first-order interaction effects.

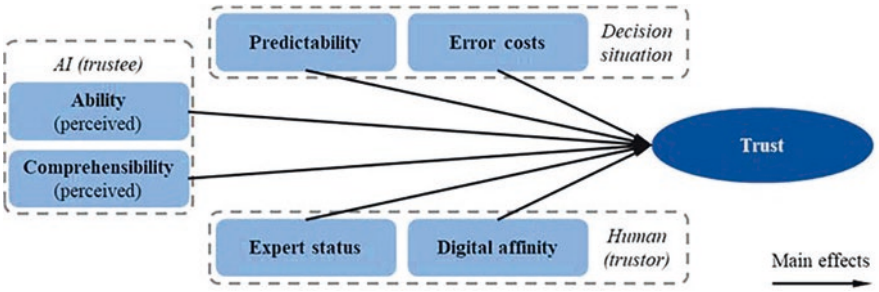


Fig. 3. Overview of the considered influencing factors on trust in AI [24]

In total there are 21 effects and thus 2^{21} theoretically possible multiple linear regression models, due to different combinatorics of these effects with each other.

To select the best model a set of evaluation criteria is used. For the model fit the Akaike information criterion (AIC) and Bayes information criterion (BIC) are reported (the smaller the better) [45]. As evaluation of multi-collinearity, the variance inflation factor (VIF) is used (should be smaller than ten) [46]. To report the determination of the models the adjusted R-squared (adj. R^2) is calculated (the higher the better) [47]. Finally, the p-value indicates the significance of each effect (should be smaller than 5%). The selection of the best regression model leads to the explanation of the factors impact on trust and thus can be used for the future design of CPM systems.

4 Results

To execute multiple linear regression, the most common and central technique is the ordinary least squares (OLS) method [48]. However, OLS estimators are very sensitive to violation of fundamental assumptions [48, 49]. Since the assumptions are not fully met,[1] robust regression with MM estimators is used. It shows the best performance in comparisons of different robust regression methods [49]. As tool, the free software R [50] is used, which is best suited for robust regression [48].

From the 2^{21} theoretical possible regression models there are 40,070 practical reasonable models, due to the hierarchy constraints of interaction effects (if an interaction effect is included in a model, the corresponding main effects also have to be included) [51]. All these models are evaluated with the mentioned criteria. Table 1 exemplarily shows the results for the best five models, according to AIC and BIC. They are the only models that are among the best ten models according to AIC and simultaneously among the best ten models according to BIC. The joint consideration of AIC and BIC values avoids the selection of an overfitted (low AIC) or underfitted (low BIC) model [52].

Table 1. Extract of the evaluation results for the best five models

Model no	Main effects (significant ones)	Interactions (significant ones)	AIC	BIC	Max. VIF	Adj. R^2
12	3 (3)	- (-)	957.8	979.1	1.32	0.408
28	4 (3)	- (-)	958.8	984.3	1.33	0.410
268	3 (2)	1 (0)	958.4	984.0	1.35	0.407
4108	3 (3)	1 (0)	957.5	983.0	1.35	0.410
4124	4 (3)	1 (0)	958.1	987.9	1.36	0.412

[1] First, the dependent variable *Trust* is not normally distributed, according to Shapiro–Wilk's test (p-value $< 2.2e{-}16$, should be $> 5\%$) and optical inspection of density and QQ plot. Second, homoscedasticity is just met as per Fligner-Killeen test ($p = 0.8$, should be $> 5\%$). Third, the linearity is proofed by the correlation matrix (using Spearman Rho), where all main effects show significant correlations (p-value $< 5\%$).

While the multicollinearity is no problem (VIF< 10) for all models, four of the five best models (no. 28, 268, 4108, 4124) face issues of non-significant effects, see Table 1, column two and three (p-value> 5%). The most promising model is model no. 12. It consists of only three main effects (Eq. 1, with a scaling of the factors from 0 to 5) and will be examined below.

$$trust = 0.589 + 0.438 * ability + 0.134 * comprehensibility + 0.156 * digital\ affinity \quad (1)$$

5 Discussion

All good models have one thing in common: Neither the factor *predictability* nor the factor *error costs* (both from the group *decision situation characteristics*) seem to make a decisive contribution to the explanation of trust, on the basis of the data collected. Even in the moderately good models in which they are included, their effects are very small and not significant. In contrast, all good models contain the three factors *ability*, *comprehensibility* and *digital affinity*, like the best model no. 12.

Considering the group *AI (trustee) characteristics*, both factors contribute to the explanation of trust. Here *ability* is already known in IS literature as a traditional and strong influencing factor [32]. The regression coefficients of model no. 12 (Eq. 1) show that *ability* has the strongest effect, by increasing it one unit, *trust* increases by 0.438 units. The factor *comprehensibility* is a novel trust antecedent, that has become increasingly important since humans interact with intelligent and autonomous systems [53], as the efforts of governments and companies to explainable AI underline [27]. The effect on *trust* is not as large as the one of *ability*, as the regression coefficient of 0.134 shows (Eq. 1).

In the group of *human (trustor) characteristics* there is also one known and one new factor. The known factor is the *expert status*, also named competence [41–43]. In this study, it is only part of a few good models (no. 28, 4124), but there it has only a small and not significant effect. For this reason, the contribution of the *expert status* to the explanation of trust can be regarded as very small, up to negligible, on the basis of the analyzed data. This is different with the *digital affinity*. It describes the degree to which a person has talent, interest and pleasure in learning and using (the latest) digital methods and technologies. Yet, it has not been considered in IS trust research. However, it is included here in all good models and shows a moderate as well as statistically significant effect (Eq. 1). Thus, it is the most surprising factor of this study.

6 Conclusion

This article examines which factors affect trust in an AI and how they affect it, on the basis of empirically collected data. Using the data-driven modelling, model no. 12 (Eq. 1) has proven to be the best and most useful model to explain trust in AI in the field of production management. This model is a first step towards the *Trust Model*

within the CPM framework. Here, the factor of *digital affinity* stands out, as it has not yet been included in any set for trust factors (see e.g. set for robots [43] or for automation [42]). However, this model currently explains about 40% of the trust variance (Table 1, adj. R^2). Thus, there is a need for further research to improve the model. For example, further influencing factors, that were not considered in this study, could be added, like the visual representation of the AI [41].

Summing up, the guiding question "What does human trust in an AI assistant depend on in production management decisions?" is partly answered. The results will serve as first design guidelines for a socially sustainable human-AI interaction in production management. For future cyber production management systems with autonomous and independent decision making by AI, the inclusion of further factors and the derivation of specific guidelines are necessary to ensure human centred design.

References

1. Kuhnle, A., Röhrig, N., Lanza, G.: Autonomous order dispatching in the semiconductor industry using reinforcement learning. Procedia CIRP **79**, 391–396 (2019)
2. Dannapfel, M., et al.: Human machine cooperation in smart production: evaluation of the organizational readiness. Int. J. Mech. Eng. Robot. Res. **8**(2), 327–332 (2019)
3. McAfee, A., Brynjolfsson, E.: Machine, Platform. Crowd. W.W. Norton & Company, New York (2017)
4. Silver, D., et al.: Mastering the game of go with deep neural networks and tree search. Nature **529**(7587), 484–489 (2016)
5. Moravčík, M., et al.: DeepStack: expert-level artificial intelligence in heads-up no-limit poker. Science **356**(6337), 508–513 (2017)
6. Dellermann, D., et al.: Hybrid intelligence. Bus. Inform. Syst. Eng. **61**(5), 637–643 (2019)
7. Russell, S.J., Norvig, P.: Artificial Intelligence, 3rd edn. Pearson Education, London (2016)
8. World Economic Forum: Deep Shift: Technology Tipping Points and Societal Impact. World Economic Forum, Geneva (2015)
9. Burggräf, P., Wagner, J., Koke, B.: Artificial intelligence in production management. In: International Conference on Information Management and Processing (ICIMP), pp. 82–88. IEEE, London (2018)
10. Kuhnle, A., et al.: Intelligente Produktionsplanung und -steuerung. wt Werkstattstechnik online **107**(9), 625–629 (2017)
11. Burggräf, P., et al.: Adaptive remanufacturing for lifecycle optimization of connected production resources. Procedia CIRP **90**, 61–66 (2020)
12. Gabel, T., Riedmiller, M.: Adaptive reactive job-shop scheduling with reinforcement learning agents. Int. J. Inf. Technol. Intel. Comput. **24**(4), 14–18 (2008)
13. Qu, S., et al.: Optimized adaptive scheduling of a manufacturing process system with multi-skill workforce and multiple machine types. Procedia CIRP **57**, 55–60 (2016)
14. Burggräf, P., et al.: Performance assessment methodology for AI-supported decision-making in production management. Procedia CIRP (2020)
15. Thomas, C., Liang, G.: The rise of the machines. https://www.mckinsey.com/industries/semiconductors/our-insights/the-rise-of-the-machines-how-chinese-executives-think-about-developments-in-artificial-intelligence (2016). Accessed 30. Apr. 2020

16. Markt und Mittelstand: Das verdienen Geschäftsführer im Mittelstand. https://www.markt-undmittelstand.de/gehalt/das-verdienen-geschaeftsfuehrer-im-mittelstand-1255071/ (2017). Accessed 30. Apr. 2020
17. Hartnett, K.: A new law to describe quantum computing's rise?. https://www.quantamagazine.org/%20does-nevens-law-describe-quantum-computings-rise-20190618 (2019). Accessed 30. Apr. 2020
18. Hsu, J.: Estimate: Human brain 30 times faster than best supercomputers. https://spectrum.ieee.org/tech-talk/computing/networks/estimate-human-brain-30-times-faster-than-best-supercomputers (2015). Accessed 30. Apr. 2020
19. AI Impacts: Preliminary prices for human-level hardware. https://aiimpacts.org/preliminary-prices-for-human-level-hardware/#easy-footnote-bottom-12 (2015). Accessed 30. Apr. 2020
20. Ranz, F., Hummel, V., Sihn, W.: Capability-based task allocation in human-robot collaboration. Procedia Manuf. **9**, 182–189 (2017)
21. von Foerster, H.: KybernEthik. Merve-Verl., Berlin (1993)
22. Windischer, A.: Kooperatives Planen. Universität Zürich, Zürich (2003)
23. Platte, L., Schönefeld, K., Hess, F.: Künstliche Intelligenz in der Produktionstechnik – eine Kränkung des Ingenieurs? FERRUM **91**, 94–101 (2019)
24. Saßmannshausen, T.M.: Vertrauen in Entscheidungen künstlicher Intelligenz im Produktionsmanagement. Shaker, Düren (2019)
25. Hancock, P.A., Stowers, K.L., Kessler, T.T.: Can we trust autonomous systems? In: Ayaz, H., Dehais, F. (eds.) Neuroergonomics, p. 199. Academic Press, London (2019)
26. de Visser, E.J., Pak, R., Shaw, T.H.: From 'automation' to 'autonomy': the importance of trust repair in human-machine interaction. Ergonomics **61**(10), 1409–1427 (2018)
27. Fjeld, J., et al.: Principled artificial intelligence. https://cyber.harvard.edu/publication/2020/principled-ai (2020). Accessed 30. Apr. 2020
28. Hancock, P.A.: Imposing limits on autonomous systems. Ergonomics **60**(2), 284–291 (2017)
29. High-Level Expert Group on Artificial Intelligence: Ethics guidelines for trustworthy AI. https://ai.bsa.org/wp-content/uploads/2019/09/AIHLEG_EthicsGuidelinesforTrustworthyAI-ENpdf.pdf (2019). Accessed 9. Mai 2020
30. Hancock, P.A.: Automation: how much is too much? Ergonomics **57**(3), 449–454 (2013)
31. Rousseau, D.M., et al.: Not so different after all: a cross-discipline view of trust. AMR **23**(3), 393–404 (1998)
32. Mayer, R.C., Davis, J.H., Schoorman, F.D.: An integrative model of organizational trust. Acad. Manag. Rev. **20**(3), 709–734 (1995)
33. Söllner, M., Pavlou, P. A., Leimeister, J. M.: Understanding trust in IT artifacts. https://doi.org/10.2139/ssrn.2475382 (2013). Accessed 6 Juni 2020
34. Lee, J.D., See, K.A.: Trust in automation: designing for appropriate reliance. Hum. Factors **46**(1), 50–80 (2004)
35. Söllner, M., et al.: Understanding the formation of trust in IT artifacts. In: Proceedings of the International Conference on Information Systems (ICIS), pp. 1–18. Orlando Florida (2012)
36. Hancock, P.A., Billings, D.R., Schaefer, K.E.: Can you trust your robot? Ergon. Des. **19**(3), 24–29 (2011)
37. Hengstler, M., Enkel, E., Duelli, S.: Applied artificial intelligence and trust—the case of autonomous vehicles and medical assistance devices. Technol. Forecast. Soc. Chang. **105**, 105–120 (2016)
38. Parasuraman, R., Riley, V.: Humans and automation: use, misuse, disuse. Abuse. Hum. Factors **39**(2), 230–253 (1997)

39. Drnec, K., et al.: From trust in automation to decision neuroscience. Front. Hum. Neurosci. **10**, 290 (2016)
40. Mayring, P.: Qualitative Content Analysis: Theoretical Foundation, Basic Procedures and Soft-Ware Solution. Gesis, Klagenfurt (2014)
41. Siau, K., Wang, W.: Building trust in artificial intelligence, machine learning, and robotics. Cutter Bus. Technol. J. **31**(2), 47–53 (2018)
42. Schaefer, K.E., et al.: A meta-analysis of factors influencing the development of trust in automation. Hum. Factors **58**(3), 377–400 (2016)
43. Hancock, P.A., et al.: A meta-analysis of factors affecting trust in human-robot interaction. Hum. Factors **53**(5), 517–527 (2011)
44. Saßmannshausen, T. M.: Data from an online survey on initial trust in artificial intelligence for production management tasks. https://doi.org/10.17632/pg8hnttzkm.1 (2020). Accessed 12. Mai 2020
45. Schreiber, J.B., et al.: Reporting structural equation modeling and confirmatory factor analysis results. J. Edu. Res. **99**(6), 323–338 (2006)
46. Hair, J.F., et al.: Multivariate Data Analysis, 7th edn. Pearson, Harlow (2014)
47. Susanti, Y., et al.: M estimation, S estimation, and MM estimation in robust regression. Int. J. Pure Appl. Mathemat. **91**(3), 349–360 (2014)
48. Wilcox, R.R., Keselman, H.J.: Modern regression methods that can substantially increase power and provide a more accurate understanding of associations. Eur. J. Pers. **26**(3), 165–174 (2012)
49. Yu, C., Yao, W., Bai, X.: Robust linear regression. https://arxiv.org/pdf/1404.6274 (2014). Accessed 12. Mai 2020
50. R Core Team: R: A language and environment for statistical computing. https://www.R-project.org/. R Foundation for Statistical Computing, Vienna, Austria (2019)
51. Hao, N., Zhang, H.H.: A note on high dimensional linear regression with interactions. https://arxiv.org/pdf/1412.7138 (2015). Accessed 12. Mai 2020
52. Burnham, K.P., Anderson, D.R.: Multimodel inference. Sociol. Methods Res. **33**(2), 261–304 (2004)
53. Matthews, G., et al.: Individual differences in trust in autonomous robots. In: IEEE Transactions on Human-Machine System (early access), 1–11 (2019)

Autoconfiguration of a Vibration-Based Anomaly Detection System with Sparse a-priori Knowledge Using Autoencoder Networks

J. Hillenbrand(✉) and J. Fleischer

wbk Institute of Production Science, 76131 Karlsruhe, Germany
jonas.hillenbrand@kit.edu

Abstract. This paper presents a method for machine component supervision with little to none prior knowledge of the machine, operating conditions and wear behavior. A hybrid approach based on unsupervised learning methods, consisting of an autoencoder network and clustering, to identify machine states and possible failure preceding anomalies is proposed. In order to cope with information sparsity, the model parameters of the unsupervised methods are derived automatically based on data distribution and a physical motivation. The approach was validated on a dataset of artificially introduced bearing faults. The gained clustering results show a general usability of the approach for condition monitoring with vibration data.

Keywords: Unsupervised learning · Anomaly detection · Autoencoder

1 Introduction

"With the proliferation of massive amounts of unlabeled data, unsupervised learning algorithms […] have gained popularity among researchers and practitioners" [1], this popularity stems from the way unsupervised learning methods can be applied, where classical supervised learning techniques cannot be used. For example, if no dataset with ground truth in form of corresponding labels exists, supervised learning and its classification or regression capabilities cannot be applied. Problems where only the given input, but not the corresponding target values are given, are so called unsupervised problems. Goals of unsupervised learning are [2]: discover patterns or groups of data with similar structure (Clustering), determine the data distribution and reduce the dimensionality of the problem without relying on user input for labelling.

However, applying unsupervised learning also comes with challenges, as it has not seen much attention as the more prominent techniques in supervised learning [3]. A common challenge is the tuning of hyperparameters to configure the model to properly fit the dataset. Another obstacle is the lack of interpretation or credibility of the clustering result, because of missing ground truth. In the end, a human component

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 518–527, 2021.
https://doi.org/10.1007/978-3-662-62138-7_52

that evaluates the results is still compulsory. Amongst unsupervised learning techniques, autoencoder networks (AEN) are an emerging technique for anomaly and fault detection.

This paper aims at decreasing the time effort and expert knowledge required for the configuration of anomaly detection systems using an AEN and clustering approach. Therefore the functionality of AENs is introduced and important research works are reviewed, followed by the introduction and validation of the new approach.

1.1 Autoencoder Network for Feature Extraction

A rather new approach for solving the problem of proper feature extraction in the context of condition monitoring is the use of an AEN. Although, early implementations of AENs can be traced back until 1987 [4], its application in fault or anomaly detection systems shows currently increasing popularity amongst researchers. The advantage of an AEN is that it does not require special knowledge on signal analysis or feature design [5]. An AEN is a neural network, which is trained using unsupervised learning. The network consists of two parts, an encoder (denoted by E) and a decoder (denoted by D). Using backpropagation it learns a representation of lower dimensionality by minimizing the difference between input x_i and its reconstructed output $\widehat{x}_i$, see Eqs. (1–3). During training the targeted output samples $\widehat{x}_i$ are equivalent to the input samples x_i. Details on formula deduction and computation of AENs can be found in [6, 7]. The transfer functions and model weights are represented by h and W, biases by b and the cost function is J, including regularization parameters λ, β and Ω.

Instead of time consuming feature extraction and engineering, an AEN can be used to identify a compressed representation of the data used to train it [8]. Moreover, AENs deliver a feature based representation in an automated way and the error returned for prediction of a new input can be used to decide whether the data was well reconstructed and hence can be used for anomaly detection.

$$z_j = h^E(W_{ji}^E x_i + b_j^E) \tag{1}$$

$$\widehat{x}_i = h^D(W_{ij}^D z_j + b_i^D) \tag{2}$$

$$J = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - \widehat{x}_i\right)^2 + \lambda \cdot \Omega^{E,D} + \beta \cdot \Omega^{Sparsity} \tag{3}$$

1.2 Autoencoder for Anomaly Detection

The use of AENs is presented in several works for anomaly detection in the context of speech, computing and also condition monitoring. Most works apply AENs to

automate and optimize the feature extraction of raw data to accomplish dimensionality reduction: Marchi et al. proposed a denoising autoencoder for feature extraction, whose features are then fed to a Long Short-Term Memory (LSTM) recurrent neural network for acoustic novelty detection. A denoising AEN is applied in order to force the hidden layer to retrieve more robust features and avoid simply learning the identity [9]. The authors in [10] proposed a new architecture for clustering of time series data, named deep temporal clustering. In their proposal, features are automatically extracted by stacked AENs and then fed into a clustering layer [8, 11]. demonstrate condition monitoring for bearing and power plant faults with AENs and extreme learning machines for automated feature extraction and later fault classification. Other works in the context of monitoring bearing faults propose the application of AENs to increase the accuracy of classifiers used after automatic feature extraction by the AEN [12–15]. These approaches show a general applicability of AENs for anomaly detection. They acquire the necessary hyperparameters offline and then use their trained models to monitor real-time data. The approach performs equally well as current state of the art hand crafted feature extraction methods. An alternative approach by [16] in the context of high performance computing attempts using the reconstruction error of an AEN as sign for anomalous data. Rather than learning the anomalous states, the network learns a representation of the healthy state and considers increasing reconstruction errors as indicator for anomalies. This approach has not been applied in production science, especially for monitoring machine vibrations. In general, obstacles for the proliferation of AEN models are correct tuning of hyperparameters and the model setup during an initial offline phase, which are both complex and time consuming. Combining the ideas introduced in [10, 16], this paper delivers a new hybrid approach to configure an AEN for training and use clustering to distinguish between normal and abnormal data. Its application concentrates on vibration data, though it can be generalized for similar time series data.

2 Hybrid Approach with Autoencoder and Clustering

The approach in this paper consists of two unsupervised learning methods to solve the problem of condition monitoring in unknown operation conditions. Examples for such conditions could be fatigue damage on bearing grooves, breakage of ball screw raceways due to high cyclic loads or change of lubrication conditions in rolling element contacts. It evolves around the five steps depicted in Fig. 1, namely data streaming or acquisition (i), data separation into smaller windows (ii), then frequency content analysis (iii) for proper configuration of the AEN in step (iv) and finally the evaluation of the AEN reconstruction by a clustering method (v). These steps are iteratively conducted to supervise the ongoing machine operation. Neglecting the first step, as it depends on the used sensor system, the remaining general and transferable steps are discussed in more detail.

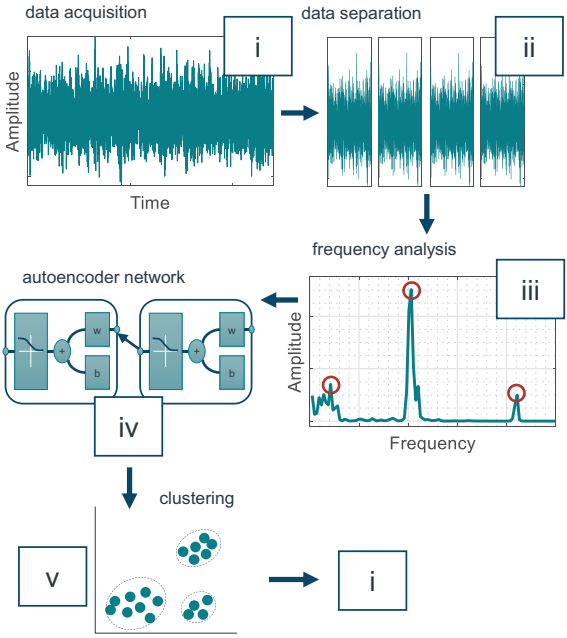


Fig. 1. Flow chart of anomaly detection system

2.1 Autoconfiguration of Autoencoder Using Frequency Domain

In order to simplify the usage of AENs, the configuration of hyperparameters can be further automated, foremost by automatic determination of neurons in the hidden layer and required input data size. The idea of the approach presented in this paper is based on Fourier theory, where each signal $f(t)$ can be described as an infinite series sum of sine and cosine waves, see Eq. (4).

$$f(t) = a_0 + \sum_{n=1}^{\infty} (c_n \cos(2\pi f_n t) + s_n \sin(2\pi f_n t)) \tag{4}$$

Especially, with vibration data the assumption that the original data consists of an infinite number of sinusoid waveforms holds truth, as the contemplated system can be abstracted and simplified as multi mass transducer system. The response of such systems is certainly affected by wear and changing operation conditions. In order to detect these changes the AEN is trained to learn the initial state of the system. Therefore, in the encoder stage of an AEN, the data will be compressed to a lower number of features, here N_f, than the input data size N. The approach assumes that N_f , representing the number of neurons in the hidden layer of the encoder stage, can be estimated by analyzing the frequency spectrum of the input data. The spectrum is analyzed for main frequencies in the spectrum using fast fourier transform (FFT).

Those main frequencies are identified automatically using the following routine (steps 1–5):

- Compute the FFT $\{f, p\}$ of x_i with $i = 1 \ldots N$, where $N = f_S/f_{res}$, f_s is the sampling rate and f_{res} the targeted frequency resolution (with default $1\,Hz$). f contains the discrete frequency bins and p the corresponding spectral amplitudes.
- Search $\{f, p\}$ for local maxima $\{\widehat{f}, \widehat{p}\}$.
- Neglect all maxima with $\widehat{p} < 2 \cdot std(p)$ and maxima should be at least $2f_{res}$ apart, where std is the standard deviation.
- This results in $\widehat{N}$ distinct main frequencies which will be used to reconstruct $f(t)$ as $f(t) \approx a_0 + \sum_{n=1}^{\widehat{N}} (c_n \cos(2\pi \widehat{f}_n t) + s_n \sin(2\pi \widehat{f}_n t))$.
- The reconstruction contains $N_{AE} = 1 + 3\widehat{N}$ unknown parameters $(a_0, c_n, s_n, \widehat{f}_n)$, which also represent the amount of neurons in the hidden layer of the encoder stage, allowing a compression rate of $c = \frac{N_{AE}}{N}$.

In order to demonstrate this process, it is conducted for a noisy sine wave:

$$f(t) = A \cdot \sin(2\pi f_{ex}(t + t_0)) + n(t) + o, \text{with } t \in [0, 1] \tag{5}$$

Where A, f_{ex}, t_0 and o are arbitrary values representing amplitude, frequency (here 33 Hz) and offset, whereas $n(t)$ adds white noise to the signal. Processed according to the steps 1 to 5, the spectrum is gained in step 1 and, due to formula (5), a single main frequency identified (using steps 2–4). Based on the identified frequency the number of hidden neurons in the encoder stage will be set to $N_{AE} = 4$ (step 5). After acquiring the number of hidden neurons an AEN can be trained in order to reconstruct the original data. As training data an artificial dataset consisting of four sets of 1000 arbitrary sine waves with different frequencies was generated (33 Hz, 250 Hz, 831 Hz and 2000 Hz). These sine waves varied in amplitude, phase, offset and noise level and for each frequency series a separate AEN was trained. MATLAB® is used for the AEN implementation and training using the Deep Learning Toolbox while adhering to Eq. (1–3) [6].

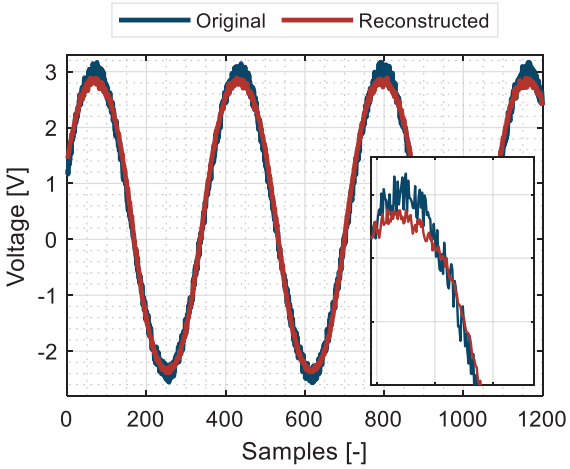


Fig. 2. Reconstructed sine wave with 33 Hz by AEN

The successful reconstruction of the original data can be found in Fig. 2. The network was able to learn the correct shape of the sine wave and compressed this information within the hidden layer comprising of four neurons. It can also be seen that this compression has a denoising effect.

2.2 Enhancing the Reconstruction Using Frequency Domain

The objective of an AEN is minimizing the difference between its input x_i and reconstructed output $\widehat{x}_i$. A measure used to quantify this error is the mean squared error (MSE) described in Eq. (5). It represents the deviation between reconstructed and original data over the signal length.

$$MSE = \frac{1}{N}\sum_{n=1}^{N}(x_i - \widehat{x}_i)^2 \tag{6}$$

$$x_i = \{x_{Time,j}, x_{Freq,k}\}, \text{ where } x_{Freq,k} = fft(x_{Time,j}) \tag{7}$$

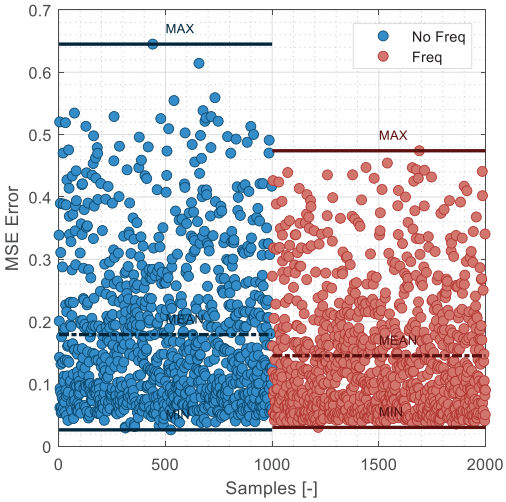


Fig. 3. MSE error for example sine dataset

If not specified otherwise, the data fed to the AEN is normalized and hence the MSE is as well. Therefore the MSE can be considered for comparison of how well the trained network performs comparing different signal windows. To further enable the AEN to learn the signal representation in terms of fourier series coefficients in addition to the time series data, the data is also represented in the frequency domain and added to the training input vector (see Eq. (7)). The usefulness of adding additional features to the input vector can be seen in Fig. 3. The figure shows the MSE distribution for all sine waves of a specific frequency comparing input vectors with (red) and without frequency information (blue). Maximum and mean error can be reduced

using additional features, whereas the minimum error for both input options stays unchanged due to the same settings for regularization to avoid overfitting.

In respect to other frequencies in the dataset, the comparison in Fig. 4. is gained. The mean MSE reduction achieved for the tested data lies between 19 and 30%. The minimum reconstruction error cannot be reduced, but using the additional frequency features achieves better generalization. Data augmentation by terms of increasing the number of features, here by adding the frequency spectrum, is a common measure to improve the training accuracy [17]. Including the frequency domain also proves beneficial in case of the AEN.

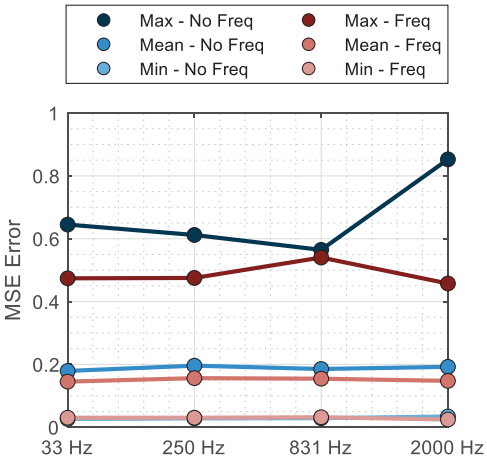


Fig. 4. MSE comparison for data with and without frequency features

2.3 Identification of Changing Operation Conditions with Clustering

Once an AEN is configured and trained according to the presented approach, it can be used to evaluate new incoming data. The only output used from the AEN is the MSE for each new signal window. During system supervision this approach delivers a tuple of time and MSE values $\{t_i, MSE_i\}$ for each new signal window. The AEN will predict a low reconstruction error for already seen data, whereas anomalous states can be identified by rising MSE values. In order to distinguish between normal and anomalous state a corresponding threshold is required [16]. Utilizing such a threshold introduces a new hyperparameter for the model, which requires human decision making to define it.

In contrast, a clustering algorithm instead of defining a threshold depending on existing knowledge of the supervised system is used in the context of this paper. This hybrid approach must deal with unseen data to start with and continuously new incoming data. Assuming no a-priori knowledge, varying numbers of clusters are expected, which are unknown at the time, when the clustering result is being computed. Among state of the art clustering algorithms this requirement reduces the number of adequate candidates, e.g. DBSCAN [18], MeanShift [19] or OPTICS [20]. In

a first attempt DBSCAN was applied, other clustering methods will be considered in future works. The remaining hyperparameters for adjusting DBSCAN are estimated automatically, as suggested in [21]. For further details on the DBSCAN algorithm see [18].

2.4 Description of Datasets with Changing Operation Conditions

In order to validate the performance of this hybrid approach, a dataset with artificial bearing faults provided by [22] was used. The bearing dataset consists of acceleration sensor data sampled with 12 kS/s at constant speed. Along with baseline data, representing the normal state, raceway damage was artificially introduced on the bearing groves of different extent to simulate fatigue. The AEN is only trained with a portion of the baseline data, before it is fed with more data with and without damage. The dataset was prepared with 1414 baseline signal windows, each containing 1200 samples, hence 100 ms of data with a possible frequency resolution of 10 Hz by FFT. Signals with damage sum up to 1213 windows of the same size and sampling rate.

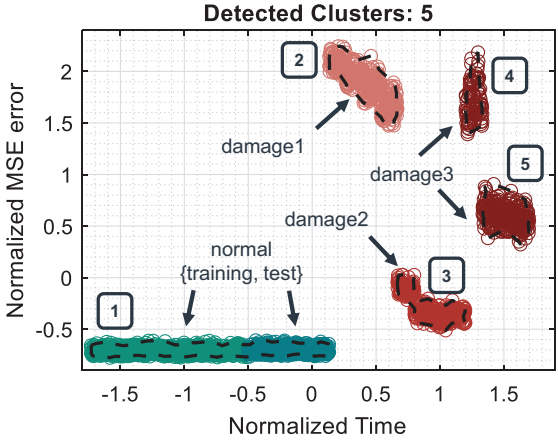


Fig. 5. Clustered result of normalized MSE plot for damaged bearings

2.5 Evaluation of New Approach

Using the bearing dataset, the baseline data was split into training and test data with ratio 2:1. Following the steps 1–5 in the introduced routine, the number of hidden layer neurons is estimated to 31. With this setting the training data is used to train an AEN with time and frequency domain. Figure 5 displays the resulting MSE plot for training, test (normal state) and three different damage classes. The ground truth is shown by different colors (including four distinct classes {normal, damage1, damage2, damage3}). Based on the experiments in [22] the damage classes range from faults of diameter 0.18 to 0.54 mm on the inner raceway of the bearings. Those damages were artificially applied using electro-discharge machining. The identified clusters are marked by the dashed black border and a number. The results obtained for

cluster 1 show that the AEN was trained well, as training and test data are assigned the same cluster. Furthermore, the MSE values allow precise separation of the damage classes. As soon as data of these classes is fed into the AEN, it assigns them to a new cluster. It can be observed, that the approach delivers five clusters, instead of four, but a visual interpretation of this data distribution, without knowledge of ground truth would also suggest five distinct clusters. As the AEN was only trained with samples from the normal state, he did not learn that the data distribution in cluster four and five actually belongs to the same class. The wrong assignment of datapoints in cluster four and five may arise from changes of the measurement data within the class {damage3} of the bearing dataset.

Throughout the whole configuration of the anomaly detection system described in Fig. 1 only one hyperparameter was set manually, the frequency resolution f_{res}. All other hyperparameters, including those of the AEN and the DBSCAN algorithm, have been estimated based on the data distribution in the training samples.

3 Conclusion

Discovering information and structure in uncharted, unlabeled data for anomaly detection purposes can be a tedious and complex task. In order to simplify this task, this paper proposed a hybrid approach, consisting of AEN and a clustering method, to detect anomalies in vibration data. In terms of reducing the effort for model design and network tuning, a novel autoconfiguration routine for AENs was introduced. It enables an adequate choice of hyperparameters, namely the number of hidden neurons, without requiring user experience. Moreover, as mentioned in other works, AENs acquire features autonomously without prior knowledge of the system. In terms of model interpretation, DBSCAN was used to cluster the reconstruction error output of the AEN. Hence, no threshold selection is necessary for distinguishing between normal and anomalous data, as the method provides this decision as part of the clustering result. This fact shall be used to identify anomalies in the incoming signal without explicit user interaction. Thereby, it enables supervision of machine conditions and recognition of changing machine states. So far, the approach lacks interpretability due to the unsupervised structure, it cannot tell why the anomaly was detected, but only provides the binary decision normal or abnormal. The interpretation of changing states or temporal development of clusters has not yet been addressed and is intended for future works.

Acknowledgement. This research work is funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project 388141462.

References

1. Celebi, M.E., Aydin, K.: Unsupervised Learning Algorithms. Springer, Cham (2016)
2. Bishop, C.M.: Pattern Recognition and Machine Learning Information Science and Statistics. Springer, New York (2009)

3. Seif, G.: An easy introduction to unsupervised learning with 4 basic techniques. https://towardsdatascience.com/an-easy-introduction-to-unsupervised-learning-with-4-basic-techniques-da7fbf0c3adf (2019). Accessed 15 May 2020
4. Ballard, D.H.: Modular learning in neural networks. In: AAAI (ed.) Sixth National Conference on Artificial Intelligence, vol. 1, pp. 279–284. Los Altos, California (1987)
5. Goodfellow, I., Bengio, Y., Courville, A.: Deep Learning. MIT Press, Cambridge (2016)
6. Mathworks Train Autoencoders: https://de.mathworks.com/help/deeplearning/ref/trainautoencoder.html. Accessed 18 Apr 2020
7. Ng, A., Ngiam, J., Foo, C.Y., et al. UFLDL Tutorial – Autoencoders: https://ufldl.stanford.edu/tutorial/unsupervised/Autoencoders/. Accessed 18 Apr 2020
8. Roy, M., Bose, S.K., Kar, B., et al.: A stacked autoencoder neural network based automated feature extraction method for anomaly detection in on-line condition moni-toring. In: IEEE (ed.) Symposium Series on Computational Intelligence (SSCI), pp. 1501–1507 (2018)
9. Marchi, E., Vesperini, F., Eyben, F., et al.: A novel approach for automatic acoustic novelty detection using a denoising autoencoder with bidirectional LSTM neural networks. In: IEEE (ed.) International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1996–2000 (2015)
10. Madiraju, N.S., Sadat, S.M., Fisher, D., et al.: Deep temporal clustering: fully unsu-pervised learning of time-domain features. https://arxiv.org/abs/1802.01059 (2018)
11. Michau, G., Fink, O., Hu, Y., et al.: Feature learning for fault detection in high dimen-sional condition monitoring signals. J. Risk Reliab. **234**(1), 104–115 (2019). 10.1177/1748006X19868335
12. Zhang, S., Ye, F., Wang, B., et al.: Semi-supervised learning of bearing anomaly detection via deep variational autoencoders. https://arxiv.org/pdf/1912.01096.pdf (2019)
13. Meng, Z., Zhan, X., Li, J., et al.: An enhancement denoising autoencoder for rolling bearing fault diagnosis. Measurement **130**, 448–454 (2018). 10.1016/j.measurement.2018.08.010
14. Ren, L., Sun, Y., Cui, J., et al.: Bearing remaining useful life prediction based on deep autoencoder and deep neural networks. J. Manuf. Syst. **48**, 71–77 (2018). 10.1016/j.jmsy.2018.04.008
15. Sohaib, M., Kim, J.M.: Reliable fault diagnosis of rotary machine bearings using a stacked sparse autoencoder-based deep neural network. Shock Vib. **2018**, 1–11 (2018). 10.1155/2018/2919637
16. Borghesi, A., Bartolini, A., Lombardi, M., et al.: Anomaly detection using autoencod-ers in high performance computing systems. In: AAAI Conference on Innovative Applications (2019)
17. Wen, Q., Sun, L., Song, X., et al.: Time series data augmentation for deep learning: a survey. https://arxiv.org/pdf/2002.12478v1(2020)
18. Ester, M., Kriegel, H-P, Sander, J., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise (1996)
19. Fukunaga, K., Hostetler, L.: The estimation of the gradient of a density function, with applications in pattern recognition. IEEE Trans. Inform. Theory **21**(1), 32–40 (1975). 10.1109/TIT.1975.1055330
20. Ankerst, M., Breunig, M.M., Kriegel, H.P., et al.: OPTICS: Ordering points to identify the clustering structure. SIGMOD Rec. **28**(2), 49–60 (1999). 10.1145/304181.304187
21. Mathworks: Estimate neighborhood threshold: DBSCAN. Phased array system toolbox. https://de.mathworks.com/help/phased/ref/clusterdbscan.clusterdbscan.estimateepsilon.html. (2020). Accessed 25 Apr 2020
22. Case Western Reserve University Bearing Data Center: https://csegroups.case.edu/bearingdatacenter/home. Accessed 20 Apr 2020

Combining Process Mining and Machine Learning for Lead Time Prediction in High Variance Processes

M. Welsing[1(✉)], J. Maetschke[1], K. Thomas[1], A. Gützlaff[1], G. Schuh[1], and S. Meusert[2]

[1] Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
M.Welsing@wzl.rwth-aachen.de
[2] MAN Truck & Bus SE, Vogelweiherstraße 33, 90441 Nürnberg, Germany

Abstract. Machine learning offers a high potential for the prediction of manufacturing lead times. In practical operations the lack of defined processes and high-quality input data are a major obstacle for the use of machine learning. The method of process mining creates a better transparency of such workflows and enriches related data. This paper develops a method, which combines the benefits of machine learning and process mining with the goal of high accuracy lead time prediction. The method is focused on high variance processes and verified with a case study containing real industrial data from heavy engine assembly processes.

Keywords: Lead time prediction · Process Mining · Machine Learning

1 Introduction

The manufacturing industry has been driven by mass customization, resulting in an increasing variation of product variants. Manufacturers have reacted to this by creating high model-mix production processes within existing production facilities [1]. However, it is self-evident that a variance in the product portfolio does lead to greater variety in the production processes. This process variation intensifies the difficulties of measuring and predicting the lead time for each process and the entire production order. Knowing the exact lead times is crucial for every manufacturer, since both production program planning and customer satisfaction are dependent on clocklike processes and delivery [2].

Fortunately, new technological developments offer solutions to master these new challenges. These new solutions are closely linked with usage of data, which is made available by digitalized production environments. From the field of data science both process mining (PM) and machine learning (ML) provide many tools to measure or predict lead times, if they are fed with the right data. If ML is utilized to predict the

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 528–537, 2021.
https://doi.org/10.1007/978-3-662-62138-7_53

lead times of production processes, it requires a complete and standardized dataset of manufacturing times for each process step. Obtaining these times or timestamps can be rather difficult in a multi-process environment, where every product variant leaves an almost unique trail of timestamps [3]. PM strength is to create better transparency in such a complex process landscape and could therefore help to filter a manufacturing timestamp dataset into a more uniform structure [4]. Doing so would create a better feature for ML algorithms to operate on and lead to a higher quality in lead time predictions. This paper aims to link the process insights obtained from PM with ML algorithms. Especially since the performance of a ML model is heavily dependent on the right input data, the information gained with PM can help to improve the model [5].

In the first part of this paper the fundamentals of ML, PM are explained and a short overview of the scientific literature of using these to predict lead time is given. The second part explains an approach on how to integrate the PM results into ML model, with the goal to create a higher accuracy in predicting lead times. Finally, the approach is validated on a data set from the heavy engine manufacturing industry.

2 Related Work

This section covers an overview of related work. A benchmark of relevant publications, that cover ML cases with lead time prediction is presented. This is followed by a comprehensive round up of the history of process mining and its application in manufacturing.

2.1 Lead Time Prediction with Machine Learning

The application of Machine Learning has risen in the past years across various industry sectors. ML algorithms aim to optimize a defined performance criterion by learning from existing data [6]. ML can be separated into two categories: supervised and unsupervised ML. Supervised ML uses existing labels to either classify or regress these labels. Unsupervised ML does not require labels since it aims to discover unknown patterns [7]. Yet the development of a ML model requires a combination of business understanding and knowledge about the right algorithms. For the domain of production and logistics a wide range of case studies have been investigated by the scientific community [8]. Lead time prediction itself had only limited observations in the research community [9]. The most relevant papers are discussed in the following section.

Lingitz et al. compare the performance of different ML algorithms for the prediction of lead times. However, they do not consider high variance processes and have no need for PM based process information [10]. The study presented by Meidan et al. also evaluates different ML algorithms, but solely considers waiting times [11]. Gyulai et al. focus on the characteristics of the product to predict lead times in a flow-shop environment, but ignore the complexity of the processes involved [12]. The works of both Pfeiffer and Ötzurk conduct the same effort of comparing the accuracy

of predicted lead times. It must be mentioned, that both authors employ only simulated data as an input for their models [13, 14]. Alenzi et al. show that support vector machines perform better than neural networks in the task of predicting order flow time. However merely simulated data and no real shop floor information was used [15].

The research of Mori et al. observes the probability distribution of operating conditions, but focuses mostly on the challenges of operating with binarized variables [16]. Schuh et al. showcase a framework as well as a study on real industry data on how ML algorithms can be used for the prediction of transition times. Although the case study is heavily process oriented PM is not used for feature enhancement [17, 18]. An interesting approach of enriching the ML input data with new features is presented by Windt et al., who apply methods adopted from the gene expression analysis [19].

2.2 Process Mining in Manufacturing

The modelling and visual illustration of business processes is a well-established research domain [20]. Process mining describes the automated modelling of such processes based on event data [21]. Although the first PM applications were used in health care and software development industries, the importance of PM for production and logistics has been recognized [20]. PM has been successfully deployed to different applications in the manufacturing industry.

While all authors prove the usefulness of PM, many do not use the data gained by PM to create further insights with ML [22–24]. Pospíšil et al. continue to analyze the PM data with ML algorithms but only a low variance manufacturing process was addressed [25]. The characteristics of high variance processes were mostly considered in the field of logistics [26]. Becker et al. combine the benefits of PM and ML in a logistics scenario. However, they use ML to enrich the data for PM and not vice versa [27]. Another approach on gaining insights on manufacturing processes with ML is presented by Wallis, who uses data mining techniques to enhance the planning efficiency of assembly tasks. Nevertheless, he does not use actual PM tools for his approach [28]. It can be concluded, that PM has not yet been utilized for feature enhancement in the context of ML models, which predict the lead time of manufacturing processes.

It can be stated that PM is an effective tool to visualize and model a manufacturing process. However, for deeper analysis, beyond the calculation of mean, minimum and maximum process times as well as simple drill-down and filtering functions, PM is often limited. Especially for cases with a high amount of features, the correlation within these features can be difficult to observe. ML algorithms can overcome this type of challenge. However, every ML algorithm needs a clear definition of variables on which it should optimize. In high variety processes, such variables are often not clearly defined and can only be obtained with the operation of PM. Particular the definition of sub-processes can help to enhance the results of ML-algorithms, since many features only have an effect in certain process steps. For these reasons the following section will present a method on how the advantages of PM and ML can be combined to predict lead times more precisely.

3 Method

The method of predicting lead times with the combined usage of PM and ML can be structured into 4 separate steps. Depending on the results of each steps, iterations are possible. An outline of the method can be obtained from Fig. 1. It is apparent how PM is utilized as an extra step to enrich the input data for the ML model. With the integration of a data understanding in the first step and the implementation in the last step the method adapts success factors from exiting approaches like the Cross Industry Standard Process for Data Mining (CRISP-DM) or the Knowledge Discovery in Databases (KDD) [29, 30]. However, the combination of the middle steps present a unique approach for the specific problem presented.

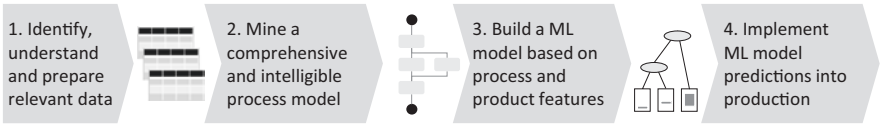


Fig. 1. Method for lead time prediction

The goal of the first step is to identify, understand and prepare all relevant data, which will be required for further analysis. In order to identify the relevant data sources all stakeholders should be interviewed. The full comprehension is crucial for hereinafter steps like the feature engineering needed for both PM and ML. The last task of the first step is the data preparation. For this, the relevant information of the different sources must be merged into bigger data frames. For the following PM the data should be transferred into an event log, which includes an order or case ID, an event name and timestamp. For a lead time prediction, the data should be structured in a way, that all features are matched with the relevant case or order id. Since some features are to be discovered within PM, the complete data preparation cannot be fully executed at this point.

The second part of the method should result in a comprehensive and intelligible process model. It is a common goal for PM to go from spaghetti to lasagna models [31]. This means that a detailed process model is often not understandable for humans and should be reduced to a simpler model, without sacrificing too much information about the process. Both heuristic and attribute based filters are helpful to perform this task on process information about manufacturing. The definition of a single start and end event, as well as the investigation of the most common events have proven to be highly effective. Additionally, it is impactful to group events together and eliminate process shortcuts. The resulting process models ought to be discussed iteratively with the domain experts. This ensures a higher quality of the final model and helps to keep all stakeholders involved. Also, this exchange is eminently suited to communicate first insights of the project. Additionally, the obtained PM model can be used to replay the existing event log and make first estimations about the lead times. These times can be used as a benchmark against the upcoming ML model results.

The defined sub-process steps and lead times for each step present a valuable input for the third part of this method. In this part a ML model will be built with features including the process and product information. Initially the right ML algorithms should be selected. This is very important since there is no optimal algorithm for all kinds of predictions [32]. The prepared and existing data should be investigated to decide on the right algorithm. For numeric lead time predictions all variations of regression algorithms are suitable. For non-numeric lead time prediction with labels like “too late” or “on time” classification algorithms are a proper fit. If the existing features are mostly numeric, support vector machines and naïve Bayes models should be considered. Otherwise for categorical features decision trees and their extensions like bagged, boosted and random forests are endorsed [33]. As soon as the right algorithm is selected, the next task of feature engineering begins. Although ML offers many algorithms to define the relevant features, the domain knowledge and business understanding is still crucial for this step. Therefore, this method proposes, that firstly the algorithms are used to detect the relevant features. After that, the non-relevant features are eliminated and set to ignore one by one. Based on this new data set the ML algorithm can be applied again.

In the last part of the method the lead time predictions are implemented into production. The relevant stakeholders should review all findings from the data and process analysis. This is important since the experience of the engineers and plant managers is needed to put the findings of the data analysis into perspective. When the model and the predicted lead times pass this final examination, the times can be used for future production planning and control. One option is to compute the prediction times once and enter them into the relevant Manufacturing Execution System (MES) or Production Planning and Control System (PPC). Since most default times in such systems are based on estimations made by experts, the utilization of ML-based times has the potential to improve the accuracy of those times. Additionally the ML-model can give valuable insights on what criteria and characteristics should be used to differentiate the predefined lead times. Another option is to integrate the whole ML-model into the planning system and compute the lead time for every new order. This option offers the additional potential, that a higher degree of differentiation becomes possible, since classical MES applications mostly work with a given structure for lead times, which is limited to certain differentiation features.

4 Validation

The presented method has been validated with a real industrial use case. The analyzed manufacturing process includes the assembly, testing and painting of heavy engines. The complete data set has a size of 4GB and covers the process records of one year. All critical information has been sanitized in order to protect the company’s interests. None of the mutations interfere with the interpretation or proof of the validity of the method.

As part of the first method step three days were spent on the shop floor to observe all relevant process steps and to get in contact with the on-site experts. Additionally, workshops with the relevant management stakeholders were conducted.

The second phase of the method kicked off the process mining. In Fig. 2 two different process models are shown, which were both created via PM. The first model includes all available event data without any filters. This model has many similarities with the so-called spaghetti processes. It created a lot of responses from all stakeholders, who expected a more streamlined process picture. However, the process model shows the need for PM, since the unfiltered process information is not a sufficient basis for ML algorithms. With no defined reference process, it is not clear which process lead time should be predicted and what data can be used to train the model on. The core idea of the model is to split the production process into reasonable parts for a higher prediction accuracy. Therefore, a process model as presented on the left side of Fig. 2 is too detailed for any serious prediction. For these reasons, the process data and event log were filtered with heuristic methods as well as time and attribute-based filters. This was combined with the input of the process engineers and shop floor personal, who verified the plausibility of the process models. The final process model is pictured on the right side of Fig. 2 and resembles much more the often-quoted lasagna process. With a process layout like this the usage of ML is finally possible, since clearly defined starting and ending point events are available, which can be used as features for the following ML application.

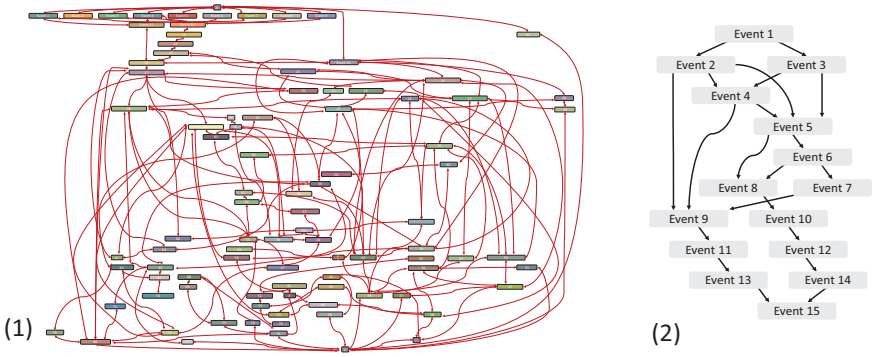


Fig. 2. Process models created with process mining before (1) and after the use of filters (2)

In the third part of the method, ML models were trained with the process data of the newly defined process model. The total input data consisted of roughly 1500 orders. This is a downsized proportion of the overall data set, due to time filters and other limitations imposed by the process mining in step two. Over 100 features from the bill of materials (BOM) were used as features for every order. Additional key order data like the engine type, customer data and so forth were added. The ML models were used to predict both the lead time of the overall process as well as times for the sub processes between each event. Different models like the Support Vector Machines, Naïve Bayes algorithm and different variation of decision trees were used

to predict lead times. All models performed well on the data, with decision tree-based methods like the random forest and bagging taking the top places. A comparison of the PM enhanced ML predictions with the ML predictions, that did not benefit from the PM transformation, was not possible, because the data did not have enough consistent features before the PM transformation to create any meaningful predictions. Tree-based methods had the crucial characteristic that they were much easier to understand for everybody. This was a vast advantage in order to achieve acceptance for the prediction results among all stakeholders. For that reason, the final algorithm, that was used, was a cross-validated decision tree. One of the models can be seen in Fig. 3. It highlights all important features and gives information about the standard deviation and the underlying amount of training data. The model is readable as any decision tree model. For example an engine order with the anonymized typ:B and Feature_1:O would land on the third branch from the left and receive a lead time prediction of 1500 time units. To create the tree models the rpart package in R was used [34].

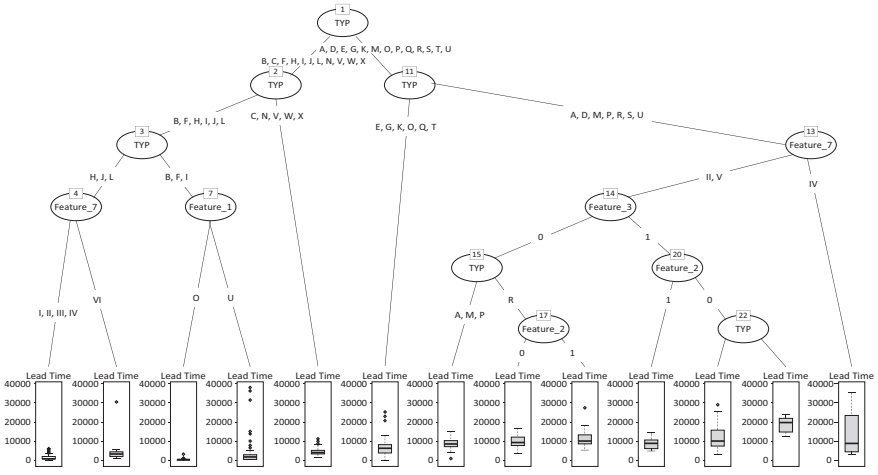


Fig. 3. Tree model for the prediction of the overall lead time

The performance of the tree model presented in Fig. 3 was evaluated with a CP table, which is part of the rpart package. The table is shown in Fig. 4. The values in this table are a common measurement for the performance of regression tree models. For a complete explanation of all values we refer to the works of Therneau et al. [34]. The relative error can be compared to the value of 1 minus the coefficient of determination in linear regressions. The xerror is related to the predicted residual error sum of squares (PRESS) statistic. Both values improve mostly as the size of the tree model increases. However, the amount of branches on the model should not increase too much, since this will lead to an overfitting, which can be observed with the xstd values. Furthermore, this will cause difficulties for the implementation of the predicted lead times into the existing production scheduling process. A problem, which can be solved with a full integration of the ML-model into the MES. The final relative

error of 0.46343 and xerror of 0.55551 are an improvement compared to the differences between the scheduled and realized lead times in the past. Due to confidentiality reasons, the amount of improvement cannot be disclosed.

CP	nsplit	rel error	xerror	xstd
0.4075858	0	1.00000	1.00185	0.099468
0.0453691	1	0.59241	0.66591	0.100772
0.0348406	2	0.54705	0.60097	0.097937
0.0066486	3	0.51220	0.55061	0.096514
0.0051076	7	0.48156	0.56362	0.096930
0.0051008	9	0.47134	0.57973	0.096953
0.0014952	10	0.46624	0.55399	0.096412
0.0013130	11	0.46475	0.55559	0.096314
0.0010000	12	0.46343	0.55551	0.096327

Fig. 4. CP Table for the performance evaluation of the overall lead time model

In the last phase of the method the information from the created ML models were fed back into the business environment. The predicted lead times were used to update the master data in the ERP and MES software. Additionally, a workflow was created to monitor the development of predictions and real data. Overall the combination of PM and ML made it possible to achieve a higher level of precision in the planning of orders and reducing the divergences between planned and actual delivery dates.

5 Conclusion

Machine learning models are highly depended on good input data. In this paper we have presented and verified a new method, which uses process mining in order to enhance the quality of process data in a way that it can be used to create better ML models. The method does not ignore the business context and offers technical help for the choice of algorithms. The feasibility of the method is proven with a case study, which uses real industrial data. It must be noted that method only works for environments in which a minimum of production data is already being recorded. The approach is intended for high variance processes and needs to be tested for different workflows. Likewise, the method focuses on processes with a high share of manual labor and is not yet scrutinized for a highly automated context.

Acknowledgements. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

The authors would also like to thank MAN Truck & Bus SE for their kind support.

Literatur

1. ElMaraghy, H., Schuh, G., ElMaraghy, W., Piller, F., Schönsleben, P., Tseng, M., Bernard, A.: Product variety management. CIRP Ann. **62**, 629–652 (2013)
2. Ivanov, A., Jaff, T.: Manufacturing lead time reduction and its effect on internal supply chain. In: Sustainable Design and Manufacturing 2017, pp. 398–407 (2017)
3. Heaton, J.: An empirical analysis of feature engineering for predictive modeling (2017)
4. Ballambettu, N.P., Suresh, M.A., Bose, R.P.J.C.: Analyzing process variants to understand differences in key performance indices. In: Advanced Information Systems Engineering, pp. 298–313 (2017)
5. Rose, L.T., Fischer, K.W.: Garbage in, garbage out: having useful data is everything. Measur. Interdisc. Res. Perspect. **9**, 222–226 (2011)
6. Mannila, H.: Data mining: machine learning, statistics, and databases. In: Proceedings of 8th International Conference on Scientific and Statistical Data Base Management, pp. 2–9 (1996)
7. Kotsiantis, S.B., Zaharakis, I., Pintelas, P.: Supervised machine learning: a review of classification techniques. Emerg. Artif. Intell. Appl. Comp. Eng. **160**, 3–24 (2007)
8. Choudhary, A.K., Harding, J.A., Tiwari, M.K.: Data mining in manufacturing: a review based on the kind of knowledge. J. Intell. Manuf. **20**, 501–521 (2009)
9. Cheng, Y., Chen, K., Sun, H., Zhang, Y., Tao, F.: Data and knowledge mining with big data towards smart production. J. Ind. Info. Integr. **9**, 1–13 (2018)
10. Lingitz, L., Gallina, V., Ansari, F., Gyulai, D., Pfeiffer, A., Sihn, W., Monostori, L.: Lead time prediction using machine learning algorithms: a case study by a semiconductor manufacturer. Procedia CIRP **72**, 1051–1056 (2018)
11. Meidan, Y., Lerner, B., Rabinowitz, G., Hassoun, M.: Cycle-time key factor identification and prediction in semiconductor manufacturing using machine learning and data mining. IEEE Trans. Semicond. Manuf. **24**, 237–248 (2011)
12. Gyulai, D., Pfeiffer, A., Nick, G., Gallina, V., Sihn, W., Monostori, L.: Lead time prediction in a flow-shop environment with analytical and machine learning approaches. IFAC-PapersOnLine **51**, 1029–1034 (2018)
13. Pfeiffer, A., Gyulai, D., Kádár, B., Monostori, L.: Manufacturing lead time estimation with the combination of simulation and statistical learning methods. Procedia CIRP **41**, 75–80 (2016)
14. Öztürk, A., Kayalıgil, S., Özdemirel, N.E.: Manufacturing lead time estimation using data mining. Eur. J. Oper. Res. **173**, 683–700 (2006)
15. Alenezi, A., Moses, S.A., Trafalis, T.B.: Real-time prediction of order flowtimes using support vector regression. Comp. Oper. Res. **35**, 3489–3503 (2008)
16. Mori, J., Mahalec, V.: Planning and scheduling of steel plates production. Part I: estimation of production times via hybrid Bayesian networks for large domain of discrete variables. Comp. Chem. Eng. **79**, 113–34 (2015)
17. Schuh, G., Prote, J.-P., Molitor, M., Sauermann, F., Schmitz, S.: Databased learning of influencing factors in order specific transition times. Procedia Manuf. **31**, 356–362 (2019)
18. Schuh, G., Prote, J.-P., Sauermann, F., Franzkoch, B.: Databased prediction of order-specific transition times. CIRP Annals. **68**, 467–470 (2019)
19. Windt, K., Hütt, M.-T.: Exploring due date reliability in production systems using data mining methods adapted from gene expression analysis. CIRP Annals. **60**, 473–476 (2011)
20. Bandara, W., Gable, G.G., Rosemann, M.: Factors and measures of business process modelling: model building through a multiple case study. Eur. J. Inf. Syst. **14**, 347–360 (2005)

21. van der Aalst, W.: Process mining: discovering and improving Spaghetti and Lasagna processes. In: 2011 IEEE Symposium on Computational Intelligence and Data Mining (CIDM), pp. 1–7 (2011)
22. Rozinat, A., Jong, I.S.M. de, Gunther, C. W., van der Aalst, W.M.P.: Process mining applied to the test process of wafer scanners in ASML. IEEE Trans. Sys. Man Cybern. Part C (Appl. Rev.) **39**, 474–79 (2009)
23. Park, J., Lee, D., Zhu, J.: An integrated approach for ship block manufacturing process performance evaluation: case from a Korean shipbuilding company. Int. J. Prod. Econ. **156**, 214–222 (2014)
24. Tu, T.B.H., Song, M.: Analysis and prediction cost of manufacturing process based on process mining. In: 2016 International Conference on Industrial Engineering, Management Science and Application (ICIMSA), pp. 1–5 (2016)
25. Pospíšil, M., Mates, V., Hruška, T., Bartík, V.: Process mining in a manufacturing company for predictions and planning. Int. J. Adv. Softw. **6**(3 & 4), 2013 (2013)
26. Knoll, D., Reinhart, G., Prüglmeier, M.: Enabling value stream mapping for internal logistics using multidimensional process mining. Expert Syst. Appl. **124**, 130–142 (2019)
27. Becker, T., Intoyoad, W.: Context aware process mining in logistics. Procedia CIRP **63**, 557–562 (2017)
28. Wallis, R., Erohin, O., Klinkenberg, R., Deuse, J., Stromberger, F.: Data mining-supported generation of assembly process plans. Procedia CIRP **23**, 178–183 (2014)
29. Wirth, R., Hipp, J.: CRISP-DM: Towards a standard process model for data mining. In: Proceedings of the 4th International Conference on the Practical Applications of Knowledge Discovery and Data Mining, pp. 29–39
30. Fayyad, U., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. AI Mag. **17**, 37 (1996)
31. van der Aalst, W., Adriansyah, A., Medeiros: Process Mining Manifesto. In: Business Process Management Workshops, pp. 169–94 (2012)
32. Mitchell, T. M.: Machine Learning. Singapore (1997)
33. Kass, G.V.: An exploratory technique for investigating large quantities of categorical data. Appl. Stat. **29**, 119 (1980)
34. Therneau, T. M., Atkinson, E. J.: An introduction to recursive partitioning using the RPART routines (1997)

Development of a Temperature Strategy for Motor Spindles with Synchronous Reluctance Drive Using Multiple Linear Regression and Neural Network

M. Weber(✉), F. He, M. Weigold, and E. Abele

Technische Universität Darmstadt, Otto-Berndt-Str. 2, 64287 Darmstadt, Germany
m.weber@ptw.tu-darmstadt.de

Abstract. One central aspect of future high speed machining is the knowledge of the thermal behaviour of the machine tool and its motor spindle. A new temperature control for motor spindles with energy efficient synchronous reluctance drive is developed. In the first instance a finite element method model (FEM) is set up. This FEM aims to analyse the use case of nearly constant bearing temperatures within a defined range throughout the machining operation. By means of design of experiments (DOE), selected operating points of the speed-torque-characteristics are simulated with FEM considering different cooling parameters such as volume flow and inlet temperature. Machine learning algorithms are used to model the input-output-relationship in order to reduce the complex thermal motor spindle FEM. The applied algorithms are multiple linear regression and artificial neural network. The concept of temperature strategy, the FEM simulation results and the thermal models using machine learning algorithms are presented.

Keywords: Temperature strategy · Motor spindles · Machine learning

1 Introduction

The challenges of high speed machining are tooling, balancing, thermal and dynamic behaviours, and reliability of the machine tools [1]. Due to the cutting parameter setting with high feed rates and rotational spindle speeds, thermal conditions play a crucial role in attaining high machining accuracy. The thermal error is estimated to be around 40% and 70% of the total machine error [2]. Especially the time-varying temperature distribution of the spindle structure is causal to the thermal displacement error [3]. In addition, time variance of the ambient temperature in the workshop affects machining accuracy [4]. Bossmanns described the problem of predicting the thermal-mechanical behaviour by naming the wide application range of motorized spindle units for high speed machining. Through different cutting processes like

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 538–548, 2021.
https://doi.org/10.1007/978-3-662-62138-7_54

finishing, roughing and drilling, the spindle speed varies, and the bearing arrangement and its preload define the dynamic response behaviour [1]. To take new steps towards further limits in high speed cutting, machine learning algorithms should be applied to simulation data of complex physical models to increase their value-added benefits in manufacturing businesses. The scenario of the simultaneously running thermal and deformation models to the cutting process will contribute to a higher degree of condition monitoring and new control loop application.

2 State of the Art

The approaches for thermal modelling of motorised spindle units can be summarised in four groups: thermal resistance networks, finite difference method (FDM), finite element method (FEM) and machine learning (ML).

Brecher et al. [5] presented a thermal resistance network to study the thermal behaviour of an externally driven grinding spindle. The resulting maximal temperature difference between the model and experimental validation is about 1.5 °C at 7,000 rpm. Bossmanns and Tu [1] developed a thermal model using finite difference method (FDM) to characterize the heat transfer coefficients and heat sinks of a motorized milling spindle. The maximal temperature difference in steady-state operation at 25,000 rpm is 3 °C. Ma et al. [3] used a 3D FEM regarding thermal effects on the resulting contact resistance and stiffness of bearings in transient simulation to study the thermal deformation. The model predicts the bearing temperature in steady-state at 15,000 rpm with less than 1 °C and the axial deformation with less than 2 µm.

Jian [6] used a back propagation neural network (BPNN) to investigate the thermal deformation of a motor spindle in steady-state operation at a maximal speed of 3,000 rpm. The thermal BPNN model has a coefficient of determination (R^2) of 0.9936 and a root mean square error (RMSE) of 0.000562. Yang et al. [7] developed a three-dimensional thermal error model of a motor spindle with maximum speed of 10,000 rpm using least square support vector machine (LS-SVM). The trained model has a R^2 of 0.999 and a RMSE of 0.427. By active compensation in the machine tool the axial and radial (X, Y) thermal errors were improved by 90%, 83% and 66%. Zhang et al. [4] developed an active coolant strategy for thermal balance control of a motor spindle with a differentiated multi-loops bath recirculation system. Mori et al. [8] implemented a machine tool spindle cooling system with on-off control including the pump cut-off during the off-phase at a machining centre and a turn-mill centre. The experimental results under load condition showed that the energy consumption is reduced by 14% to 25%. The effect of thermal displacement fluctuation of the tool centre point of the motor spindle can be controlled by setting different lengths of the on–off interval between 0.4 µm and 3.0 µm. Li et al. [9] developed a variable cooling oil volume method for an externally driven spindle of a machining centre with 10,000 rpm maximum speed. The flow rate is controlled by an implemented thermal multivariate regression model for three axes (X, Y, Z) on an Arduino that sets the driving power frequency of the oil pump. The model predictive R^2 is 0.9675. The control method is validated by both, experimental machining operation and measurements.

The deformation of the spindle can be reduced between 38 and 70% on the Y- and Z-axis. Liu et al. [10] built a transient thermal load model based on heat-fluid–solid coupling FE simulation of a motor spindle. Using extreme learning machine theory, the relationship between the simulated temperatures and the corresponding measured load conditions, as rotational speed and bearing temperatures, was trained in a single hidden layer feedforward neural network. The results will be used for a motor spindle coolant strategy to reduce thermal deformation errors.

A shortcoming of nearly all thermal models is their validity to a limited speed range of the motor spindle. Load conditions like cutting operation or the electrical torque are not considered. For accurate modelling, enough experimental operation data or data obtained by specific experiment planning should be used. Most of the thermal models were trained without consideration of time variance of ambient temperature and external heat transfer through free convection. Spindle housing or placement in headstock and thermal conduction into the machine tool structure are often neglected. All these points influence the model accuracy and the industrial applicability of the models.

3 Application and Approach

3.1 Motor Spindle with Synchronous Reluctance Motor

A motor spindle with synchronous reluctance motor of maximal 30,000 rpm (SynRM) and rigidly preloaded bearing arrangement was developed to take high-speed machining to a new level [11]. The front and rear hybrid angular contact ball bearings are mounted in a tandem O arrangement (see Fig. 2) and Oil-air lubrication is used. The design of the non-locating bearing with a sliding bush is no longer necessary. As a result, the spindle design is more compact. The rigid bearing adjustment over the whole shaft length results in high dynamic stiffness and consequently minimal vibration tendency in the operating speed range. From an economic point of view the added values for machine tool users are higher energy efficiency, machining accuracy and maximum productivity [12]. SynRM have reduced iron losses and no copper losses in the rotor compared to an asynchronous rotor which results in a maximum shaft temperature reduction of 50%. This leads to less axial shaft displacement and increased accuracy [12]. Furthermore, the advantage of the SynRM enables the rigidly preloaded bearing adjustment while additional thermal preload is still induced. The thermally induced preload problem and thus the bearing failure must be prevented [13]. For this purpose, a condition monitoring and active control of the bearing temperature is needed.

3.2 Temperature Control Strategy

In Fig. 1 the system of the temperature strategy with a closed loop control is shown. The advantage of the presented approach is the flexibility in the adaptation of the closed loop control depending on the desired operation mode. For example, fast heating up of the spindle system can be realised, resulting in a short warm up time [14].

Another application of the temperature control strategy is the active control of the bearing preload, which allows the bearing load to be controlled and the bearing life to be positively influenced. Advantages regarding spindle condition monitoring, lower failure probability and maintenance costs are offered to the machine users.

The input values of the temperature control module are the set point values related to the chosen operating mode. The set points can be the average temperatures of the inner and outer bearing ring. The developed thermal model of the motor spindle is used as a condition observer in the closed-loop control. A measured temperature value can be applied for continuous model validation. The process variables of the multi-variable control, flow rate and inlet temperature, are set by a PID controller. Its control para-meters are set depending on the operating point which can be realised by a BPNN [4]. An inverter pump in the chiller sets the flow rate by controlling the speed of the electric drive in alignment with the measured volume flow in the cooling circuit. The disturbance variables are rotational speed and torque which determine the internal heat generation. The heat generated by the cutting process is neglected as the heat is transferred through convection at the rotating tool clamping chuck and the cooling lubricant [1].

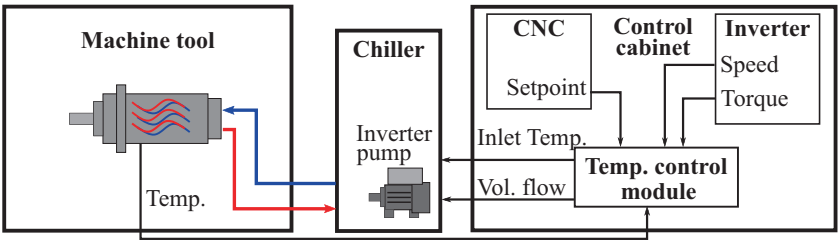


Fig. 1. Temperature strategy for motor spindle units realised in a closed loop.

3.3 Methodical Approach

The presented results were elaborated parallel to the production and assembly phase of the motor spindle. For this reason, no experimental results are shown. The focus is on the development of a methodology to reduce complex physical models. Following the methodical approach, two simulation models of the motor spindle (see Fig. 2) are used.

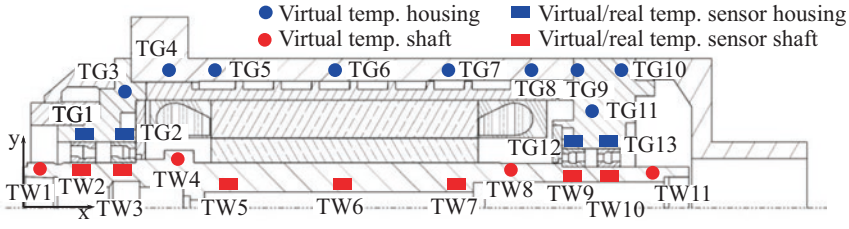


Fig. 2. Virtual und real measuring points of the temperature.

The first one is a mechanical rotor dynamic model of the motorized spindle shaft and the bearing arrangement using MESYS [15]. This software is specific to the calculation of the bearing load condition. In dependence of the spindle speed and given temperature distribution the operating contact angle, the average size of the Hertzian pressure ellipse in the outer and inner contact zone between balls and bearing rings and the resulting bearing forces are calculated. This is mandatory in order to predict the thermally induced preload of the bearing set [1]. With ANSYS® Mechanical this type of accurate bearing arrangement calculation is not provided. The above-mentioned parameters are used to calculate the heat flux density of the bearings. These in turn are input variables of the thermal model.

The second simulation model is a thermal steady-state and magnetostatic model of the motor spindle using ANSYS®. For reasons of symmetry, only one electrical pole of the SynRM is modelled in three dimensions. In order to represent the simulated input-output relationship in the best possible way, considering all main effects, and to minimize the effort of computer experiments, statistical design of experiments is applied after the models are built up. The last step in the methodical process is the modelling of the input–output relationship coming from the FEM model using machine learning algorithms.

4 Thermal FEM Model of the SynRM Motor Spindle

4.1 Thermal FEM Model in ANSYS®

The thermal steady-state FEM of the motor spindle with SynRM and rigidly preloaded angular contact ball bearings is set up following the procedure described in Gebert [16] with respect to the explanations taken from Bossmanns and Tu [1]. The relevant aspects are presented in greater detail in the following section.

The core and armature winding losses of the SynRM are calculated in a separate electromagnetic ANSYS® Maxwell simulation. The eddy current losses are calculated based on the specific losses for each region under consideration of the harmonic effect according to Steinmetz equation:

$$P_{\mathrm{e}} = \sigma_{e1.5}\left(\frac{f}{50\,\mathrm{Hz}}\right)\left(\frac{B}{1.5\,\mathrm{T}}\right)^2 m_{\mathrm{c}} k_{\mathrm{p}} k_{\mathrm{t}} \tag{1}$$

Where f is the supplied frequency of the stator phase current and thus proportional the rotational speed of the SynRM. $\sigma_{e1.5}$ is the specific eddy current loss at $f = 50\,\mathrm{Hz}$ and $B = 1.5\,\mathrm{T}$, m_{c} is the mass of the core section, k_{p} is the pile factor and k_{t} is the sheet metal treatment factor. The hysteresis losses are obtained from:

$$P_{\mathrm{hys}} = \sigma_{\mathrm{hys}1.5} k_{\mathrm{hys}}\left(\frac{f}{50\,\mathrm{Hz}}\right)^2\left(\frac{B}{1.5\,\mathrm{T}}\right)^2 m_{\mathrm{c}} k_{\mathrm{p}} k_{\mathrm{t}} \tag{2}$$

Where $\sigma_{\mathrm{hys}}1.5$ is the corresponding specific eddy current loss and k_{hys} is the factor considering the harmonic effect. The square of the magnetic flux density B

is proportional to the square of the stator phase current I_S and thus to the electrical torque of the SynRM M_{el} [17]:

$$M_{el} = \frac{3}{2}\frac{p}{2}\left(L_d - L_q\right)I_S^2 \sin 2\theta \tag{3}$$

where p is the number of pole pairs, L_d and L_q are the inductances of the direct and quadrature axis and θ is the current angle. The bearing friction losses are calculated according to Gebert [16]. The frictional torque equations due to the lubricant boundary condition (4) and the bearing load (5) are:

$$M_0 = k_{0z} f_0 (\vartheta n)^{\frac{2}{3}} d_m^3 10^{-7} \tag{4}$$

$$M_1 = f_1 Q_l d_m \tag{5}$$

where k_{0z} is a bearing dependent coefficient based on experimental studies of Gebert [16], f_0 and f_1 are coefficients with respect to the lubricant and bearing type. ϑ is the kinematic viscosity of the lubricant oil. d_m is the mean diameter of the bearing. Q_l is the bearing load. The bearing friction losses are determined by:

$$P_b = n(M_0 + M_1)/9550. \tag{6}$$

Air friction losses are considered for every rotating surface of the shaft and the electrical steel sheet lamination stack.

The heat transfer via forced convection is calculated for every rotating surface of the shaft and the rotor [16]. A heat sink represents the surrounding air. The heat transfer is defined by free convection [1] and thermal radiation according to Stefan-Boltzmann-Law. The ambient temperature is constant at 20 °C. The water cooling circuit is the second heat sink. Heat is dissipated through forced convection. Stationary heat conduction through all parts of the motor spindle is considered. The heat transfer in the roller bearings is modelled as described in Gebert [16].

The temperature dependency of the bearing load, the viscosity of the lubricant and of the armature winding resistance is considered using iterative simulation steps including the mechanical, electrical and thermal simulation. The convergence criterion is defined to be less than 3 °C of the average temperature difference of the motor spindle parts from the previous to the current iterative step.

4.2 Design of Computer Experiments

DOE is applied to identify the best input parameter sets for obtaining the main factor effects on the input-output relationship. Different fractional factorial designs based on Box Behnken design (BBD), Taguchi orthogonal array design (OAD) and Monte Carlo method (MCM) were investigated using the Maximin criterion [18] and the entropy criterion [19] (see Table 1). The aim of the Maximin criterion is to maximise the smallest value of eutectic distances between test points. The result is the smallest possible Maximin value. The higher the entropy of an experimental design, the more information it contains. The best fractional factorial design is the OAD. It has

the highest entropy which is more important than the Maximin criterion in terms of modelling. It is composed of 25 factor combinations including 5 factor levels for each factor shown in Fig. 3. The factors are speed, torque, volume flow and inlet temperature. Five factor levels are applied to model the second-order terms in the underlying model equations with the highest possible accuracy. Regarding the speed levels below 10,000 rpm, the heat generation is assumed to be equal or less. This marks the area of the base speed range with the operation at the current limit. The cooling parameter levels are defined by the technical limits of the chiller.

Table 1. Results of the quality criteria of the three experimental designs

Criterion	BBD	OA	MCM
Maximin	19.55	20.94	25.23
Entropy	1.517	1.599	0.792

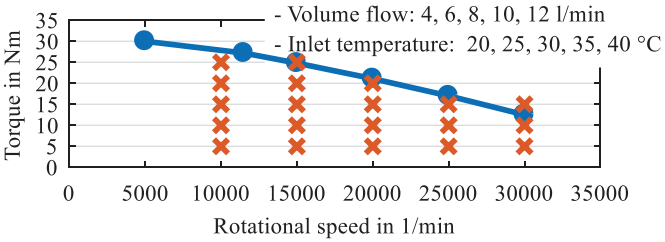


Fig. 3. Design of experiment according to Taguchi orthogonal array design showing the four factors, its five factor levels and the speed-torque characteristic in blue.

4.3 FEM and MESYS Simulation Results

In Fig. 4, an exemplary presentation of the thermal distribution of the factor combination with $n = 30{,}000\,\text{rpm}$, $M = 10\,\text{Nm}$, $\dot{V} = 8\,\text{l/min}$ and $T_0 = 25°\,\text{C}$ is shown. The simulation results of the inner and outer bearing ring temperatures for all 25 factor combinations are displayed in Fig. 5. The average bearing temperatures in Fig. 5 can be specified as the set point value for the temperature control strategy.

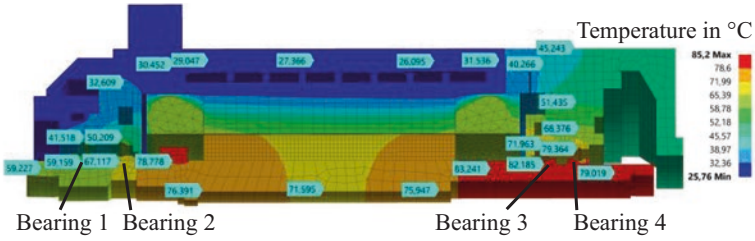
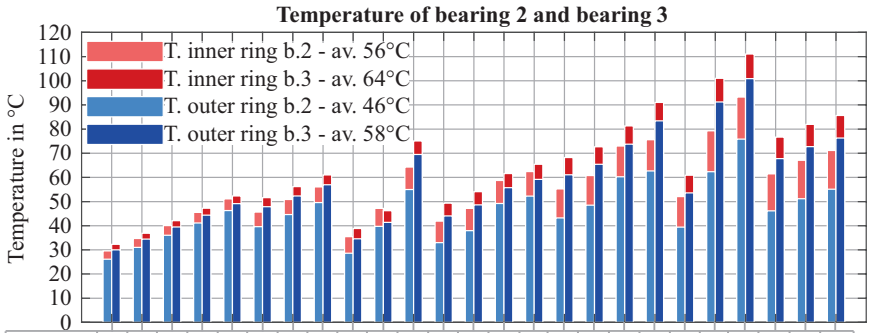


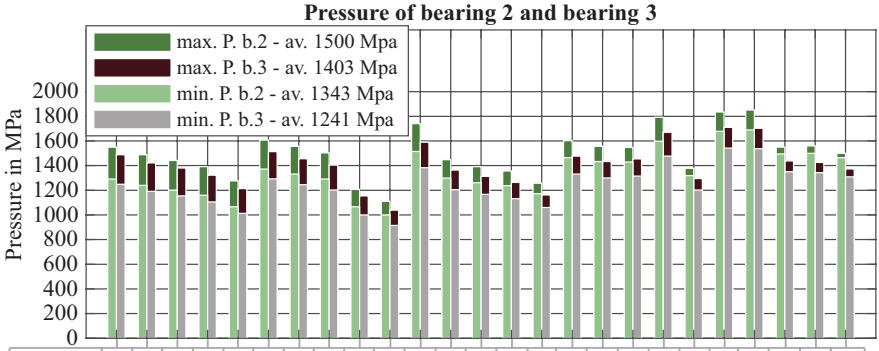
Fig. 4. Simulation result in ANSYS® showing the temperature distribution in the motor spindle at operating point: 30,000 rpm, 10 Nm, 8 l/min, 25 °C.



Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
n in 1000/min	10	10	10	10	10	15	15	15	15	15	20	20	20	20	20	25	25	25	25	25	30	30	30	30	30
M in Nm	5	10	15	20	25	5	10	15	20	25	5	10	15	20	25	5	10	15	5	15	5	10	5	10	12
V̇ in l/min	4	6	8	10	12	6	8	10	12	4	8	10	12	4	6	10	12	4	6	8	12	4	6	8	10
T_0 in °C	20	25	30	35	40	30	35	40	20	25	40	20	25	30	35	25	30	35	40	20	35	40	20	25	30

Fig. 5. FEM simulation results of the second and third bearing temperatures.

The corresponding bearing pressure of the second and third bearing can be seen in Fig. 6. The maximum permissible contact pressure of 2000 MPa for the second and third bearing is not achieved for any factor level combination. This means that it is possible to control the bearing temperature by varying the volume flow and the inlet temperature within the factor limits. Nevertheless, extreme combinations of factors like number 19,21 and 22 should be avoided in the control system in order to keep the risk of a thermally induced preload problem low.



Number	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
n in 1000/min	10	10	10	10	10	15	15	15	15	15	20	20	20	20	20	25	25	25	25	25	30	30	30	30	30
M in Nm	5	10	15	20	25	5	10	15	20	25	5	10	15	20	25	5	10	15	5	15	5	10	5	10	12
V̇ in l/min	4	6	8	10	12	6	8	10	12	4	8	10	12	4	6	10	12	4	6	8	12	4	6	8	10
T_0 in °C	20	25	30	35	40	30	35	40	20	25	40	20	25	30	35	25	30	35	40	20	35	40	20	25	30

Fig. 6. Result of the maximal and minimal pressure of the second and third bearing.

5 Machine Learning Model of the SynRM Motor Spindle

The complex thermal FEM model is reduced to a model which can be calculated more time efficiently by using machine learning algorithms. The thermal multiple linear regression model is compared to the thermal neural network model.

5.1 Multiple Linear Regression Model

In a first step the initial model equation system is derived from the analytical Eqs. (1) to (6). This results in the following theoretical linear regression model:

$$\begin{aligned} \boldsymbol{T}_\mathrm{i} = {} & \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 n_\mathrm{i} + \boldsymbol{\beta}_2 M_\mathrm{i} + \boldsymbol{\beta}_3 \dot{V}_i + \boldsymbol{\beta}_4 T_{0,\mathrm{i}} + \boldsymbol{\beta}_5 n_\mathrm{i} M_\mathrm{i} + \boldsymbol{\beta}_6 \dot{V}_i T_{0,\mathrm{i}} \\ & + \boldsymbol{\beta}_7 n_\mathrm{i}^2 + \boldsymbol{\beta}_8 M_\mathrm{i}^2 + \boldsymbol{\beta}_9 n_\mathrm{i} M_\mathrm{i}^2 + \boldsymbol{\beta}_{10} (n_\mathrm{i} M_\mathrm{i})^2 + \boldsymbol{\varepsilon}_\mathrm{i} \end{aligned} \tag{7}$$

with $\mathrm{i} \in \{1, \ldots, N_\mathrm{train}\}$ as index for an observation from the training data set where $N_\mathrm{train} = 25$. Here the following applies: $\boldsymbol{T}_\mathrm{i}, \boldsymbol{\beta}_0, \ldots, \boldsymbol{\beta}_{10}, \boldsymbol{\varepsilon}_\mathrm{i} \in \mathbb{R}^{1\times 24}$ and $n_\mathrm{i}, M_\mathrm{i}, \dot{V}_i, T_{0,\mathrm{i}} \in \mathbb{R}$. $\boldsymbol{T}_\mathrm{i}$ is the vector of the 24 temperature values obtained by the thermal FEM model. $\boldsymbol{\beta}_0, \ldots, \boldsymbol{\beta}_{10}$ are the regression coefficient vectors and $\boldsymbol{\varepsilon}_\mathrm{i}$ is the vector of normally distributed errors. The thermal model is developed using the statistics software R [20], with the specific package leaps [21]. A model selection is performed automatically to find the set of necessary predictors that fits best. To assess the goodness of fit of the model the criterion of Mallow's C_p-statistic is evaluated. The final predictive thermal multiple linear regression model consists of the following predictors:

$$\widehat{T}_\mathrm{j} = \widehat{\boldsymbol{\beta}}_0 + \widehat{\boldsymbol{\beta}}_1 n_\mathrm{j} + \widehat{\boldsymbol{\beta}}_2 M_\mathrm{j} + \widehat{\boldsymbol{\beta}}_3 \dot{V}_\mathrm{j} + \widehat{\boldsymbol{\beta}}_4 T_{0,\mathrm{j}} + \widehat{\boldsymbol{\beta}}_5 n_\mathrm{j} M_\mathrm{j} + \widehat{\boldsymbol{\beta}}_6 \dot{V}_\mathrm{j} T_{0,\mathrm{j}} + \widehat{\boldsymbol{\beta}}_7 {n_\mathrm{j}}^2 + \widehat{\boldsymbol{\beta}}_8 {M_\mathrm{j}}^2 \tag{8}$$

with $\mathrm{j} \in \{1, \ldots, \mathrm{N_{val}}\}$ as index for an observation from the comparison data set. Here the following applies: $\widehat{T}_\mathrm{j}, \widehat{\boldsymbol{\beta}}_0, \ldots, \widehat{\boldsymbol{\beta}}_8 \in \mathbb{R}^{1\times 24}$ and $n_\mathrm{j}, M_\mathrm{j}, \dot{V}_\mathrm{j}, T_{0,\mathrm{j}} \in \mathbb{R}$. $\widehat{T}_\mathrm{j}$ is the vector of the predicted temperatures. $\widehat{\boldsymbol{\beta}}_0, \ldots, \widehat{\boldsymbol{\beta}}_8$ are the estimated regression coefficients. This model is compared with respect to the thermal FEM model. The Monte Carlo method was used to determine five factor combinations as comparison data set. In line with this, the data training ratio is 80%. R^2 is 0.9946 and RMSE is 0.9954. A second approach to determine the prediction accuracy of the model is used. For this purpose, the mean value and standard deviation of the temperature differences of all 24 temperature output values are calculated (see Table 2). The average prediction accuracy of the thermal multiple linear regression model of the total 5 times 24 temperature output values is $\overline{\overline{\Delta T}} = -0.15 \pm 0.99\,^\circ\mathrm{C}$.

5.2 Neural Network Model

The observations from the thermal FEM model provide 25 simulation results, that form the basis for training a resilient backpropagation neural network using R and its package neuralnet [22]. To compare the possible neural networks, an automated model selection determining the number of hidden layers and neurons depending on

the choice of input variables is performed. This comparison shows that the most accurate network is composed of two hidden layers with four and eight neurons. The corresponding input variables are n, M, $\dot{V}$ and T_0. The model comparison is carried out in analogy to the regression model. R^2 is 0.9900 and RMSE is 1.3572. The average prediction accuracy of the thermal neural network model is $\overline{\overline{\Delta T}} = -0.15 \pm 1.36°C$.

Table 2. Average temperature difference $\overline{\Delta T}$ of the calculated FEM results and the corresponding standard deviation $\sigma_{\Delta T}$

	Monte Carlo operation points as comparison data set				Multiple linear Regression model		Neural network model	
Nr	n in rpm	M in Nm	$\dot{V}$ in l/min	T_0 in °C	$\overline{\Delta T}$ in °C	$\sigma_{\Delta T}$ in °C	$\overline{\Delta T}$ in °C	$\sigma_{\Delta T}$ in °C
1	21,000	5	7	26	1.18	0.99	1.27	1.03
2	14,000	5	11	26	−0.71	0.73	−0.71	0.44
3	6,000	14	8	34	−0.22	0.43	−1.58	0.63
4	12,000	9	9	25	−0.62	0.26	−0.53	0.31
5	27,000	10	11	33	−0.38	0.92	0.80	1.44

6 Conclusion

In this article a thermal multiple linear regression model and a thermal neural network model for a motor spindle with synchronous reluctance motor are developed and compared based on simulation results of a thermal FEM model. The use of the models as state observers in a temperature control strategy is presented. The aim is to enable a rigidly preloaded bearing arrangement as a function of speed, torque and the cooling parameters volume flow and inlet temperature. The thermal models can predict the temperature distribution in the entire motor spindle with an accuracy of $\overline{\overline{\Delta T}} = -0.15 \pm 0.99\,°\mathrm{C}$ respectively $\overline{\overline{\Delta T}} = -0.15 \pm 1.35\,°\mathrm{C}$. This means the thermal multiple linear regression model is selected for use in the closed loop control. As an outlook, the next step will be the experimental validation of the models. Furthermore, investigations are carried out to take other disturbance variables into account, such as the time-varying ambient temperature and transient heat transfer. The practical industrial implementation of the temperature strategy is carried out at a machining centre in a subsequent research project.

References

1. Bossmanns, B., Tu, J.F.: A thermal model for high speed motorized spindles. Int. J. Mach. Tools Manuf. **39**, 1345–1366 (1999)
2. Donmez, M.A., Hahn, M.H., Soons, J.A.: A novel cooling system to reduce thermally-induced errors of machine tools. CIRP Ann. **56**(1), 521–524 (2007)
3. Ma, C., Yang, J., Zhao, L., Mei, X.S., Shi, H.: Simulation and experimental study on the thermally induced deformations of high-speed spindle system. Appl. Therm. Eng. **85**, 251–268 (2015)
4. Zhang, Y., Liu, T., Gao, W., Tian, Y., Qi, X., Wang, P., Zhang, D.: Active coolant strategy for thermal balance control of motorized spindle unit. Appl. Therm. Eng. **134**, 460–468 (2018)

5. Brecher, C., Shneor, Y., Neus, S., Bakarinow, K., Fey, M.: Thermal behavior of externally driven spindle: experimental study and modelling. Engineering **7**, 73–92 (2015)
6. Jian, B.L., Guo, Y.S., Hu, C.H., Wu, L.W., Yau, H.T.: Prediction of spindle thermal deformation and displacement using back propagation neural network. Sens. Mater. **32**(1), 431–445 (2020)
7. Yang, J., Shi, H., Feng, B., Zhao, L., Ma, C., Mei, X.S.: Thermal error modeling and compensation for a high-speed motorized spindle. Int. J. Adv. Manuf. Technol. **77**(5–8), 1005–1017 (2015)
8. Mori, K., Bergmann, B., Kono, D., Denkena, B., Matsubara, A.: Energy efficiency improvement of machine tool spindle cooling system with on-off control. CIRP J. Manuf. Sci. Technol. **25**, 14–21 (2019)
9. Li, K.Y., Luo, W.J., Hong, X.H., Wei, S.J., Tsai, P.H.: Enhancement of machining accuracy utilizing varied cooling oil volume for machine tool spindle. IEEE Access **8**, 28988–29003 (2020)
10. Liu, T., Liu, D., Zhang, Y., Shang, P., Gao, W., Bai, X., Zhang, J., Zhang, D.: Temperature detection based transient load/boundary condition calculations for spindle thermal simulation. Int. J. Adv. Manuf. Technol. **108**, 35–46 (2020)
11. Weber, M., Weigold, M.: High speed synchronous reluctance drives for motor spindles, 116. In: Proceedings of the 15th International Conference on High Speed Machining, Prague (Czech Republic), 8.–9. October 2019, pp. 1–7 (2019)
12. Abele, E., Daume, C., Sielaff, T.: Lagerungsalternativen durch Einsatz von Synchronreluktanztechnologie in Motorspindeln. wt Werkstattstechnik online **107**(7/8), 492–499 (2017)
13. Stein, J.L., Tu, J.F.: A state-space model for monitoring thermally-induced preload in anti-friction spindle bearings of high-speed machine tools. Transactions of the ASME. J. Dyn. Syst. Meas. Control. **6**(3), 372–386 (1994)
14. Weber, M., Helfert, M., Unterderweide, F., Abele, E., Weigold, M.: Synchronreluktanz-Motorspindeln in Werkzeugmaschinen. wt Werkstattstechnik online **109**(1/2), 72–80 (2019)
15. MESYS AG: MESYS software for mechanical engineering, Zurich, Switzerland. https://www.mesys.ag/ (2018). Accessed 14 May 2020
16. Gebert, K.: Ein Beitrag zur thermischen Modellbildung von schnelldrehenden Motorspindeln. Shaker Verlag, Aachen (1997)
17. Matsuo, T., Lipo, T. A.: Field oriented control of synchronous reluctance machine. In: Proceedings of IEEE Power Electronics Specialist Conference – PESC '93, Seattle, WA, USA, pp. 425–431 (1993)
18. Morris, M.D., Mitchell, T.J.: Exploratory designs for computational experiments. J. Stat. Plan. Inference **43**(3), 381–402 (1995)
19. Fang, K.-T.: Design and Modeling for Computer Experiments. Computer Science & Data Analysis. Chapman & Hall/CRC, New York (2005)
20. R Core Team: R: A language and environment for statistical computing. R Foundation for statistical computing, Vienna, Austria. https://www.R-project.org/ (2013). Accessed 14 May 2020
21. Lumley, T., Miller, A.: leaps: Regression Subset Selection. R package version 3.0. https://CRAN.R-project.org/package=leaps (2017). Accessed 14 May 2020
22. Fritsch, S, Guenther, F., Wright, M. N: neuralnet: Training of neural networks. R package version 1.44.2. https://CRAN.R-project.org/package=neuralnet (2019). Accessed 14 May 2020

Concept for Predicting Vibrations in Machine Tools Using Machine Learning

D. Barton(✉) and J. Fleischer

wbk Institute of Production Science, Karlsruhe Institute of Technology (KIT),
Kaiserstraße 12, 76131 Karlsruhe, Germany
david.barton@kit.edu

Abstract. Vibrations have a significant influence on quality and costs in metal cutting processes. Existing methods for predicting vibrations in machine tools enable an informed choice of process settings, however they rely on costly equipment and specialised staff. Therefore, this contribution proposes to reduce the modelling effort required by using machine learning based on data gathered during production. The approach relies on two sub-models, representing the machine structure and machining process respectively. A method is proposed for initialising and updating the models in production.

Keywords: Machine tools · Machining · Vibration · Machine learning

1 Introduction

The performance of machine tools, measured in terms of productivity, availability, product quality, and production costs, depends strongly on the choice of process settings for the machining process. Vibrations during the process are an important factor when defining process settings, as excessive vibrations have a significant effect on tool wear and surface quality. In extreme cases they may cause tool breakage and damage to the machine. A prediction of these vibrations enables an informed choice of settings to achieve the best possible productivity while taking into account quality and cost requirements. [1, 2]

This contribution examines existing experimental and numerical methods to characterise and predict machine vibrations. These existing methods have not achieved widespread use in industry due to high costs and limited scalability. A new concept is proposed to predict vibrations based on a regression model, thus automating the costly modelling process (Fig. 1). The enhanced machine tool could then assist its user in evaluating settings.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 549–558, 2021.
https://doi.org/10.1007/978-3-662-62138-7_55

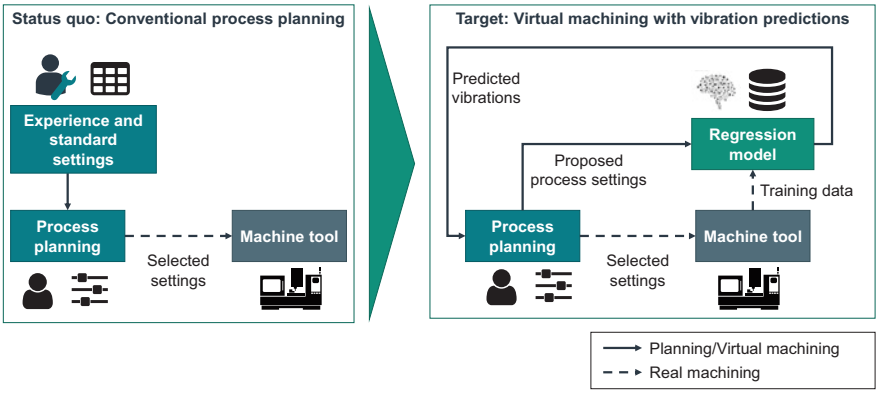


Fig. 1. Enhanced process planning using virtual machining

2 Existing Modelling Methods for Machine Tool Vibrations

In the design process of machine tools, numerical simulation models are required in order to predict the behaviour of structures that have not yet been built. With existing machines, it is possible to identify vibration properties experimentally. In both cases, a model of the machine can be integrated into a simulation of the process-machine interaction.

2.1 Numerical Simulation of Machine Vibrations

Vibration properties of mechanical structures, including machine tools, can be simulated using the finite element method (FEM). Once such a model has been built for a given machine type, predictions can be made concerning static properties (e.g. stiffness) and dynamic behaviour (natural frequencies and modes) [3]. Before complex FEM models can be used to assist in finding suitable process settings, a smaller surrogate model needs to be derived from the original using order reduction methods or machine learning, thus reducing the required computing resources [4, 5].

In order to predict the amplitude of vibrations or to increase the accuracy of the prediction, the model must contain a high level of detail including accurate damping properties. As there are no suitable models to predict damping properties in complex machines, these must be determined experimentally and added to the FEM model, thus increasing the cost of modelling [3]. The relevant mechanical properties of machine components, especially relating to friction and damping, vary widely among machines of the same type. Additionally, due to wear, these properties change over time for a given machine. Thus a model that was tuned for one particular machine will not always be valid for another machine of the same type or a different state of wear [6].

2.2 Experimental Identification and Data-Driven Approaches

Instead of simulating the structure of a machine tool using the methods described above, the vibration behaviour can be determined by performing experiments on a real machine. Currently a method known as experimental modal analysis (EMA) is typically used. This requires the machine structure to be excited by an external force, e.g. using an electromagnetic shaker. The transfer function mapping the measured excitation force to the resulting vibration is known as the frequency response function (FRF). This describes the machine's dynamic stiffness as a function of the excitation frequency [7].

EMA is a costly process requiring specialised equipment and machine downtime, therefore it is not usually performed on every machine. Recent research has focused on identifying machine properties during operation [8–10]. These approaches aim to use cutting forces to replace the external excitation, however they still require targeted experimentation and thus a loss of production time.

Other approaches use machine learning and experimental data to map process settings directly to machining results related to the machine's vibration properties. Using an artificial neural network (ANN), Karkalos et al. predict surface roughness after milling [11]. Several studies focus on predicting the stability of processes, i.e. the occurrence of chatter, thus representing a stability lobe diagram (SLD): Cheruruki et al. use ANN for chatter prediction in turning [12], Friedrich et al. estimate a stability lobe diagram using support vector machines (SVM) and ANN [13], Denkena et al. use kernel interpolation [14]. All these approaches model a setup with a single combination of one machine and one tool. Postel et al. propose a hybrid approach for stability prediction relying on ensemble transfer learning, showing potential for deployment to a broader range of machines and tools [15]. Denkena et al. propose a process planning approach that relies on machine learning models to predict surface roughness in turning operations [16].

2.3 Simulation of Process-Machine Interactions

Once the machine has been modelled using either an FEM simulation or an experimentally determined FRF, this model must be combined with a physical model of the cutting process itself in order to predict the vibrations occurring during operation. This may be achieved by performing a geometric-physical simulation. In this time-domain simulation, the frequency response function of the machine structure is approximated by a set of harmonic oscillators. In each time step, the intersection of geometric models representing the cutting tool and workpiece is determined. Based on this intersection and the cutting speed, empirical cutting force models predict the resulting force, which is applied to the oscillators in the structure model to predict the machine deformation. For the following time step, the workpiece geometry is updated by subtracting the intersection with the tool, and the tool position is updated based on the feed rate and deformation [1, 17].

Geometric-physical simulation is conventionally used to assist when planning machining processes. Recent studies also aim to simulate in real time, parallel to the real machining process. Saadallah et al. train a surrogate model with machine learning

using simulation results [18]. Finkeldey et al. use pre-calculated simulations results to deliver online predictions, taking into account the effect of tool wear during the process, by switching between several process models, each describing a different state of tool wear [19].

3 Regression Model of Vibration Amplitudes in Machining

All the methods presented above require significant expert knowledge, costly specialised measurement equipment and targeted experiments. To achieve widespread use in an industrial setting, an ideal modelling method for machine tools should be entirely automated. Instead of costly experiments, the method should rely on data acquired during regular operation of the machine tool in production. The approach presented in this contribution aims to fulfil these requirements by setting up a regression model to map process settings to the corresponding vibrations measured on the main spindle.

3.1 Overall Architecture

Unlike the existing data-driven approaches mentioned above, the scope of this contribution is to predict vibration amplitudes through a regression model rather than determine whether a process will be stable. The aim is to deliver predictions for multiple tools, materials and machines, while taking into account changes over time in machine behaviour. The present concept covers stable vibrations resulting from open-loop dynamical behaviour. In future work, the authors plan to add a closed feedback loop to the model in order to include the influence of the vibrations on process forces, thus aiming to cover self-excited vibrations.

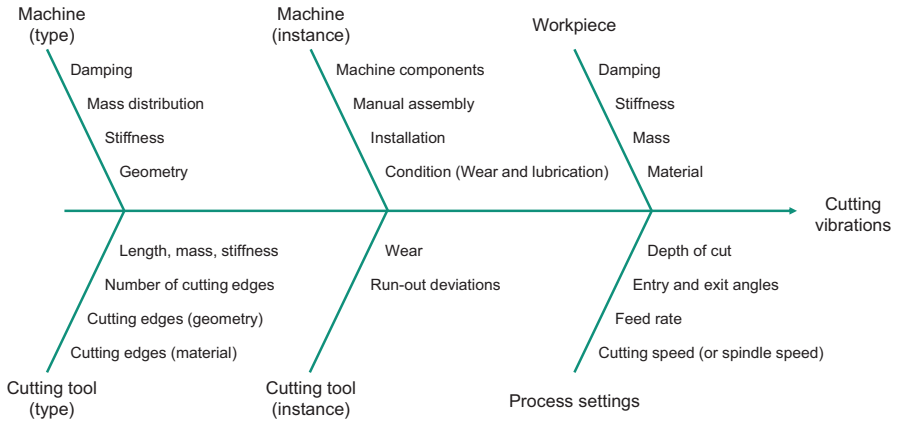


Fig. 2. Cause-and-effect diagram for cutting vibrations in milling [7, 20–22]

An overview of factors influencing cutting vibrations during machining processes is given in Fig. 2 using the example of milling. Additionally, machine tools display

non-cutting vibrations due to other excitations [22]. Given the aim of keeping the model up-to-date throughout the lifecycle of machine tools, the relevant time scales must be considered (Table 1). Depending on the rate of change over time, different strategies are required to take the change into account in the regression model. The factors with the highest rate of change over time (cutting conditions, machine pose, tool wear) are included in the input data of the model, while factors that change slowly are covered by updating the model.

Table 1. Changes in machine behaviour over time.

Approximate time scale	Cause of change
Days - years	Machine wear and maintenance work
Days	Model update
Minutes - seconds	Tool wear
Seconds - milliseconds	Cutting conditions, machine pose

Special consideration is attached to the cutting tool: while properties of the machine itself may remain constant over the course of several process steps and workpieces, machine tools and especially machining centres are typically used with multiple tools. A tool may potentially be used on multiple machines. The tools differ in their geometry (in particular the number of teeth or blades). Additionally the preferred process settings will depend on the tool and the workpiece material. These tool-specific properties influence the frequencies at which the machine structure will be excited by the process. In order to accurately model the machine structure for a wide range of frequencies, the corresponding model should be trained with data from several different tools. Therefore it is proposed to separate the model into two parts: a model of the cutting process and a model of the machine structure's dynamic stiffness (Fig. 3). The process model maps process settings and conditions to the cutting force, which the machine model then maps to the measured vibration. Thus the machine model can be trained using all the available data for a given machine, whereas each process model is trained using the data gathered for this individual process, potentially spanning multiple machines. A machine model is specific to a machine instance, whereas a process model describes a generic combination of a workpiece material with a tool type.

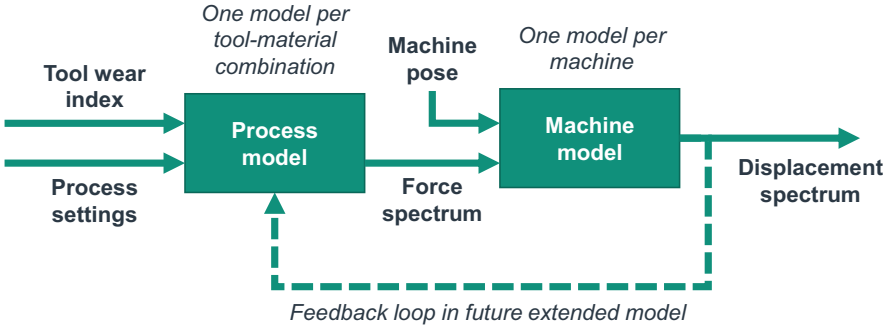


Fig. 3. Model architecture

The effects of workpiece mass and workpiece stiffness on the dynamic behaviour of the machine are not considered in the present article, based on the assumption of small and stiff workpieces. The effects of the tool mass, stiffness and length are also not taken into account, based on the assumption of stiff tools of equal length. However, the concept may be extended to consider modifications in dynamic behaviour based on the principles of receptance coupling [23].

3.2 Input and Output Formats

Inputs for the process model are the cutting conditions, in the case of milling these are the spindle speed n, feed rate v_f, axial depth of cut a_p, and tool engagement angle φ. A wear index for the tool instance, based on the tool's previous load history, is introduced as an additional input. The output from the process model consists in three force signals (one for each direction in space). In accordance with the state of the art, excitations and the resulting vibrations are modelled in the time-frequency domain, therefore the measured force (if available) and acceleration data are subjected to a Fast Fourier Transform (FFT). The FFT is performed with a sliding time window, taking into account the conflicting aims of time and frequency resolution. In training and applying the model, each time step is considered as a separate data point, in combination with the cutting conditions applying at that point in time. The force spectrum in each direction is used directly as a multi-dimensional input and output vector respectively, whereas the measured acceleration spectrum in each direction is integrated twice to obtain a displacement spectrum. In order to consider the effect of machine pose, the position of the feed axes is used as an additional input to the machine model.

3.3 Implementation of Sub-models

Each sub-model (process model and machine model) constitutes a supervised learning task. Given the aim of predicting vibration amplitudes within a continuous range, regression is required. According to studies on a wide range of tasks, suitable machine learning algorithms include random forests, ANN, and SVMs [24]. The process model has a low-dimensional input and high-dimensional output, whereas the inputs and outputs of the machine model are both high-dimensional. In this respect, the required models show similarities to generative models and autoencoders respectively, suggesting a neural network may be suitable. On the other hand, the problem may be converted into multiple models with a single output, each corresponding to the amplitude of force or displacement in a given frequency band. An alternative approach consists in using the frequency as an additional input, thus representing the entire spectrum within a single model with a single output and successively applying the model to individual frequency bands. The latter approach has the advantage of allowing a flexible frequency resolution, which may be used to focus on the most relevant frequency bands (e.g. frequencies with significant amplitudes in a given process).

4 Deployment, Transfer and Update of Models

The present approach is designed to be applied to multiple processes and machines, and adapt to changes in machine behaviour. Therefore a strategy is required for the initial training of models, their transfer to further machines and processes, and model updates based on new data.

4.1 Algorithm Selection and Hyperparameter Tuning

As described above, several algorithms appear to be suitable for the implementation of the sub-models, and each algorithm provides hyperparameters that must be determined. To enable generalisation and transfer, these choices should not be made based on data from only one machine or one process. A dataset should be constituted that is representative for the group of machines and processes to which the approach is to be applied (e.g. all machining centres within a factory, or all machining centres produced by a particular machine vendor). Ideally, this wide ranging set of data is acquired from real production. However, this data isn't generally available, therefore it may be preferable to train with simulated data based on existing methods and validate with measured data.

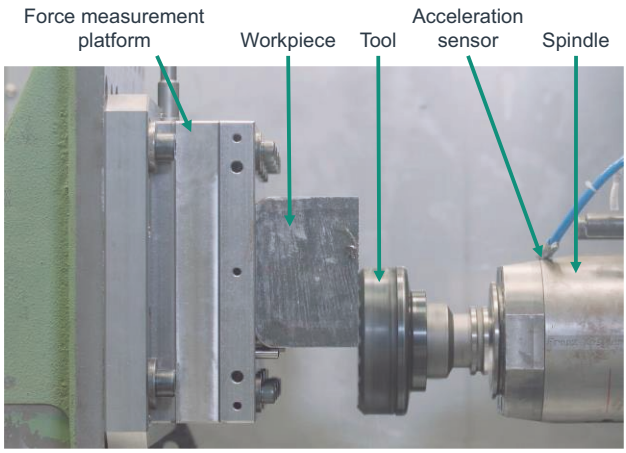


Fig. 4. Experimental setup for initial modelling of cutting processes

4.2 Initial Training of Sub-models

In general, a dynamic measurement of process forces is impractical in industrial production. If a valid machine model is available, this can be used to train the models of processes that are performed using this machine. Conversely, if process models are available for processes that collectively provide a sufficiently broad excitation spectrum, these can be used to train the machine model based on data gathered during these processes. If neither sub-model is available, force measurement must be

performed for one machine. This may be based on a direct force measurement (Fig. 4) or an indirect measurement relying on available signals such as motor currents. After this initial effort, further machines may be modelled as described in the next section.

4.3 Transfer of Models and Model Update

Machines of the same type do not present identical vibration behaviour, however the similar stiffness and mass distribution lead to similar natural frequencies. In order to exploit the similarity between machines of the same type, it is proposed to use pre-trained machine models from similar machines where available, and adapt them by continuing training with data from the specific machine, thus achieving a form of transfer learning. Likewise, it is proposed to transfer process models from processes where the workpiece material and tool type are similar.

During manufacturing operations, the data required for training the models is collected continuously. Periodically, the collected data is reviewed with respect to the accuracy of the existing model and the suitability of the data for training (e.g. broad excitation spectrum). Based on these criteria, the update of a sub-model may be triggered. In this case, the sub-model in question is trained using the newly collected data, while the weights of the other sub-models are frozen. When updating process models, data from multiple machines should be used if available.

5 Conclusions

In this contribution, an overall concept is presented for predicting vibrations in machine tools, consisting of two sub-models: a process model and a machine model. Both the final output (displacement spectrum) and the data transmitted from the process model to the machine model (force spectrum) are represented in the frequency domain. By relying on data acquired during regular manufacturing operation, the approach has the potential to enable a more widespread use of vibration prediction in process planning for machining.

Additional data is required to evaluate the models in detail and select optimal implementations of the sub-models. Further work may also focus on extending the model to consider the influence of self-excited chatter and vibrations not directly caused by the cutting force.

References

1. Altintas, Y., Kersting, P., Biermann, D., Budak, E., Denkena, B., Lazoglu, I.: Virtual process systems for part machining operations. CIRP Ann. (2014). https://doi.org/10.1016/j.cirp.2014.05.007
2. Brecher, C., Esser, M., Witt, S.: Interaction of manufacturing process and machine tool. CIRP Ann. (2009). https://doi.org/10.1016/j.cirp.2009.09.005
3. Schwarz, S.: Prognosefähigkeit dynamischer Simulationen von Werkzeugmaschinenstrukturen. Dissertation. Forschungsberichte IWB, vol. 313 (2015)

4. Bonin, T.: Moderne Ordnungsreduktionsverfahren für die Simulation des dynamischen Verhaltens von Werkzeugmaschinen. Dissertation. Forschungsberichte IWB, Band 306 (2015)
5. Pfrommer, J., Zimmerling, C., Liu, J., Kärger, L., Henning, F., Beyerer, J.: Optimisation of manufacturing process parameters using deep neural networks as surrogate models. Procedia CIRP (2018). https://doi.org/10.1016/j.procir.2018.03.046
6. Reuß, M., Dadalau, A., Verl, A.: Friction variances of linear machine tool axes. Procedia CIRP (2012). https://doi.org/10.1016/j.procir.2012.10.021
7. Brecher, C., Weck, M.: Werkzeugmaschinen, Fertigungssysteme 2. Konstruktion, Berechnung und messtechnische Beurteilung, 9th edn. VDI-Buch. Springer Vieweg, Berlin (2017)
8. Berthold, J., Kolouch, M., Wittstock, V., Putz, M.: Broadband excitation of machine tools by cutting forces for performing operation modal analysis. MM SJ (2016). https://doi.org/10.17973/MMSJ.2016_11_2016164
9. Putz, M., Wittstock, V., Kolouch, M., Berthold, J.: Investigation of the time-invariance and causality of a machine tool for performing operational modal analysis. Procedia CIRP (2016). https://doi.org/10.1016/j.procir.2016.04.052
10. Li, B., Cai, H., Mao, X., Huang, J., Luo, B.: Estimation of CNC machine–tool dynamic parameters based on random cutting excitation through operational modal analysis. Int. J. Mach. Tools Manuf. (2013). https://doi.org/10.1016/j.ijmachtools.2013.04.001
11. Karkalos, N.E., Galanis, N.I., Markopoulos, A.P.: Surface roughness prediction for the milling of Ti–6Al–4V ELI alloy with the use of statistical and soft computing techniques. Measurement (2016). https://doi.org/10.1016/j.measurement.2016.04.039
12. Cherukuri, H., Perez-Bernabeu, E., Selles, M.A., Schmitz, T.L.: A neural network approach for chatter prediction in turning. Procedia Manuf. (2019). https://doi.org/10.1016/j.promfg.2019.06.159
13. Friedrich, J., Torzewski, J., Verl, A.: Online learning of stability lobe diagrams in milling. Procedia CIRP (2018). https://doi.org/10.1016/j.procir.2017.12.213
14. Denkena, B., Bergmann, B., Reimer, S.: Analysis of different machine learning algorithms to learn stability lobe diagrams. Procedia CIRP (2020). https://doi.org/10.1016/j.procir.2020.05.049
15. Postel, M., Bugdayci, B., Wegener, K.: Ensemble transfer learning for refining stability predictions in milling using experimental stability states. Int. J. Adv. Manuf. Technol. (2020). https://doi.org/10.1007/s00170-020-05322-w
16. Denkena, B., Dittrich, M.-A., Stamm, S.C., Prasanthan, V.: Knowledge-based process planning for economical re-scheduling in production control. Procedia CIRP (2019). https://doi.org/10.1016/j.procir.2019.03.238
17. Wiederkehr, P., Siebrecht, T.: Virtual machining: capabilities and challenges of process simulations in the aerospace industry. Procedia Manuf. (2016). https://doi.org/10.1016/j.promfg.2016.11.011
18. Saadallah, A., Finkeldey, F., Morik, K., Wiederkehr, P.: Stability prediction in milling processes using a simulation-based machine learning approach. Procedia CIRP (2018). https://doi.org/10.1016/j.procir.2018.03.062
19. Finkeldey, F., Hess, S., Wiederkehr, P.: Tool wear-dependent process analysis by means of a statistical online monitoring system. Prod. Eng. Res. Dev. (2017). https://doi.org/10.1007/s11740-017-0773-0
20. Quintana, G., Ciurana, J.: Chatter in machining processes: a review. Int. J. Mach. Tools Manuf. (2011). https://doi.org/10.1016/j.ijmachtools.2011.01.001

21. Denkena, B., Hollmann, F. (eds.): Process Machine Interactions. Predicition and Manipulation of Interactions between Manufacturing Processes and Machine Tool Structures. Lecture Notes in Production Engineering. Springer, Berlin (2013)
22. Guo, M., Ye, Y., Jiang, X., Wu, C.: Comprehensive effect of multi-parameters on vibration in high-speed precision milling. Int. J. Adv. Manuf. Technol. (2020). https://doi.org/10.1007/s00170-020-05441-4
23. Park, S.S., Altintas, Y., Movahhedy, M.: Receptance coupling for end mills. Int. J. Mach. Tools Manuf. (2003). https://doi.org/10.1016/S0890-6955(03)00088-9
24. Caruana, R., Karampatziakis, N., Yessenalina, A.: An empirical evaluation of supervised learning in high dimensions. In: Proceedings of the 25th International Conference on Machine Learning, pp. 96–103 (2008)

Automated Profiling of Energy Data in Manufacturing

C. Kaymakci[1(✉)] and A. Sauer[1,2]

[1] Fraunhofer Institute for Manufacturing Engineering and Automation IPA, Nobelstr. 12, 70569 Stuttgart, Germany
can.kaymakci@ipa.fraunhofer.de
[2] Institute for Energy Efficiency in Production, University of Stuttgart, Nobelstr. 12, 70569 Stuttgart, Germany

Abstract. In order to offer energy flexibility in energy markets in short time slots a fast and efficient processing and analysis of data from shop floor to production planning and control is necessary. To this end and to gain more knowledge, different datasets and sources have to be integrated. This paper proposes a conceptual architecture and a method for profiling energy data of manufacturing systems. This includes datasets from information systems as well as physical sources such as sensors, actuators or machine data. Real-life data often come with quality problems like missing and invalid values, outliers or duplicates. The key concept is to automatically identify the necessary metadata for including the dataset in an environment where further analysis and integration of datasets can take place. Moreover, a web service for profiling and visualizing data is implemented.

Keywords: Big data in manufacturing · Data integration · Energy analytics · Energy efficiency · Data profiling

1 Introduction

An increasingly important part of Germany's energy mixture is covered by decentralized power supply based on renewable energies [1]. Since renewable energy sources can be volatile, it is necessary to strike a balance between energy supply and demand [2]. In today's industrial sector, the potential for flexibility is already used to increase demand-side energy flexibility [3]. To enable energy flexibility in energy markets in short time slots the data flowing from shop floor to production planning and control need to be processed and analyzed fast and efficiently [2, 4, 5]. Different energy data streams or energy datasets from different sources must be aggregated, filtered, transformed and analyzed. The challenge lies primarily in the volume, variety and velocity of data from various sources and information systems [6]. One of the main challenges is to integrate different data sources, such as sensors, manufacturing execution systems or external data, e.g. weather forecasts [7].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 559–567, 2021.
https://doi.org/10.1007/978-3-662-62138-7_56

The identification and integration of data from different systems or sources are the main challenges in machine learning and optimization applications [8]. For that reason this paper presents a conceptual data profiling architecture for energy data sources from the manufacturing industry, which identifies metadata and important features of data sources and datasets. The proposed architecture is not an integration platform or a new standard but extracts information and knowledge of datasets including a descriptive profile (e.g. mean, minimum, maximum, standard deviation) derived from the data source. By profiling energy data or other sources from manufacturing and analyzing the data structure, the declarative definition of different data sources into target systems can be improved. The raw dataset is loaded into a stateless two-component view-controller web application. In a first step, a controller for processing and detecting metadata handles the dataset. The dataset and the generated profile is visualized to get a full data profile afterwards. For ingesting data the innovative **data handshake** procedure is introduced that analyzes the structure of the data by rule-based or unsupervised learning methods before ingesting the dataset into subsequent information systems. Using the proposed architecture, data scientists in the manufacturing industry do not need to waste their time on data discovery, data collection or data preprocessing. Hence, generating knowledge and machine learning models for insights into manufacturing processes becomes easier for non-programmers.

The next section explains the necessity of analyzing data in manufacturing and correspondingly the importance of profiling data sources for further analysis. Additionally, it is necessary to evaluate the concept of automated data profiling to get an overview of the different approaches for profiling data and its quality metrics. The third section will present the proposed conceptual architecture of a web-based data profile. The profile operates directly from the data given by a data source. Section four presents an implementation of the architecture where real data is processed and profiled. The last section summarizes the conceptual architecture and shows possibilities in further work.

2 Related Work

The main goal of analyzing energy data in manufacturing systems as defined by DIN EN 50001:2018 [9] is establishing an energy management system to improve energy consumption and energy efficiency, which is the ratio of the output of a system and its energy consumption. Energy management is also used for managing energy flexibility [10]. Therefore, it is necessary to collect energy data from energy consumers with substantial energy consumption. The goal of evaluating energy data is to monitor and understand the energy usage of manufacturing systems [11]. Bränzel et al. [11] define the most important elements of energy data evaluation as a calculation of energy consumption, time-series analysis, development of energy performance indicators (e.g. specific energy consumption) and the correlation of data from different systems. Furthermore, the collected energy data have to be integrated with other data sources from manufacturing systems (e.g. throughput, room temperature, etc.), that could be correlated to energy consumption [11]. Most of the collected energy data

in manufacturing can be considered as a measurement of energy consumption, called time series data [11]. A time series is the output of continuously measured observations of physical objects [12]. Especially in manufacturing systems sensors measure attributes such as temperature, position of an object, or energy consumption of physical objects (e.g. machines or other actuators) [7].

Data profiling minimizes and optimizes both project runtimes and efforts in handling existing data. The term data profiling can, in sum, be characterized as a set of tools and methods to gain knowledge about the metadata of a given data source for further processing [13]. The task of data profiling can be divided into three subtasks: import, profiling by computation and output. The relevant purposes for data profiling are data exploration and data integration.

One of the most innovative approaches in profiling data to support data scientists in data exploration and integration can be found in GOODS from Google [14]. Termite represents and embeds all rows, columns and entities as points where distance metrics are used for profiling [15]. KAYAK [16] is a framework supporting data scientists in data preparation and profiling. Aurum [8] handles problems of data discovery by profiling datasets to an enterprise knowledge graph called EKG. The EKG creates a signature for every dataset for identification purposes. Nevertheless, these approaches are more generic for all types of datasets. Google, for example, handles over 26 billion datasets with requirements quite different from that of manufacturing companies [14].

Gschwandtner et al. [17] present "Know Your Enemy" (KYE) for revealing quality issues in time series data. KYE automatically detects quality issues and problems hidden in the dataset. A taxonomy for quality problems in time series data can be found in [18]. TimeCleanser [19] uses visual analytics and data cleansing for time series data to detect and clean quality issues. Other approaches for profiling time series data and its quality metrics exist in [20–22]. Nevertheless the approaches do not consist an automated detection of content-based metadata such as physical units or the topology of measurements.

3 Data Profile Architecture

In this chapter the more holistic work of [14–16] is combined with the detection of quality issues in time series data in [18]. Therefore, an architecture that automatically profiles time series data considering data like energy consumption in manufacturing is introduced. Additionally a new method called "data handshake" for detecting parameters for ingesting structured manufacturing data such as log files or tabular data is developed. The architecture can also be extended to unstructured and heterogeneous data like images, video or sound. The data profile architecture can be considered as the key enabler for further data processing and machine learning applications.

The goal of software architecture is to describe and organize a system by defining the different components, their modules and relations. It can be considered as a plan for implementing a software system [23]. The proposed architecture is function-oriented and describes the functional components and their relations in terms

of data and information flows. Technological aspects such as programming language or software tools are abstracted to generate a concept for different implementation strategies. The main goal of the proposed architecture is to detect the structure and metadata of a dataset by visualizing the data profile. Data scientists as well as energy managers in manufacturing companies may use the architecture for machine learning and optimization. It is considered as an enabler to start the analytical process by profiling the data. The general outcome of the metadata and profiling process is a data structure called Data Profile Table. The Data Profile Table creates a repository of metadata that is visualized in a data profile to get more information out of the data when using it for machine learning in manufacturing or other applications.

The energy profile architecture is divided into two main components to be seen in Fig. 1: a controller to handle incoming and unknown data as well as a profile dashboard showing the profiling results.

The controller consists of three different modules called **Data Identification**, **Data Integration** and **Meta Processing**. After processing the raw energy data into a data profile, it is pushed to the profile dashboard. The data profile can be visualized on a **Profile Table** and an **Interactive Profile Board** to explore the data in different time dimensions. The next paragraph describes the different modules of the components.

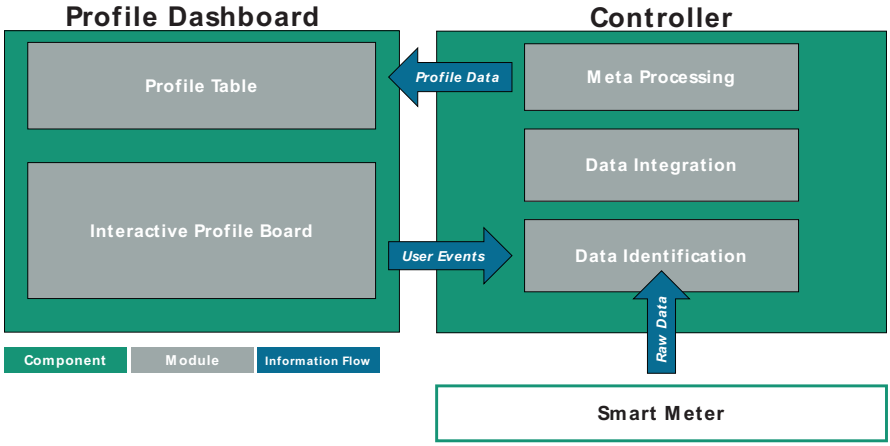


Fig. 1. Conceptual architecture of the data profile dashboard

To integrate data into subsequent tools and applications, it is necessary to detect the separator, the format of values (e.g. decimal sign), and the format of the time stamp (technical metadata). The **Data Identification** module is the key module. It identifies the technical metadata responsible for ingesting and integrating the dataset. When the metadata are detected, the **Data Integration** module integrates the data into a format, which can be processed and analyzed from other modules. The main functionality of the **Meta Processing** module is the calculation of quality metrics and other metadata such as the memory usage. This includes visualization data (e.g. resampled versions) or other features derived from the original data.

After identifying, integrating and processing the data, the metadata can be presented in the profile page component. First of all, a **Profile Table** presents all defined

quality metrics in a tabular format, where each source is defined as a row and the quality metrics as columns. Consequently, the user gets a graphical overview of the metrics for inspecting quality issues such as missing values. The **Profile Table** also shows a small sample of the data to explore the dataset. Visualizations are shown in the **Interactive Profile Board,** where the settings can be adjusted according to the selection of different views and sampling methods. Furthermore, it is possible to switch between different sources. The **Interactive Profile Board** includes functions as zoom in/out or image downloads.

Automatically profiling technical metadata for ingesting and integrating raw data into subsequent systems makes **Data Identification** the most important module of the proposed architecture. The **Data Identification** module can be described as the ingestion layer of the architecture identifying the format of the data. This requires different services.

The *DataStructureIdentifier* service detects the underlying structure of data by analyzing the incoming raw data. The detection algorithm is able to identify common data formats. These are converted into a readable tabular data format. The conversion into a widely used data format is a key aspect of the *DataStructureIdentifier*. It is handled by the "Data Handshake" method (see Fig. 2). The Data Handshake can be compared to handshakes from TCP/IP connections, where two participants agree to communicate with each other by exchanging information to create a connection for transferring data. A Data Handshake is defined as a list of objects in a dataset, where the general structure can be determined. To transform the incoming data into a table-like structure the columns must be discovered. A key pain point while ingesting data from different data sources are the different input parameters to be detected before importing data. (1) In a first step, the controller requests and receives a small slice of data as a pure string. (2) Afterwards, the sample data is analyzed by rule-based methods and natural language processing. The outcome of this step is the actual format and structure of the dataset. (3) The last step of the Data Handshake is to ingest the full dataset. The complete dataset is parsed from the ingestion engine by requesting the raw data from the source. For that purpose, the tabular data format

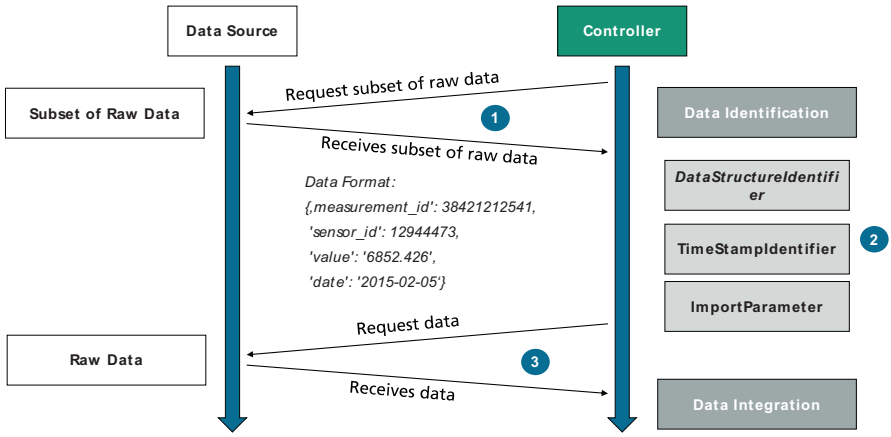


Fig. 2. Data handshake process

(e.g. DataFrame or DataTable) is used which enables high interoperability between different data science tools such as Pandas or R.

Another important service of the **Data Identification** module is called *TimeStampIdentifier*. The *TimeStampIdentifier* service handles the right ingestion of the timestamp column. One of the main problems in ingesting time-series data sources into other systems is the lack of interoperability in terms of timestamp data types varying dependent on different regions or software tools. Thus, the module detects the right timestamp format and converts it to the UNIX data type for time. The *TimeStampIdentifier* works on two levels. First, it checks the patterns and the technical metadata of the columns by using rule-based or unsupervised machine learning methods. Afterwards the timestamp column is parsed to detect the right format of the given timestamp. Nevertheless, there is a chance that the right timestamp column or the format cannot be detected. Therefore, an expert system does not throw an error but interacts with the user, indicating timestamp formats. The right format is chosen from the suggestions of the expert system. In addition, the expert system can actively learn from user feedback.

4 Exemplary Implementation Based on Energy Data

A prototype for importing and profiling historical energy consumption data in different data formats such as CSV, JSON or XML is implemented as a service-oriented web application. The application is programmed as an easily accessible web service for detecting quality issues in the dataset. The prototype contains the Profile Table and the Interactive Profile Dashboard in the frontend and the controller functionalities in the backend where data is processed for the Profile Table and the Interactive Profile Dashboard. Above all, the data handshake represents the backbone and key innovation of processing, detecting all important metadata by using just a small sample of the data source. Using only a small sample means efficient time and memory usage. When considering a CSV file, it is important to detect the delimiter/separator, the timestamp column, and the data types. After deriving this information from the data source, it is possible to import the whole dataset and the profiling can be continued.

An example for the profile table of energy consumption data from a transformer can be found in Fig. 3. The data is measured from a production hall and imported by a REST API or a CSV file. The Profile Table consists of different profile and quality metrics as columns. In this example, it is possible to see data types, the memory usage of the columns, the amount of null values, the descriptive statistics (mean, min, max, standard deviation, etc.) and the amount of unique values. This table can be extended in a modular way by implementing new metrics into the backend of the system. By investigating the quality metrics of data (e.g. missing data) the production or energy manager can see the value of data for further processing.

Profile Table

column_name	col_data_type	non_null_values	unique_values_count	mean	std	min	25%	50%	75%	max	%_of_non_nulls	null_values
energy_median	float64	31132	3051	0.29	0.27	0	0.13	0.21	0.34	4.31	99.99678797417532	1
energy_mean	float64	31132	26752	0.36	0.3	0	0.19	0.28	0.44	3.86	99.99678797417532	1
energy_max	float64	31132	3588	1.18	0.8	0	0.66	1.03	1.52	9.14	99.99678797417532	1
energy_count	int64	31133	30	47.81	2.77	0	48	48	48	48	100	0
energy_std	float64	31036	31030	0.25	0.19	0	0.13	0.21	0.32	2.35	99.6884334950053	97
energy_sum	float64	31132	27163	17.39	14.29	0.01	8.9	13.48	21.15	185.32	99.99678797417532	1
energy_min	float64	31132	881	0.12	0.14	0	0.05	0.08	0.14	1.97	99.99678797417532	1

Fig. 3. The profile table of energy data

The Interactive Profile Dashboard is implemented as a visualization service. From different sources and plots, the service can ingest data as a standardized visualization. One of the key advantages is the visual analysis and comparison of energy consumption between different sources. In the developed implementation different sampling methods are applied for profiling different time intervals on a daily, monthly or yearly basis. Furthermore, it is possible to get the mean or the sum of a defined interval. The interactivity feature consists of zoom-in/out functions, auto-zoom or saving functionalities. Figure 4 shows an example of visualizing a machine in a defined time slot. The configuration of the time slot can be done without any programming. It is possible to adjust the resampling method of the data by interactively adjusting the profile. In datasets with multiple sources, the user can switch between different sources to explore differences in the consumption of energy (e.g. different days, weekdays/weekends).

Fig. 4. Weekly resampled energy data from a machine

The implementation allows to automatically create a data profile by its quality metrics and its data. To this end, the controller processes incoming raw data of different formats and interfaces in an efficient way without using the whole dataset. The profile of the energy consumption data can be used for exploration and further processing. Additionally, the data profile is the starting point for machine learning applications or other optimization methods.

5 Conclusion and Future Work

Different sources, formats, types and communication protocols complicate the ingestion and integration of multiple data sources. In this paper, a conceptual architecture for Automated Data Profiling of energy data is presented by considering time-oriented energy consumption data in manufacturing. The architecture is designed for energy data but also other data with temporal dependencies such as machine data or process data. The concept of the data profile is divided into two modules. The processing and calculation of technical metadata as well as the visual analysis or profiling of the data. These modules of the data profile are presented in the profile dashboard and can be explored by the user to derive the next steps. The presented data profile can be used as key information source for applying methods of machine learning or other time series analysis. The architecture has been developed and implemented for energy consumption data. An application of the profile dashboard could be the forecast of energy consumption by ingesting data from the shop floor level. Another example of using the data profile as a data exploration platform could be the control of process parameter dependent on the energy supply.

Further work will be done on integrating streaming data into the architecture by extending it to a real-time profile. Furthermore, it is planned to build a qualitative extension of the data profile described by a domain expert. Combining the processed profile data and the described profile data in a complete data profile will enable data scientists or domain experts in a manufacturing company to search, query, and analyze different datasets and data sources. One key concept could be to realize a data profile network as a connection between data sources. A data discovery tool and catalogue as in [14] can be implemented to further support data science and machine learning in manufacturing industries.

References

1. Bundesumweltamt: Erneuerbare Energien in Deutschland. Daten zur Entwicklung im Jahr 2018 (2018)
2. Roesch, M., Bauer, D., Haupt, L., Keller, R., Bauernhansl, T., Fridgen, G., Reinhart, G., Sauer, A.: Harnessing the full potential of industrial demand-side flexibility: an end-to-end approach connecting machines with markets through service-oriented IT platforms. Appl. Sci. **9**(18), 37 (2019)
3. Palensky, P., Dietrich, D.: Demand side management: demand response, intelligent energy systems, and smart loads. IEEE Trans. Industr. Inform. **7**(3), 381–388 (2011)
4. O'Donovan, P., Leahy, K., Bruton, K., O'Sullivan, D.T.J.: Big data in manufacturing: a systematic mapping study. J. Big Data **2**(1), 20 (2015)
5. Harding, J.A., Shahbaz, M., Srinivas, Kusiak, A.: Data mining in manufacturing: a review. J. Manuf. Sci. Eng. **128**(4):969–976 (2006)
6. Cui, Y., Kara, S., Chan, K.C.: Manufacturing big data ecosystem: a systematic literature review. Robot. Comput. Integr. Manuf. 62 (2020)
7. Westkämper, E., Löffler, C.: Visionen und strategische Konzepte für das System Produktion Grenzen überwinden mit Strategie und Technologie. In: Westkämper, E., Löffler, C. (eds.) Strategien der Produktion, pp. 71–237. Springer Vieweg, Berlin (2016)

8. Castro Fernandez, R., Abedjan, Z., Koko, F., Yuan, G., Madden, S., Stonebraker, M.: Aurum: a data discovery system. In: Proceedings – IEEE 34th International Conference on Data Engineering, ICDE 2018, pp. 1001–1012. Institute of Electrical and Electronics Engineers Inc. (2018)
9. International Organisation for Standardization: ISO 50001 – Energy Management (2011)
10. Sauer, A., Weckmann, S., Zimmermann, F.: Softwarelösungen für das Energiemanagement von morgen. Stuttgart (2016)
11. Dirk, B., Marko, E., Olaf, G., Schulze, J.: Energiemanagement. Springer Vieweg, Wiesbaden (2019)
12. Ziegel, E.R., Box, G., Jenkins, G., Reinsel, G.: Time series analysis, forecasting, and control. Technometrics **37**(2), 238 (1995)
13. Abedjan, Z., Golab, L., Naumann, F.: Data profiling (2018)
14. Halevy, A., Noy, N.F., Olston, C., Polyzotis, N., Roy, S., Whang, S.E.: Goods: organizing Google's datasets (2016)
15. Fernandez, R.C., Madden, S.: Termite: a system for tunneling through heterogeneous data. In: Proceedings of the ACM SIGMOD International Conference on Management of Data, pp. 1–8. Association for Computing Machinery, New York, USA (2019)
16. Maccioni, A., Torlone, R.: KAYAK: a framework for just-in-time data preparation in a data lake. In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), pp. 474–489. Springer (2018)
17. Gschwandtner, T., Erhart, O.: Know your enemy: identifying quality problems of time series data. IEEE Pacific Vis. Symp. **1**, 205–214 (2018)
18. Gschwandtner, T., Gärtner, J., Aigner, W., Miksch, S.: A taxonomy of dirty time-oriented data. Lect. Notes Comput. Sci. **7465**, 58–72 (2012)
19. Aigner, W., Gärtner, J., Kriglstein, S., Pohl, M., Suchy, N.: TimeCleanser : A visual analytics approach for data cleansing of time-oriented data categories and subject descriptors. In: Proceedings of the 14th International Conference on Knowledge Technologies and Data-Driven Business (2014)
20. Bors, C., Gschwandtner, T., Miksch, S.: Capturing and visualizing provenance from data wrangling. IEEE Comput. Graph. Appl. **39**(6), 61–75 (2019)
21. Kusumasari, T.F., Fitria: Data profiling for data quality improvement with OpenRefine. In: International Conference on Information Technology Systems and Innovation (2017)
22. Arbesser, C., Spechtenhauser, F., Mühlbacher, T., Piringer, H.: Visplause: visual data quality assessment of many time series using plausibility checks. IEEE Trans. Vis. Comput. Graph. **23**(1), 641–65 (2017)
23. Schatten, A., Demolsky, M., Winkler, D., Biffl, S., Gostischa-Franta, E., Östreicher, T.: Software-Architektur. In: Best Practice Software-Engineering: Eine praxiserprobte Zusammenstellung von komponentenorientierten Konzepten, Methoden und Werkzeugen, pp. 199–227. Spektrum Akademischer Verlag, Heidelberg (2010)

Automated Identification of Parameters in Control Systems of Machine Tools

P. Gönnheimer(✉), A. Puchta, and J. Fleischer

Wbk Institute of Production Science, Kaiserstrasse 12, 76131 Karlsruhe, Germany
philipp.goennheimer@kit.edu

Abstract. Especially in the context of Artificial Intelligence (AI) applications and increasing Overall Equipment Effectiveness (OEE) requirements, the use of data in production is gaining in importance. Applications in the field of process or condition monitoring use, for example, machine component parameters such as motor currents, travel speeds and position information. However, as the data is usually only accessible in the machine control systems in non-standard structures and semantics, while having a large number of potential variables, the identification and use of these parameters and data sources represents a significant challenge. This paper therefore presents an approach to automatically identify and assign machine parameters on the basis of time series data. For the identification, feature- and deep learning-based classification approaches are used and compared. Classification results show a general usability of the approaches for the identification of machine parameters.

Keywords: Digital manufacturing system · Identification · Machine tool

1 Introduction

Process and condition monitoring in machines and production systems is becoming increasingly important due to growing requirements on indicators such as Overall Equipment Effectiveness (OEE). Successful approaches in this area, for example for anomaly detection, are based on data from the machine's control system using parameters such as motor currents, travel speeds and position data [1].

In many systems, this control data is already available or accessible to the user, for example via OPC UA. However, the structuring and semantics of the provided data do not follow any standardization and are in most cases machine-specific. Identification and mapping of the data for further use must therefore be done manually. With an increasing number of parameters, which for many OPC UA servers of machines is in the range of several tens of thousands to over one hundred thousand nodes, this poses a great challenge even for experts. In order to facilitate the identification and mapping of these parameters, this paper presents an approach for the automated identification using Machine Learning (ML) as part of a holistic user assistance system.

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 568–577, 2021.
https://doi.org/10.1007/978-3-662-62138-7_57

2 State of the Art and Objectives

2.1 Universal Machine Technology Interface

In order to standardize machine tool information models and therefore facilitate the identification of parameters, the universal machine technology interface (umati) has been developed [2]. Umati provides a specification for the structure and semantics of certain parameters but does not include drive data such as motor currents or position data, which are decisive for central approaches in the area of predictive maintenance and therefore for increasing OEE [1]. Furthermore, umati is limited to OPC UA and therefore does not provide a solution for non-OPC capable machines.

2.2 Rule-based Identification

An existing approach on automated parameter identification comprises five rules with which data can be examined for plausibility and, in the event of a match, assigned to parameters. The rules include comparisons on G-code, range, gradients, correlations and naming of nodes. However, the tool requires a large amount of domain knowledge, which additionally has to be made available, leading to limitations regarding a comprehensive holistic assistance system [3].

2.3 Time Series Classification

Due to the diversity of signals of machine tools, purely rule-based approaches require immense domain knowledge and implementation effort. In order to achieve a universal approach for identification, ML-based classification approaches have proven to be suitable. There are many different approaches to time series classification using ML, e.g. dynamic time warping (DTW) [4], feature-based approaches [5] and deep learning approaches [6]. Deep learning describes Artificial Neural Networks (ANN) with a number of hidden layers [7]. One possibility of a deep neural network is the fully convolutional network (FCN) [6]. An FCN is a Deep Convolutional Neural Networks (CNN) consisting of five parts. The input layer takes the time series into the network and passes it into the second part. Part two to four each consist of one convolutional layer. After each convolutional layer a batch normalization [8] and a Rectifier Linear Unit (ReLU) [9] are applied. The result of the ReLU is passed on to the next layer. The output layer gets its input from part four. It consists of a global average pooling layer [10] and a softmax layer. The softmax layer normalizes the sum of the inputs to one. This produces a probability distribution of the membership probability for each class [6].

As there is no comprehensive suitable solution for the automated identification of key parameters in control systems of machine tools yet, such as motor currents and position data, the objective of the approach presented in this paper is to develop a strategy to identify these parameters by means of ML and rules. This aims to increase the fundamental availability of data in production for improved transparency, data evaluation and consequently also to increase OEE.

3 Approach for the Identification of Control System Data

3.1 Experimental Setup

This paper obtains the data used from a machine tool test bed used at wbk. The machine tool has one vertical and two horizontal linear axes. For each axis the set-point and the actual position are available as well as the torque generating current and the control difference of each axis. In addition, for the spindle the current and torque are given. The cycle parameter, which increases by 1 with each timestep, and a binary control parameter are also specified. In total there are 16 target parameters in the dataset.

Two different processes on the machine tool are used to train and evaluate the identification approach. Both processes represent different processing steps of milling an aluminum workpiece. The first one is a small data set with 12 sections. The second is a large data set with 25 sets. Process 2 is also more than 10 times longer than process 1. The data captured from the machine tool is sampled with 500 Hz. To obtain training and test data sets from the process data, signals from the process data are separated into windows of the same length. The examined signal lengths are 5000 (10 s), 1000 (2 s) and 200 (0.4 s) data points. In addition, z-normalization [11] is applied to each input signal. This is necessary to represent the signals without units and offsets. Every time series has the condition that the standard deviation is not equal to zero, otherwise the z-normalization is not applicable.

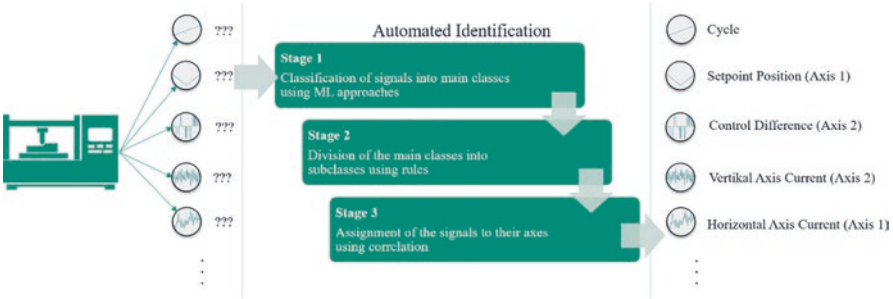


Fig. 1. Identification process of machine tools

The identification process is divided in three stages. The first stage consists of one of the ML approaches classifying the signals. To improve the accuracy of the first stage, classification can be performed multiple times. A feature-based approach was chosen as they have proven to be faster than other approaches [5]. The feature-based approach with the best performance investigated was Random Forests [12]. A deep neural network was also investigated. The structure used here is a CNN. CNNs are the most popular deep neural networks for time series classification [13]. The FCN as a CNN was chosen because of its excellent performance compared to other approaches in time series classification with the manageable depth of the network and the necessary preprocessing [6]. The second stage divides existing classes again into

subclasses. Stage 3 establishes the connection between the classes in order to detect signals of the same axis. The identification process is illustrated in Fig. 1.

The goal is to classify relevant signals in stage 1 with the shortest possible signal length, whereas in stage 2 and 3 the focus is on the detection of the correlation. Independently of the experimental setup examined here, the strategy can be transferred to other types of machine tools, for example with more linear axes. However, only those signals can be identified that can be classified into the parameter classes under investigation. With additional knowledge such as the number of axes, the robustness of the identification process could be improved.

3.2 Setup of Feature-Based Approach

The first of two ML approaches representing the first stage in the identification process is feature-based. The data set is divided into five classes of parameters: binary parameters (Binary), control differences (CTRL_DIFF), currents and torques (CUR/TOR), cycle (CYC) and positions (POS).

In addition to the z-normalization, 40 features are extracted from every time series during data preprocessing. These features include different parameters of signal processing, electrical engineering and statistics. The most important feature for all signal lengths is the maximum of the second derivative. Other important features include the waveform length, the median frequency and the median. A random forest classifier with 20 learners is trained with the training data set consisting of the extracted features. 20 learners showed an appropriate tradeoff between complexity and accuracy of the models. The splitting criteria of the trees is the Gini's index [14]. The maximum number of splits of the trees is limited by the number of observations n in the data set, so the maximum number of splits is n-1. Splits take place until the nodes are pure, i.e. the observations of a node belong to the same class. The feature-based approach with the second stage is shown in Fig. 2.

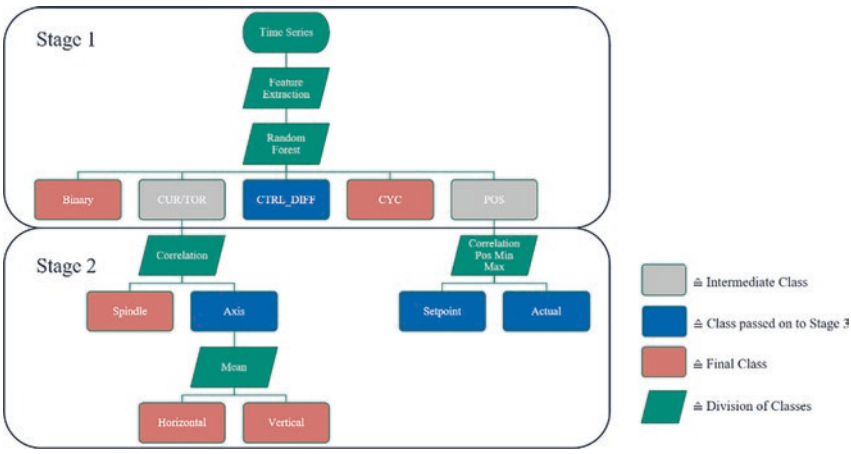


Fig. 2. Feature-based approach including stage 2.

3.3 Setup of Deep Learning Approach

The second ML approach uses deep learning. In contrast to the feature-based approach, this approach covers four classes: binary parameters (Binary), control differences (CTRL_DIFF), currents and torques (CUR/TOR) and positions and cycle (POS/CYC).

The combination of class cycle and positions results from the fact that the nets trained here, which use the class division of the feature-based approach, have all cycle signals assigned to the class positions. The ANN architecture shown here is the architecture of an FCN. The FCNs used here correspond to the structure from [12]. The nets are trained with the optimization algorithm Adam [15]. The loss function is the cross entropy. The nets are trained for 1000 epochs. The batch sizes for the models are 1500 for the signals with a signal length of 200, 1000 for a signal length of 1000 and 300 for a signal length of 5000. The learning rate is 0.001. Apart from the z-normalization of each individual time series, no data preprocessing is executed. After determining the membership probabilities as a result of the networks, a plausibility check is performed. The plausibility check only accepts the class if the score is greater equal 0.8. Otherwise, a new signal from the same data source is classified. The limit of 0.8 was chosen in order to compensate for the lack of subdivision of the cycle and position class of the deep learning approaches. The deep learning approach with the second stage is shown in Fig. 3.

Since the described methods of ML do not allow a further classification of the 16 target parameters with adequate accuracy, rules are used to relate the signals to each other.

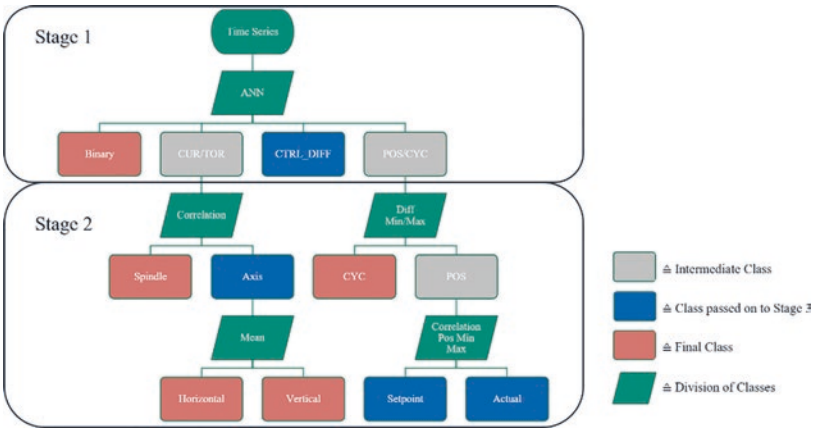


Fig. 3. Deep learning-based approach including stage 2.

3.4 Setup of Stage 2: Rule-Based Assignment

In the second stage of the parameter identification, the classes' currents and torques as well as positions are subdivided again, see Fig. 2. The class currents and torques

is divided into the subclasses spindle and axes. To achieve this separation, the correlations between all signals of the class are determined. This requires simultaneously recorded signals. To calculate the correlation, the Pearson correlation coefficient is used. For each correlation calculation, simultaneously recorded signals are required. The signals of torque and current of the spindle form a pair with increased correlation due to their direct relation, which is assigned to the subclass spindle. The other signals are the axis current signals which are assigned to axis class. To avoid misalignments, it is advisable to carry out a plausibility check to verify the hypothesis. The limit here was set to 0.8, which is derived from the available data sets. A prerequisite for the plausibility check is that the standard deviation is not equal to zero for all signals to be compared. A further subdivision is made in the axis class to detect the horizontal and vertical axis currents. For this, the signals are considered without z-normalization. The characteristic used here is the magnitude of the mean value of the signals. The value of the vertical axis must be greater than the value of the horizontal axis, because the vertical axis counteracts the acceleration due to gravity. Unlike the previous subdivision, this subdivision does not require signals that were recorded at the same time. Offsets for the current signal are untypical, therefore no normalization is necessary.

Before the subclassification of the position class is determined, the position signals from the class cycle and positions of the deep learning approach are required. Therefore, every signal of the class is considered to be mean-free, but not z-normalized. Next, the difference between the successive data points of the signal is formed. If the maximum and minimum of the differences is 1, the signal is a cycle signal. The others are assigned to the position group, see Fig. 3. The position class of both approaches is divided into two subclasses: setpoint and actual position. By calculating the correlations between all signals in the position class, pairs are formed that have a high correlation. These are setpoint and actual position of one axis. To verify the hypothesis, a plausibility check is also performed here. Since the correlations between the associated setpoint and actual position signals are greater, the limit here is 0.95. By checking the position of the minima and maxima at the same time interval, it can be determined which of the two signals is the setpoint and actual position due to the setpoint signal being ahead of the actual position.

3.5 Setup of Stage 3: Axes Assignment

The third stage of identification relates a part of the classes, in this case the assignment of the control differences, current, setpoint and actual position signals of the axes, to each other, see Fig. 4. For this purpose, the correlation is used. The Pearson correlation coefficient is also used here for correlation calculations. In this stage the correlations should also be verified by a plausibility check. The limits defined here result from the available data in order to avoid misalignments. In this way the correlation between the classes of control difference and current axis signals is formed. The pairs with the highest correlation to each other are assigned to an axis. In addition, the correlations of the current signals are calculated with the first and second derivative of the position signals. If one of the two correlations are higher than 0.8 the signals pairs with the highest correlation are assigned to an axis. The limits of the plausibility

checks are again based on the available data to prevent misclassifications. To achieve an even higher correlation for the two assignments, it is possible to take only those parts of the control differences or position signals in which a change in value occurs. The ranges in which this does not happen are removed from the control difference/position signal and the current signals. The signal with the most movement is first used to calculate all correlations with the current signals. The maximum correlation after a plausibility check of 0.8 is matched. Next, the signal with the second largest movement is used and the correlations with the current signals are calculated. After the plausibility check the signal is reassigned. This goes on until the last signal was processed. If the plausibility check is not successful or a signal has a higher correlation to an already assigned signal, new time series must be recorded. The already assigned pairs can then be verified again. Alternatively, the correlation between the first derivative of the difference between setpoint and actual position and control difference can be calculated. Afterwards a plausibility check of 0.8 is performed to verify the result.

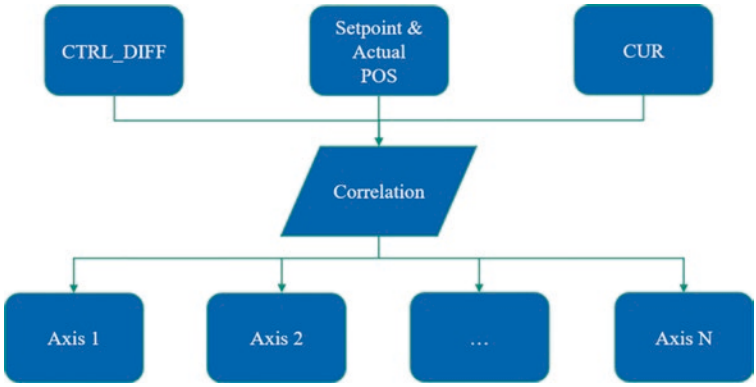


Fig. 4. Axes assignment of stage 3.

4 Results and discussion

4.1 Stage 1: Parameter Identification Using ML

Before the ML approaches are applied, the standard deviations of the signals are calculated. If they equal zero, new signals must be used. In the first stage the classifiers are trained with one of the two process data sets of the machine tool test bed. The data set of the other process is then used to evaluate the two classifiers.

If the classifiers are trained with the small data set from process 1, an increase in the accuracy of the FCN can be seen when the signal length decreases (Table 1). In addition, the random forests are superior to the FCNs for the signal lengths of 1000 and 5000. The influence of the signal length on the random forests compared to the FCNs is small. The accuracy with a length of 200 is comparable for both classifiers.

Table 1. Accuracy of the classifier trained with process 1 and validated with process 2.

Signal Length	5000	1000	200%
Random Forest	84.36%	84.78%	85.74%
FCN	72.35%	82.10%	85.86%

If the two classifiers are trained with the large data set from process 2 and are evaluated with process 1, the overall results of the classifiers are closer, although the networks have an additional plausibility check (Table 2). An increase in accuracy with decreasing signal length can be observed for both approaches. This observation is greater for the deep learning approach than for the feature-based approach.

Table 2. Accuracy of the classifier trained with process 2 and validated with process 1.

Signal Length	5000	1000	200
Random Forest	86.17%	86.46%	88.64%
FCN	84.32%	88.88%	89,66%

The results of the two processes show that the feature-based approach has a better performance for the longest signal length compared to the deep learning approach. The difference in accuracy between the signal length 5000 and 1000 is close in both evaluations. The result of the deep learning approach shows that a shorter signal length leads to a higher accuracy. Reasons for this increase may be the increase in data for shorter signal lengths and the fact that smaller signal segments occur more often in a similar form in other processes than larger segments. This observation can be shown most strongly with the FCN with a difference of more than 10% for different signal lengths for process 1. Compared to the deep learning approach, the feature-based approach has the advantage that it recognizes the cycle signal already in the first stage of identification. However, data preprocessing must take place for the feature extraction. The FCN on the other hand has the advantage that no data preprocessing is needed except for the z-normalization. The proportion of cycle signals increases in the data set with a shorter signal length. If this is considered, the classification of the signals from the present data set shows a tendency towards better classification for the shorter signal lengths 1000 and 200. Usually, the classification of binary signals achieves the highest accuracy for the models. In contrast, the signals of the control differences have the lowest accuracy.

4.2 Stage 2: Rule-Based Assignment

For a better comparability of the results, the complete processes are considered for stage 2 and 3. If correlations are calculated for all current and torque signals in both processes and the plausibility check is carried out, the correlation between the current and torque of the spindle is recognized in 97.3% of cases. The correlations in the process in which the signals could not be clearly assigned are not large enough to be accepted by the plausibility check. In this case a new time series is needed.

For the division of the current axis class into vertical and horizontal axis, the amount of the mean value is used as described in Sect. 3.5. In the present data set, the value of the vertical axis of the complete process is at least four times greater than the value of the horizontal axis in all 37 cases.

With the subdivision of the position class, all pairs of setpoint and actual positions signals belonging together are recognized by forming the correlation. When determining the class affiliation to setpoint and actual positions signal, the signals are always recognized by at least one of the two criteria, the position of the minimum and maximum. Short signal lengths are particularly suitable for this detection, as this prevents similar points in the signal path from being approached and thus the minima and maxima of the setpoint and actual position signals do not refer to the same points.

Starting the second stage with the order of the setpoint and the actual position signals, the correlation is the highest compared to all other plausibility checks in the second stage. Furthermore, the number of parameters in this class must be even. All parameters that do not have such a high correlation to a class member are reclassified.

4.3 Stage 3: Axes Assignment

The assignment between all current signals and all second derivative of the setpoint position signals is possible for process 2 and so in 25 of 37 cases. The setpoint signal is better suited for stage 3 than the actual position signal, as the values of the correlation calculations are usually greater there. If now only the moving parts of the signals are used, all signals from process 2 are still detected.

Table 3. Axis assignments process 1 first derivative of setpoint position and current signals

	Axis 1	Axis 2	Axis 3
Entire signal	11/12	0/12	0/12
Moving parts of the signal	11/12	11/12	1/12

The results of the assignment of the first derivation of the setpoint position signals and the current signals for process 1 are shown in Table 3. Thus, the assignment of axis 2 can be significantly increased if only the moving parts of the signals are taken into account. One assignment of axis 3 can also be made. The setpoint position signal of the third axis has a high correlation to the current signal of the first axis in process 1 when third axis is in motion but is never assigned incorrectly because of the previously assigning of the longest signal. This is because both axes move in the short time that axis 3 is moving. The assignment of the control difference to the current signals is also only possible for process 2. In process 1 the level of correlation is not sufficient.

Alternatively, the signals of control difference and difference of setpoint and actual position signals can be determined for process 2. If the correlation is not high enough, new signals must be recorded. The already assigned signals can then be validated with the new data.

5 Conclusion and Outlook

The use of control data for process and condition monitoring of machines is becoming increasingly important. However, the identification and mapping of the required parameters is in many cases a very complex manual process. This paper presents an approach how machine parameters can be identified automatically using ML.

Future research will focus on validating and expanding the detection strategy with further types of workpieces, tools and machines. Validation will focus on the ML approaches of the identification strategy. Among other things, the influence of a manual preselection of the machine type, for example with regard to the number of axes, will be analyzed. There is also a need for research in handling of unknown parameters, of which the data could be clustered, for instance. In addition, the parameter list will be extended and the identification approach embedded in a holistic user assistance system.

References

1. Netzer, M., Gönnheimer, P., Michelberger, J., Fleischer, J.: Skalierbarkeit von KI-Anwendungen in der Produktion. In: Fabriksoftware, pp. 51–54, Berlin (2020)
2. VDW Homepage. https://vdw.de/technik-und-normung/umati/. Accessed 2 Apr 2020
3. Gönnheimer, P., Hillenbrand, J., Betz-Mors, T., Bischof, P., Mohr, L., Fleischer, J.: Auto-configuration of a digital twin for machine tools by intelligent crawling. In: Production at the leading edge of technology, pp. 534–552. Springer, Berlin (2019)
4. Keogh, E., Ratanamahatana, C.A.: Exact indexing of dynamic time warping. Knowl. Info. Syst. **7**(3), 358–386 (2004). https://doi.org/10.1007/s10115-004-0154-9
5. Baydogan, M.G., Runger, G., Tuv, E.: A bag-of-features framework to classify time series. IEEE Trans. Pattern Anal. Mach. Intell. **35**(11), 2796–2802 (2013)
6. Wang, Z., Yan, W. Oates, T.: Time series classification from scratch with deep neural networks: A strong baseline. In: International Joint Conference on Neural Networks (IJCNN), IEEE, pp. 1578–1585 (2017)
7. Wani, M.A., Bhat, F.A., Afzal, S., Khan, A.I.: Advances in Deep Learning. Springer, Singapore (2020)
8. Ioffe, S., Szegedy, C.: Batch normalization: accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
9. Nair, V., Hinton, G.E.: Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 807–814 (2010).
10. Lin, M., Chen, Q., Yan, S.: Network in network. arXiv:1312.4400(2013)
11. Mathworks Homepage. https://de.mathworks.com/help/stats/zscore.html. Accessed 22 June 2020
12. Breiman, L.: Random forests. Mach Learn. **45**, 5–32 (2001)
13. Fawaz, H.I., Forestier, G., Weber, J., Idoumghar, L., & Muller, P.A.: Deep learning for time series classification: a review. Data Min. Knowl. Disc. **33**(4), pp. 917–963 (2019)
14. Gini, C.: Variabilita e mutabilita. Memorie di metodologia statistica, vamu (1912)
15. Kingma, D.P., Ba, J.: Adam: a method for stochastic optimization. arXiv preprint arXiv:1412.6980, **434** (2019)

Management Sciences

Process Cost Calculation Using Process Data Mining

A. Menges(✉), C. Dölle, M. Riesener, and G. Schuh

RWTH Aachen University, Aachen, Germany
a.menges@wzl.rwth-aachen.de

Abstract. In times of competitive market environments, offering customized products becomes a crucial success factor for manufacturing companies. High variance in the product portfolio leads to expanding costs along the value chain. These costs are not allocated cause-related, which results in cross-subsidization and competitive disadvantage. Process-oriented approaches proved high potential to improve cost allocation. The determination effort often is uneconomically high and results are arbitrarily imprecise. That might result in wrong decisions according the product portfolio and production system. The increasing execution of activities in information systems allows reductions in costs and time as well as increasing quality in process-oriented cost accounting. The aim of this paper is to develop a methodology that uses event data gathered in information systems to determine variant-specific process costs using Process Data Mining.

Keywords: Production system · Product portfolio management · Process cost accounting · Process data mining

1 Introduction

Manufacturing companies in high-wage countries are facing an increasingly dynamic competitive situation [1, 2], which in the globalized world often leads to a growing number of product variants [3]. The increase of external variety is accompanied by expanding internal variety [4], which leads to additional product elements. To manufacture a broad variety of parts, special effort in the production system, e.g. by new machinery, new production processes, additional jigs and tools etc., become necessary [5]. Thus, focusing on the most profitable product variants decrease costs and increase efficiency, effectivity and productivity along the value chain. The main difficulty is the profitability evaluation of a specific product variant, as the corresponding costs are often unknown [6]. The main reason therefore lies in increasing overhead costs [7, 8]. Conventional cost accounting methods allocate overhead costs in proportion to direct costs, which leads to an inaccurate evaluation of profitability, because products with high direct costs do not automatically cause high effort for realization [9]. As a

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 581–590, 2021.
https://doi.org/10.1007/978-3-662-62138-7_58

result, products with low degree of standardization are cross-subsidized by standardized products, which leads to competitive disadvantage compared to companies with a more standardized product portfolio [5].

To address these deficits, process-oriented methods such as Activity-based Costing [10], process cost accounting [11] and Time-driven Activity-based Costing [12] were developed. These methods focus on activities and processes performed for variant realization and determine the corresponding costs by resource consumption [12, 13], which leads to significantly higher cost transparency [14]. Nevertheless, process-oriented cost accounting is rarely applied. Process-oriented cost accounting is based on interviews, which means that the effort for cost estimation is extremely high while the accuracy and reliability are very low [8]. Furthermore, processes in companies are so complex that they can no longer be understood and retraced manually [15]. As a result, the determined costs are arbitrarily imprecise.

Digitization and the execution of activities in information systems offer enormous potential for data-based retracing of processes [16]. Using event data from information systems, Process Mining (PM) aims for the identification of actually executed processes [15]. While Process Mining aims for the identification of processes, the aim of this paper is to determine process costs by analyzing large sets of event data. Thereby, accuracy, trustworthiness and timeliness for the determination of process costs are increased while the effort is dramatically decreased. The calculation of process costs is performed by the development of Process Data Mining (PDM). This allows an efficient evaluation of product variant profitability and thus a trustworthy basis for decisions according the strategic alignment of the production system and the product portfolio.

Therefore, basic definitions and terminology are explained at first to provide an overall understanding of the current challenges in cost accounting as well as trends in digitization. The necessity for the research provided in this paper is derived by the introduction and discussion of existing approaches with the aim of data-based cost accounting in a second step. Based on that, the concept for process cost determination using PDM is presented in a fourth step. Finally, a critical reflection of the methodology, a conclusion and the need for further research is stated.

2 Definitions and Terminology

In this chapter, definitions and terminologies are presented in order to provide a common understanding in the context of this paper.

The **product portfolio** comprises all services, products and product variants of a company [17]. Product variants are characterized as products of similar shape and/or function with a high proportion of identical assemblies or parts [4]. Customers can configure product variants, whereas configuration means the combination of standardized components, taking into account configuration rules [18]. In this paper, mechatronic products, represented by configurable assemblies and parts [5] are focused.

Conventional **cost accounting** methods allocate overhead costs in proportion of direct costs to product variants [14]. Process-oriented cost accounting methods

consider the resources consumed along a product realization process [8], which leads to better cost results [19]. In the context of this paper, process costs are determined by resource consumption for product realization to support decisions according product portfolio and production management.

Data mining is used to identify human-understandable structures and forecast future behavior of systems [20, 21]. Therefore, six main fields of algorithms can be differentiated: classification, regression, clustering, dependency modeling, summarization and deviation detection [22, 23]. In the context of this paper, data mining methods are used to discover process-oriented costs in large sets of event data. As processes become ever more complex, it is almost impossible for companies to retrace the processes actually executed [15]. To gain transparency over processes,

Process Mining aims for discovering processes based on event data, which is gathered automatically in information systems [15, 24]. Therefore, the application fields discovery, conformance and enhancement can be differentiated. To enable PM, a minimum set of event-related information (case id, activity, timestamp) is defined and has to be tracked by process-aware information systems. This paper uses event data and the possibility to gain process-related data automatically from information systems to increase efficiency and trustworthiness of process-oriented costs determination.

Process Data Mining is a so far not established discipline in data science. The aim is to discover knowledge about processes and related disciplines in event data, i.e. process data, using data mining. In the context of this paper, PDM is used to determine variant-specific process costs based on large sets of event data.

3 State of the Art

Existing approaches already focus on the determination and cause-related allocation of variant-induced process costs. Further approaches integrate data science methods in cost accounting to reduce effort and increase accuracy. The state of the art and the need for the research provided in this paper is explained in this section by describing these approaches and analyzing them according the suitability to address the stated issue.

In 1989, Kaplan & Cooper observed increasing overhead costs and identified the need to change methodology in cost accounting. As a result, they developed *Activity-based Costing (ABC)* [9]. Building on these ideas, *Process Cost Accounting* was developed by Horváth & Mayer [11] which was further extended by the *Resource-Oriented Process Cost Accounting* due to Schuh [6, 25]. The aim of these methods is to determine costs for resource consumption of product variants along the complete value chain. The determination of process-oriented costs with these methods caused extremely high effort [8, 14]. To reduce determination and maintenance effort of process-oriented cost models, *Time-driven Activity-based Costing (TDABC)* was developed by Kaplan & Anderson [12], which mainly solved the issue of multi-variate cost drivers. At the same time, *Resource Consumption Accounting* was developed by Keys & van der Merwe [26], which focused on product-specific cost determination

by resource consumption. Although these methods led to considerable improvements in profitability evaluation, they are rarely used in practice due to the high effort and low reliability [14]. Modern cost accounting frameworks address these deficits by relying on methods from data science.

Using data-based methods is a frequently mentioned way to reduce effort and increase accuracy in process-oriented cost accounting [8, 11, 14, 25]. Rising computing power is fueling the implementation of data science over the last years. Kostakis et al. [27] used data mining to identify associations between cost-drivers. Based on quantified associations, cost drivers, which cause high determination effort, are quantified based on those with low determination effort. Yeh & Dang [28] as well as Chang et al. [29] applied data mining in product costing for the estimation of product life-cycle costs and for predicting product unit costs, respectively. Ostadi et al. [30] used fuzzy logic to lower uncertainty and ambiguity in the input parameters for TDABC with the aim to improve reliability of results. While these approaches focus on noise reduction of input data and on improvement of reliability and accuracy, the underlying data is still based on interviews or observations. To overcome these deficits and use automatically gathered and realistic data, Wynn et al. [31] adapted PM to calculate process costs based on resource consumption for activity execution. This approach seems promising to solve the conflicting goals of time, cost and quality in process-oriented cost accounting. Nevertheless, product costing is not regarded and the stated simplifications are not applicable for product cost accounting. In this paper, elements of the described approaches are combined to determine process costs based on event data using PDM.

4 Methodology

In the following, the challenges in process cost determination using existing methods are described more detailed. Based on the identified deficits, a solution is derived by the development and implementation of PDM. Process costs are the sum of activity costs of all activities within the respective process [32].

Therefore, the activities and their respective resource consumption for the regarded process are required. Wynn et al. [31] calculate the process costs for the execution of a process variant, discovered with PM, of a specific case id. This requires knowledge about the case id-specific process variant, which can be determined using PM if a case id runs through the processes once and can be used as criteria for process variant differentiation. In manufacturing companies, unique process variants cannot be retraced for a specific case id, because same case ids could proceed multiple process variants. Figure 1 exemplarily shows the activities and the real processes with five process variants ((1)–(5)) executed for a specific case id. Each process variant has its process costs, which are equal to the sum of the respective activity costs. Furthermore, some activities are performed with relevance to quantity (e.g. purchase); others aren't (e.g. design).

Using PM, it is not possible to discover process variants, as the differentiation criteria (different case ids) is missing [15]. Due to product and process standardization, it is furthermore not possible to use a specific product variant, customer order or anything else as differentiation criteria [5]. Consequently, the calculation of the costs for a single process run of the regarded case id is not possible, because the process processed is unknown. To solve this issue, PDM is developed to statistically evaluate large sets of event data in order to discover knowledge in event data generally and to determine variant-specific process costs specifically. The following sections, the relevant preconditions and the methodology of process costs determination using PDM are described.

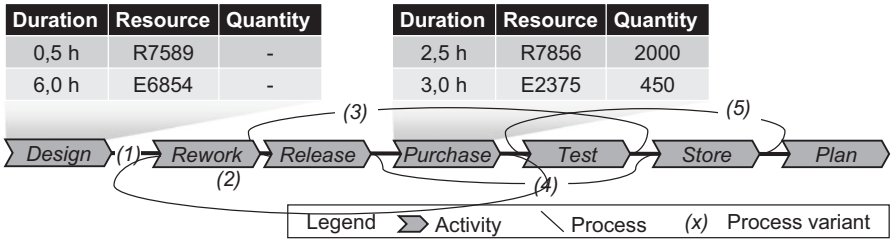


Fig. 1. Real process of a single case id in manufacturing companies

4.1 Preconditions for Process Cost Determination Using PDM

For the calculation of process costs of product elements, the preconditions according to the relevant data and information have to be defined.

For the calculation of variant-specific process costs, the **product portfolio** and the product variants are represented by the elements within the product structure, e.g. parts, assemblies etc. The product elements or their identifiers are used as case ids, so that the corresponding process costs are determined for each product element and will be allocated to the variants they are used in. Therefore, it is necessary to split the process costs in dependency of the production volume of the product element. A unique identifier and it's time-related cost rate represent the **resources** consumed by product elements. Whereas product and resource build the basis for process cost determination, the process or event data gathered in process-aware information systems form the backbone of the process cost calculation. For each event, the respective process element identifier (for cost assignment), the activity (for cost gathering), its start and end timestamp (for duration), the processed quantity of pieces (for cost distribution) and the resources which executed the activity (for event cost calculation) need to be specified. As basis for the calculation, the variables and their sets are defined. So let

- $E = \{e_i | i = 1, \ldots, m\}$ be the set of events with m the amount of events,
- $A = \{a_j | j = 1, \ldots, n\}$ be the set of activities with n the amount of activities,
- $R = \{r_k | k = 1, \ldots, o\}$ be the set of resources,
- $R^c = \{r_k^c | k = 1, \ldots, o\}$ be the set of resource costs with o the amount of resources,
- t_s be the start and t_e be the end timestamp,
- d be the duration, c be the costs and q be the processed quantity of pieces.

4.2 Determination of Process Costs Using Process Data Mining

As stated earlier, process variant discovery is not possible for equal case ids. Nevertheless, through the development of PDM, the analysis of large event data sets is used to calculate the process costs. In the following, process cost calculation is described exemplary for a product element (see Fig. 2).

First, the cost of event i is calculated by $c_i = \left(t_{e,i} - t_{s,i}\right) * r^c_{k,i}$, with $r^c_{k,i}$ the cost rate of the resource r_k consumed in event i.

For further calculation, the assignability of costs to a processed quantity of product elements becomes relevant. Either, an activity is performed for a specific quantity (case a, *qr*), which means that its costs can be divided by the processed quantity of pieces, or it is not (case b, *nqr*), which means, that the corresponding costs have to allocated to the total amount of processed pieces (q_{tot}). q_{tot} defines the reference quantity of the regarded product element.

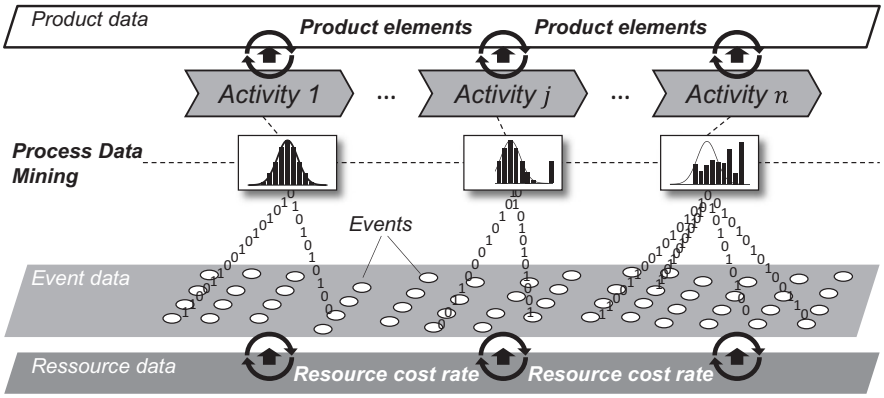


Fig. 2. Determination of variant-specific process costs using Process Data Mining.

Case a. As costs arise *without* regard to processed quantity (e.g. Fig. 1, design), the cost of activity j is calculated by the event costs of the respective activity: $c^{nqr}(j) = \sum_i c_i$.

Case b. As costs arise *with* regard to a specific processed quantity, the cost of processing a single piece of the regarded product element through activity j is calculated based on evaluation of event-specific unit costs using PDM. Therefore, event costs are normalized by $c^n_i = c_i/q_i$ at first to gain the costs of processing one piece through event i. The distribution of c^n_i might be skewed and show outliers due to inconsistent data. The cost of activity j is determined by the median $\tilde{c}^{qr}_j$ of the normalized event costs c^n_i. The median is less efficient, but more robust against outliers compared to the mean and delivers better and more reliable results for data not distributed normally [33].

Furthermore, the likelihood $p(j) = \sum_i q_i/q_{tot}$ for execution of activity j is determined as quotient of the total quantity of pieces processed by the respective activity ($\sum_i q_i$) and the total output quantity q_{tot} of the regarded product element. If activity

j is executed exactly once for processing the total amount of product elements (i.e. is processed in each process variant), the probability for activity execution $p(j) = 1$ (e.g. Fig. 1, process variant (1)). If an activity is conducted multiple times (e.g. Fig. 1, process variant (2), (3), (5)), $p(j) > 1$; if an activity is skipped (e.g. Fig. 1, process variant (4)), $p(j) < 1$. The latter apply if an activity in a process variant is run several times or not every time, respectively. Finally, the process costs c_{tot} are calculated for one unit of the regarded product element:

$$c_{tot} = \frac{c_j^{nqr}}{q_{tot}} + \sum_{j=1}^{n} \tilde{c}_j^{qr} p_j$$

An immediate property of the proposed procedure is that the approximation improves the more event data is available. It is therefore to be expected that its usability will only increase as data collection techniques improve as a result of digitalization.

4.3 Practical Example and Validation

The described methodology is applied to an exemplary case and the results are compared to conventional cost accounting methods. For simplification, two product elements A and B are regarded with a sample set of relevant data summarized in Fig. 3.

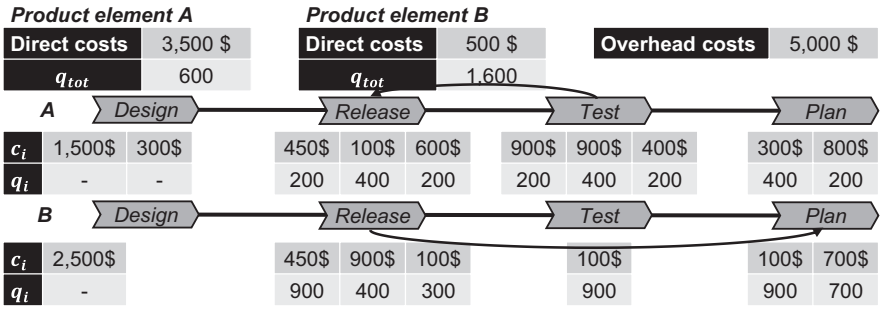


Fig. 3. Data for process cost calculation and cost accounting comparison

With conventional cost accounting, the overhead cost charge is calculated in proportion to the relative share of direct costs. This would result in:

$$\mathbf{A}: c_A = \frac{3{,}500\$}{3{,}500\$ + 500\$} * 15{,}000\$ = 13{,}125\$; \quad \mathbf{B}: c_B = \frac{500\$}{3{,}500\$ + 500\$} * 15{,}000\$ = 1{,}875\$$$

In comparison to that, process cost accounting leads to better cause-related cost allocation. It is obvious that for A the activities "release" and "test" are processed multiple times ($p(j) > 1$), whereas the activity "test" is not processed by each product element unit ($p(j) < 1$). While the stated example is based on a small set of data, Process Data Mining unleashes its potential when analyzing large sets of data.

The results of the calculation are summarized and compared in Table 1. The process costs are not equal to the total overhead costs. To allocate all overheads, the remaining overhead costs are allocated in proportion to the relative share of direct costs [14].

Table 1. Comparison of conventional and process-oriented overhead cost charge

Product element	Conventional overhead cost charge	Process costs ($)	Process-oriented overhead cost charge ($)	Deviation
A	13,125 $ (87.5%)	6,825	2,177	−31.4%
B	1,875 $ (12.5%)	5,687	311	+219.9%

The comparison shows tremendous differences between conventional and process-oriented cost calculation. By using conventional cost accounting, product element A would have been **too expensive** and product element B way **too cheap**.

5 Conclusion and Further Research

With expanding variance in the product portfolio, the knowledge about the profitability of each product variant is extremely important for producing companies. False evaluations could lead to wrong decisions in portfolio management and cause competitive disadvantage. Furthermore, the production system is aligned according to the variants to be produced. Setting up the production system for unsold products causes extremely high costs and effort, which further exacerbates the situation. Cost accounting methods that focus on the determination and cause-related allocation of costs actually incurred are extremely time-consuming, costly and at the same time inaccurate. Digitization shows enormous potential to determine and analyze activities executed for variant realization as well as the corresponding process costs. In order to evaluate big sets of event data to calculate process costs, an algorithm using Process Data Mining was developed in this paper.

The presented algorithm calculates process data without the need for activity sequencing, which is a huge advantage compared to Process Mining. Further research could focus on the implementation of the described algorithm as well as on the cause-related allocation of process costs to corresponding product variants.

Acknowledgements. „Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612".

References

1. Poprawe, R. et al.: Direct, mold-less production systems. In: Brecher, C., Özdemir, D. (eds.) Integrative Production Technology. Theory and Applications, pp. 23–111. Springer, Cham (2017)
2. Schuh, G. et al.: Concept for development project management by aid of predictive analytics. In: Kocaoglu, D.F. (ed.) Proceedings of PICMET 2016: Technology Management for Social Innovation, pp. 2040–2047. PICMET, Hawaii (2016)
3. ElMaraghy, H., et al.: Product variety management. CIRP Ann. **62**(2), 629–652 (2013). https://doi.org/10.1016/j.cirp.2013.05.007
4. Riesener, M.: Ähnlichkeitsbasierte Produktkonfiguration im Maschinenbau. Dissertation, RWTH Aachen (2015)
5. Schuh, G., Riesener, M.: Produktkomplexität managen. Strategien - Methoden - Tools, 3rd edn. Hanser, München (2017)
6. Schuh, G., Kaiser, A.: Kostenmanagement in Entwicklung und Produktion mit der Ressourcenorientierten Prozeßkostenrechnung. In: Männel, W. (ed.) Prozeßkostenrechnung, vol. 51, pp. 369–382. Gabler, Wiesbaden (1995)
7. Miller, J.G., Vollmann, T.E.: The hidden factory. Harv. Bus. Rev. **63**(5), 142–150 (1985)
8. Coenenberg, A.G., et al.: Kostenrechnung und Kostenanalyse, 9th edn. Schäffer Poeschel, Stuttgart (2016)
9. Cooper, R., Kaplan, R.S.: How cost accounting distorts product costs. Manag. Account. **69**(10), 20 (1988)
10. Cooper, R., Kaplan, R.S.: Activity-based systems. Measuring the costs of resource usage. Account. Horiz. **6**(3), 1–13 (1992)
11. Horváth, P., Mayer, R.: Prozesskostenrechnung. Der neue Weg zu mehr Kostentransparenz und wirkungsvolleren Unternehmensstrategien. Controlling **4**, 214–219 (1989)
12. Kaplan, R.S., Anderson, S.R.: Time-driven activity-based costing. SSRN J. (2003). http://dx.doi.org/10.2139/ssrn.485443
13. White, L.: Resource consumption accounting. Manager-focused management accounting. J. Corp. Acct. Fin. **20**(4), 63–77 (2009)
14. Horsch, J.: Kostenrechnung. Klassische und neue Methoden in der Unternehmenspraxis, 3rd edn. Springer Gabler, Wiesbaden (2018)
15. van der Aalst, W.: Process Mining. Springer, Berlin (2016)
16. Aalst, W.M.P.: Process-aware information systems. Lessons to be learned from process mining. In: Transactions on Petri Nets and Other Models of Concurrency, vol. 2, pp. 1–26 (2009). https://doi.org/10.1007/978-3-642-00899-3_1
17. Riesener, M. et al.: Development of a methodology to design product portfolios in accordance to corporate goals using an evolutionary algorithm. In: IEEE IEEM 2018, pp. 1466–1470
18. Link, J.: Database Marketing. In: Bliemel, F., Fassott, G., Theobald, A. (eds.) Electronic Commerce. Herausforderungen - Anwendungen - Perspektiven, 3rd edn., pp. 105–122. Gabler Verlag, Wiesbaden (2000)
19. Deloitte: Komplexitätsmanagement in der Automobilindustrie (2016)
20. Usama, F. et al.: From data mining to knowledge discovery in databases. AI Mag. **17**(3), 37–54 (1996)
21. Hastie, T., et al.: The Elements of Statistical Learning. Data Mining, Inference, and Prediction. Springer Series in Statistics, 12th edn. Springer, New York (2017)
22. Jain, A.K., Dubes, R.C.: Algorithms for Clustering Data. Prentice Hall Advanced Reference Series. Prentice-Hall, Englewood Cliffs (1988)

23. Weiss, S.M., Kulikowski, C.A.: Computer Systems that Learn. Classification and Prediction Methods from Statistics, Neural Nets, Machine Learning, and Expert Systems. Kaufmann, San Mateo (1994)
24. Weijters, T., van der Aalst, W.M.: Process mining: discovering workflow models from event-based data (2001)
25. Schuh, G.: Gestaltung und Bewertung von Produktvarianten. Ein Beitrag zur systematischen Planung von Serienprodukten. RWTH, Aachen (1989)
26. Van Der Merwe, A., Keys, D.E.: The case for resource consumption accounting: it can give you a more accurate projection of the resources you need and their costs than traditional methods can. (Cost Management). Strateg. Financ. **83**(10), 31–37 (2002)
27. Kostakis, H., et al.: Integrating activity-based costing with simulation and data mining. Int. J. Account. Inf. Manag. **16**(1), 25–35 (2008). https://doi.org/10.1108/18347640810887744
28. Yeh, T.-H., Deng, S.: Application of machine learning methods to cost estimation of product life cycle. Int. J. Comput. Integr. Manuf. **25**(4–5), 340–352 (2012). https://doi.org/10.1080/0951192X.2011.645381
29. Chang, P.-C., et al.: Forecasting of manufacturing cost in mobile phone products by case-based reasoning and artificial neural network models. J. Intell. Manuf. **23**(3), 517–531 (2012). https://doi.org/10.1007/s10845-010-0390-7
30. Ostadi, B., et al.: A combined modelling of fuzzy logic and Time-Driven Activity-based Costing (TDABC) for hospital services costing under uncertainty. J. Biomed. Inf. **89**, 11–28 (2019). https://doi.org/10.1016/j.jbi.2018.11.011
31. Adams, M. et al.: Realisation of cost-informed process support within the YAWL workflow environment, vol. 219 (2015). https://doi.org/10.1007/978-3-319-19509-4_1
32. Riesener, M. et al.: Calculation and allocation of complexity costs using process data mining. In: 2019 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 997–1001 (2019). https://doi.org/10.1109/ieem44572.2019.8978735
33. Cramer, E., Kamps, U.: Grundlagen der Wahrscheinlichkeitsrechnung und Statistik. Springer, Berlin (2020)

Refining Circulation Factories: Classification Scheme and Supporting Product and Factory Features for Closed-Loop Production Integration

J. Rickert[1,2], S. Blömeke[1,2], M. Mennenga[1,2], F. Cerdas[1,2], S. Thiede[1,2], and C. Herrmann[1,2]

[1] Institute of Machine Tools and Production Technology (IWF), Chair of Sustainable Manufacturing and Life Cycle Engineering, Technische Universität Braunschweig, Braunschweig, Germany
j.rickert@tu-braunschweig.de
[2] Battery LabFactory Braunschweig (BLB), Technische Universität Braunschweig, Braunschweig, Germany

Abstract. A Circulation Factory is a concept for a future closed-loop production system that combines product manufacturing with remanufacturing and recycling into one hybrid system for increased environmental and economic performance. So far, Circulation Factories have remained a vision. In the present paper, the initial framework for Circulation Factories is revisited and refined by analyzing real world production systems, which are already closing the loop partially. This way, a classification scheme for closed-loop production integration, and a list of product and factory features, which support closed-loop integration, are derived. The feature list is used to analyze if under current conditions the high impact case of Li-Ion batteries is a suitable scope of application for Circulation Factories.

Keywords: Circulation Factory · Closed-loop production · Circular Economy

1 Introduction

Manufacturing creates value, and is directly or indirectly responsible for s significant share of the local and global impacts on the environment. Through products and jobs, manufacturing contributes to high standards of living and prosperity. However, manufacturing and related upstream chains also require substantial amounts of (non-) renewable materials and energy, and releases emissions into air, water and soil. Therefore, manufacturing is indispensable on the one hand, but also poses substantial risk to the environment and human health on the other hand [1].

The demand for products and thus manufacturing is growing due to a rising global population, increasing global standards of affluence, and is still impelled by

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 591–601, 2021.
https://doi.org/10.1007/978-3-662-62138-7_59

linear economies, which mainly follow the principle of 'take-make-dispose' [2]. Consequently, future manufacturing needs to decrease its detrimental environmental impacts. Thus, sustainability in manufacturing has become more important, and has led to various research, which strives towards that 'products are designed, produced, distributed, used and disposed with minimal (or none) environmental and occupational health damages, and with minimal use of resources' [3]. This is in line with the Circular Economy (CE) paradigm, which comprises end-of-life management options, to cascade either complete products, or components or materials into additional life cycles. The aim of CE is to organize, engineer and control the activities, which enable the conservation of the residual economic and environmental value of unwanted or unsuitable (End-of-Use, EoU), and degraded or inoperable (End-of-Life, EoL) products [4]. The cascading options have the potential to mitigate environmental consequences related with product manufacturing by 'closing the loop': harmful waste disposal practices like landfilling are avoided, and the demand for primary components and materials is decreased by providing secondary components and materials with a lower environmental impact [5]. In general, smaller loops (re-use or re-manufacture products or components) are considered preferable, as they are usually associated with less environmental impacts in comparison to wider loops (material recycling). The protopia of a Circulation Factory (CF) closes the loops and keeps them as small as possible by integrating product manufacturing as well as remanufacturing and recycling into one hybrid system under the same roof [6].

The present paper refines the initial CF framework [6] by analyzing real world closed-loop production systems and deriving a classification scheme for closed-loop production integration, as well as product and factory features, which support closed-loop production integration. These features are used to analyze if the high impact case of Li-Ion batteries (LIBs) is a suitable scope of application for a CF.

2 Refining the Circulation Factory Framework

2.1 State of the Art on Circulation Factories

In the following, the CF foundations are compiled and illustrated with a conceptual framework. The CF concept combines manufacturing, versatile (dis-)assembly, quality inspection, material recycling and remanufacturing capabilities (Fig. 1). A CF may be organized as a flexible matrix system, in which versatile working units (workstations or production machines) can perform manufacturing and assembly, as well as disassembly and remanufacturing tasks. The individual working units are linked through a flexible material transport with an intelligent decentralized control logic, which assigns workpieces and allows for a cycle-independent flow [7]. Consequently, workpieces may choose different paths through the CF depending on their properties, subsequent tasks to be performed, and the availability of the working units, or depending on an overarching factory strategy, i.e. prioritizing output or resource efficiency. The importance of a digital factory control, i.e. continuously monitoring, mirroring and optimizing resource flows, is stressed in the defining framework.

The CF framework offers a number of potentials and challenges. According to [8] a main problem for remanufacturing is the limited availability of spare parts, whereas this is not a problem within a CF, in which spare parts can be manufactured. Essential remanufacturing/recycling information is also available, as a CF would treat its 'own' products. The CF approach promises increased environmental and economic performance by utilizing secondary materials and components with lower environmental impacts on the one hand, and by coupling and sharing technical building services (TBS) for production and retro-production on the other.

Challenging remains the absence of the highly flexible (retro-)production and material transport systems. Problematic is also the time offset between the start of the production and retro-production sections within a CF, as ramping up remanufacturing and recycling capacities is only feasible when significant amounts of EoU/EoL products are flowing back. Consequently, one of the main challenges for real world implementation of CFs are the fluctuations regarding returning products, and materials in terms of quantity and quality. While there are numerous publications dealing with economic aspects of closed-loop production such as such as production planning and coordination [9–11], the engineering has remained immature and Circulation Factories a vision.

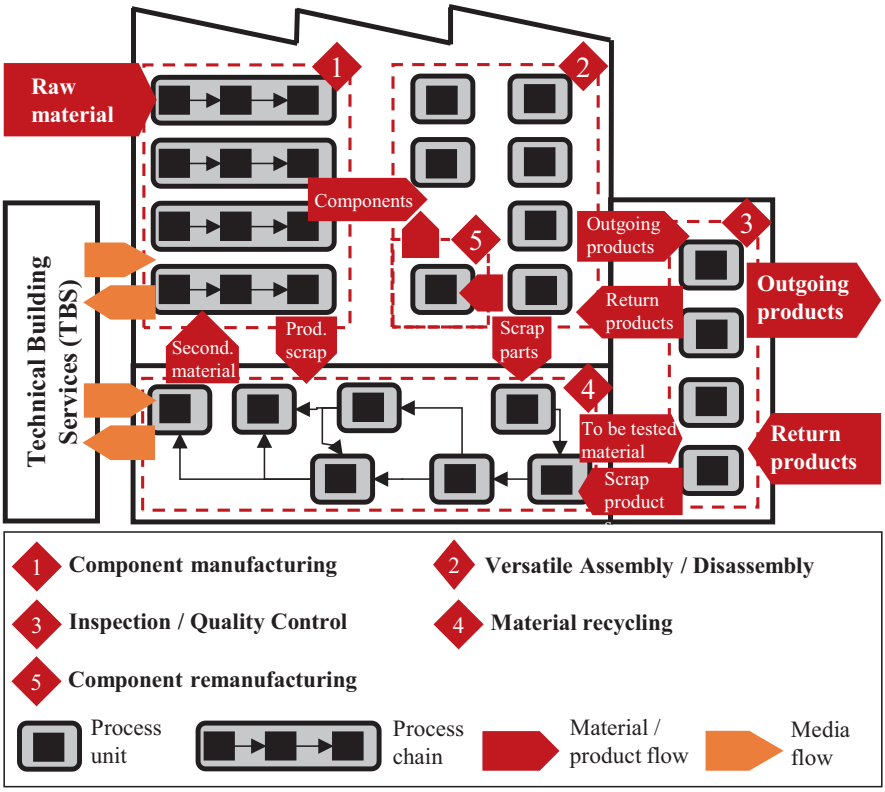


Fig. 1. Conceptual framework of the Circulation Factory. (Adapted from [6])

2.2 Closed-loop production in practice

In the following, existing approaches for closed-loop production systems are presented and analysed. *Tang and Teunter* [12] present a case study of a company, which produces and remanufactures water pumps for diesel engines. Around 120,000 water pumps are put out, and approx. 20% are remanufactured, either for new engine assembly or as a spare part. Although there are around 100 different water pump models, many of them differ only slightly, making a grouping in five product families possible. Assembly for all product families is performed on the same production line, regardless of whether the components are newly manufactured or remanufactured. *Van der Laan et al.* [13] introduce the case of a Dutch manufacturer of photocopiers, which have a modular, easy-to-disassemble structure. Modules, which satisfy quality requirements are remanufactured, and assembled with newly manufactured modules to obtain a copier, that is sold as a new product. *Schmidt et al.* [14] present the case study of Lorenz GmbH, which produces and remanufactures water meters. The company takes advantage of the fact, that water meters need to be replaced periodically by law, but the brass hydraulic components of the meters do not deteriorate in the same time. Upon return, the products are disassembled and suitable components are supplied to the assembly of new meters. Approx. 25% of outbound products use remanufactured components. Another example is an *eyeglass frame producer* from Sydney. The self-designed frame system is offered in multiple sizes and numerous colours and comprises interchangeable parts for the main frame as well as for the temples and connecting pins. The company offers a special product family, for which they source approx. 75% of the material by recycling waste plastic from the urban vicinity of central Sydney. The manufacturing chain for the special product family includes sorting and shredding of waste (beer keg lids from nearby bars and breweries, used fishing nets or milk bottle tops from cafés), injection moulding, cutting of the lenses and assembly of the frame. The complete closed-loop production is set up in a single specialised store.

2.3 Classification of closed-loop production integration

While the visionary concept is a fully integrated system of product manufacturing as well as remanufacturing and recycling, it is possible to identify manifestations of partial integration as previously shown in Sect. 2.2. In the following, the presented cases are analysed with the question in mind of how much they resemble a fully integrated CF to derive a classification scheme of closed-loop production integration as shown in Table 1.

The degree of closed-loop production or CF integration can vary in different dimensions. For example, the *depth* of closed-loop integration, which can range from no integration (only forward production) to the fully integrated system as described in Sect. 2.1 and Fig 1. In between intermediate forms exist, e.g. ranging from the supply of materials through recycling, to providing components, or the final product itself through remanufacturing as shown by the examples of Sect. 2.2. The lines between the intermediate forms are also blurry, as further combinations, e.g. production *and* component remanufacturing *and* material recycling, are possible as well. Possibly, the

CF concept can go even further in this regard and incorporate servicing mobile, in-use products. Other dimensions include:

- *Product variety* (e.g. a CF system which is able to valorise a range of products through disassembly and material recycling, whereas the forward-oriented manufacturing and assembly could only be able to produce a single product);
- *Material sourcing* (ranging from primary resources only to different combinations of primary and secondary material sources);
- *Spatial* dimension (e.g. separated, or separated, but coordinated facilities or CF 'under one roof').

Table 1. Classification scheme of closed-loop production integration

Degree of integration	No	Intermediate			Full	
Depth	Only production	+ *material* recycling	+ *component* reman.	+ *product* reman.	+ service	All
Product variety	Production of one product	Combinations of the production of one/multiple products, and the recovery of one/multiple products			Production + recovery of multiple products	
Material sourcing	Primary resources only	Primary + secondary resources of high quality (e.g. post-industrial)		Primary + secondary resources of uncertain source (e.g. post-consumer)	All; focus on secondary sources	
Spatial	Separated facilities	Separated, but coordinated facilities			'under one roof'	

All examples from Sect. 2.2 represent intermediate forms on the CF classification scheme. Regarding closed-loop integration depth, the case of eyeglass frames integrates *material* recycling, the photocopier manufacturer integrates production and *component* remanufacturing, and finally the water pumps and water meters examples integrate *product* remanufacturing. In terms of product diversity, the material recycling case has the greatest spectrum of products, which gets recovered for production. The companies from the photocopier, water pump and water meter remanufacturing examples concentrate on a narrow set of own products. Information concerning the material sourcing relationships from the presented examples are vague. Only in the water meter case a post-consumer material sourcing seems likely. All presented approaches are integrated 'under one roof'. Finally, all examples report improved economic and environmental performance through the closed-loop production integration.

2.4 Supporting Product and Factory Features

The presented examples from Sect. 2.2 show the feasibility of closed-loop production integration, but also show that the success of closed-loop production integration largely depends on product and factory features. Therefore, in the following sub-chapter, a list of supporting product and factory features is assembled and expanded.

The following product features stand out from the remanufacturing cases: high modularity, easy to disassemble joints, durability, and high residual economic value in relation to remanufacturing effort. Other features, such as a high number of reusable components, or the ability of the product to be identified are taken from publications concerning design for recycling or remanufacturing such as [15, 16].

Flexible (manufacturing) equipment can be identified as supporting factory feature from the water pump and water meter case, as it enables the assembly regardless of whether the components are newly manufactured or remanufactured. This desirable feature is extended to changeability according to [17] and applies to other factory elements such as the building shell and especially TBS as well. Factory systems become changeable, if their constituent elements adhere to the principles of universality, scalability, modularity, mobility and compatibility [17]. Scalable and modular TBS with the correct capacity for the joint supply of production and retro-production can lift economic and environmental efficiency. Current systems e.g. for heating, ventilation, and air conditioning (HVAC), or for the supply of compressed air are often oversized, inefficiently controlled and can account for up to 40% of factory energy demand [18]. Coupling waste heat flows e.g. from compressed air generation with demand from HVAC systems, or manufacturing and recycling processes can also be promising. A precondition for this coupling is the metering of resource flows and demands, which can be achieved by a wide application of sensors, and the collection, consolidation and processing of the sensor data, or in other words through digitalization. The full list of supporting product and factory features can be found in Table 2.

Table 2. List of supporting product and factory features

Product feature	Description (desirable feature manifestation)
Handleability	Workable dimensions regarding mass or size (low), Material hazardousness/toxicity (none, low),
Legal constraints	Restrictions e.g. regarding specialised storing or transport; or regarding legally limited service life, e.g. operating hours or number of use cycles (none, low); prescribed periodical exchange or calibration (high)
Complexity	Product complexity can be described through following sub-features: If considered for *remanufacturing*: Modularity (high), Number of reusable components (high), Use of standardized components (high), Use of standardized, easy to disassemble joining technologies (high) If considered for *recycling*: Recyclability of materials (high), Number of materials (low), Material dilution (low)
Durability	Wear resistance or ability to withstand multiple life cycles (high)
Diversity	E.g. in regards to product variants (low)
Back-/Forward compatibility	Ability of evolved components to be used in earlier product variants, and upgradeability of the whole product (high)
Demand cycle	Period of product demand (long)
Residual value	Residual value, either *economic* as residual component or material value (high), or *environmental* e.g. expressed as embodied energy, or in regard to criticality (high)
Availability of product information	E.g. in regards to above mentioned features like material composition; but also as the *ability of the product to be identified*, for example through product information carriers such as labels, bar codes or RFID tags (high)
Factory feature	
Changeability	*Flexibility* as the ability of a manufacturing system to switch with reasonable effort to a new, although similar product (high), *Transformability* as the ability of an entire factory to switch to another product (high); enabled by factory elements, which adhere to the principles of *universality, scalability, modularity*, *mobility* and *compatibility* [17].
Digitalization	Digitalization as the basis for decentralized, autonomous control of flexible material transport systems, and as the foundation for continuous monitoring and databased optimization of resource flows (high).

Product and factory features are have strong interdependencies. On the one hand, there are product design guidelines, which increase the producibility, on the other hand production and retro-production systems are usually designed to serve product characteristics. However, both, product and factory features, also have a strong influence on where a specific closed-loop production system can be placed on the classification scheme (Table 1).

The product features make closed-loop production feasible in the first place. In terms of closed-loop integration depth, only durable parts and products can be used for remanufacturing. The same is true concerning joining technologies: only if non-destructive disassembly is possible, the product or component can be considered for remanufacturing. Regarding spatial integration, the handleability has significant influence, as transport between facilities may be cumbersome in regards to large, heavy or hazardous products and parts. Finally, only if the residual product, component or material value or related savings potential is high enough, a closed-loop production may be considered by companies. Together with adequate factory features, such as digitally supported and changeable factory elements, a closed-loop production can be achieved. This is illustrated by the fact, that only a closed-loop production system with flexible equipment can serve a broad product variety, while being spatially concentrated. Modularity can enable a step-wise integration in terms of depth. For example, starting with a forward-production only, the capacities could be gradually expanded to remanufacturing, while making sure that the product, material and information flow is kept operational through compatible systems.

3 Application Towards Li-Ion Batteries

Li-Ion batteries are the essential component of battery electric vehicles from an economic and environmental perspective and thus represent a cornerstone for the societal pursuit towards sustainable mobility. Remanufacturing and recycling LIBs, and thus re-circulating the valuable electrode materials (e.g. NMC: Ni, Mn, Co & Li), is considered pivotal for improving the environmental outlook of electric mobility [19], and due to a demanding supply situation. Different approaches for recycling have been developed and are scaled up, without consensus on which process chain is economically or environmentally optimal [20].

The list of previously identified, supporting product and factory features (Table 2) is used to analyse if the high impact case of LIBs is a suitable scope of application for implementing a CF. An overview is given in Table 3, but not all features are addressed due to missing information.

Table 3. List of supporting product and factory features applied to LIBs

Product feature	Description
Handleability	LIBs are heavy and bulky, and some have odd shaped housings for fitting into converted combustion engine cars. LIBs pose electrical, fire and explosion, and chemical hazard and thus need specialised transport, storage and processing.
Legal constraints	Spent LIBs require specialised storing and transport by law. Transport boxes or - if the LIB has been damaged - leak proof containers are mandatory for transportation. Storing LIBs has legal constraints as they contain toxic electrolytes.
Complexity	LIBs are complex: high modularity is contrasted by being safety-critical, and permanent joining technologies such as gluing and welding. For recycling, the number of materials and material dilution are obstructive.
Diversity	LIBs are comparatively new products, without industry-wide standards as of yet [19]. Product development is still ongoing, resulting in a large variety of material combinations, formats, sizes and joining technologies.
Residual value	Not all spent LIBs can be used for 2nd life, but material value is still high.
Availability of product information	Data on the state of health may be acquired through the battery management system (BMS) and can be used for deciding on the fate of the used battery (2nd life or recycling). Product information on systems level such as detailed material composition is scarce, which is challenging for recycling companies, forcing them to build up experience.
Factory feature	
Changeability	[21] suggests, that current NMC LIB manufacturing systems contain all necessary processes, ambient conditions and production equipment for future cathode production e.g. of lithium sulphur batteries. However, the necessary change for the production of coming battery generations, such as all-solid-state batteries, requires far-reaching exchange of equipment. Manufacturing takes place on Gigafactory scale, ramping up remanufacturing and recycling capacities will only be feasible when significant amounts of spent LIBs are returning in 5–10 years. Flexible equipment for the automated disassembly of diverse products is not available yet.

Although LIBs and their materials represent high added value, future remanufacturing and recycling economics may decrease instead of improve. The reason being, that future product generations will be more cost effective, and that the most valuable metals such as Co and Ni are substituted in future generations, which also affects the environmental value. This requires significant efficiency improvements. Remanufacturing is currently immature and obstructed by LIBs being safety-critical therefore calling for high quality standards of remanufactured products, but

using permanent joints at the same time, making non-destructive disassembly hard. Regardless of whether remanufacturing or recycling is the goal, and due to forthcoming high product quantities, automated disassembly to module level will be necessary in the future. Unfortunately, using conventional automation approaches is not sufficient, as this would again result in an inflexible set up. This can be observed for example at Apple's automated iPhone disassembly [22] and is opposing the requirement of a changeable CF.

Current (technical) product features, make present-day LIBs not well suited for closed-loop production. As product development is still ongoing and new versatile equipment for automated disassembly is developed, there is yet an opportunity: if LIB manufacturers adhere to the identified features such as standardization, low diversity and back-/forward compatibility there is potential for LIB closed-loop production.

4 Summary and Outlook

The present paper refines the initial CF framework by developing a classification scheme for closed-loop production integration, and analyzing and classifying real world closed-loop production systems accordingly. The presented examples outline CF feasibility, and a positive impact on economic and environmental indicators. Product and factory features, which support closed-loop integration are derived, are used to analyze if the high impact case of Li-Ion batteries is a suitable scope of application for a CF. Currently, unfavourable, technical product features, and a lack of changeable (retro–)production equipment impedes CF implementation for LIBs.

Looking in to the future, closed-loop production systems, which fulfil at least partly the CF vision, will be necessary for companies, industries and countries to adhere to sustainability objectives. However, the planning of CFs systems will be complex due to the inherent complexity of such integrated systems. Thus, the developed scheme and the identified product and factory features can support initial CF conceptualization, e.g. by using the scheme as morphological model for finding viable combinations.

References

1. Duflou, J.R., et al.: Towards energy and resource efficient manufacturing: a processes and systems approach. CIRP Ann. **61**(2), 587–609 (2012)
2. Steffen, W., et al.: The trajectory of the anthropocene: the great acceleration. Anthr. Rev. **2**(1), 81–98 (2015)
3. Alting, L., Jørgensen, J.: The life cycle concept as a basis for sustainable industrial production. CIRP Ann. **42**(1), 163–167 (1993)
4. Herrmann, C.: Ganzheitliches Life Cycle Management. Springer, Berlin (2010)
5. Geyer, R., et al.: Common misconceptions about recycling. J. Ind. Ecol. **20**(5), 1010–1017 (2016)
6. Cerdas, F., Kurle, D., et al.: Defining circulation factories – a pathway towards factories of the future. Procedia CIRP **29**, 627–632 (2015)

7. Filz, M.-A., Gerberding, J., Herrmann, C., et al.: Analyzing different material supply strategies in matrix-structured manufacturing systems. Procedia CIRP **81**, 1004–1009 (2019)
8. Hammond, R., Amezquita, T., Bras, B.: Issues in automotive parts remanufacturing industry. J. Eng. Des. Autom. **4**(1), 27–46 (1998)
9. Kenné, J.-P., et al.: Production planning of a hybrid manufacturing–remanufacturing system under uncertainty within a closed-loop supply chain. Int. J. Prod. Econ. **135**(1), 81–93 (2012)
10. Vercraene, S., et al.: Coordination of manufacturing, remanufacturing and returns acceptance in hybrid manufacturing/remanufacturing systems. Int. J. Prod. Econ. **148**, 62–70 (2014)
11. Polotski, V., et al.: Production control of hybrid manufacturing–remanufacturing systems under demand and return variations. Int. J. Prod. Res. **57**(1), 100–123 (2019)
12. Tang, O., Teunter, R.: Economic lot scheduling problem with returns. Prod. Oper. Manag. **15**(4), 488–497 (2009)
13. Van Der Laan, E., Salomon, M., Dekker, R.: Investigation of lead-time effects in manufacturing/remanufacturing systems under simple PUSH and PULL control strategies. Eur. J. Oper. Res. **115**(1), 195–214 (1999)
14. Schmidt, M., Spieth, H., Bauer, J., Haubach, C.: 100 Betriebe für Ressourceneffizienz, vol. 1. Springer, Berlin (2017)
15. Steinhilper, R., Weiland, F.: Exploring new horizons for remanufacturing an up-to-date overview of industries, products and technologies. Procedia CIRP **29**, 769–773 (2015)
16. Sundin, E., Bras, B.: Making functional sales environmentally and economically beneficial through product remanufacturing. J. Clean. Prod. **13**(9), 913–925 (2005)
17. Wiendahl, H.-P., ElMaraghy, H.A., Nyhuis, P., Zäh, M.F., et al.: Changeable manufacturing – classification. design and operation. CIRP Ann. **56**(2), 783–809 (2007)
18. Hesselbach, J. et al.: Energy efficiency through optimized coordination of production and technical building services. In: Conference Proceedings LCE2008, pp. 624–628 (2008)
19. Harper, G., Sommerville, R., et al.: Recycling lithium-ion batteries from electric vehicles. Nature **575**(7781), 75–86 (2019)
20. Werner, D. et al.: Recycling chain for spent lithium-ion batteries. Metals **10**(3) (2020)
21. Bognar, N., et al.: Assessment of changeability in battery cell production systems. Procedia CIRP **72**, 695–700 (2018)
22. Rujanavech, C. et al.: Liam – An Innovation Story. Apple (2016)

Complexity-Oriented Description of Cyber-Physical Systems

A. Keuper(✉), C. Dölle, M. Riesener, and G. Schuh

Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
a.keuper@wzl.rwth-aachen.de

Abstract. In product development projects, companies try to meet more individual customer requirements by increasing the number of product variants. The product variants are no longer solely created physically, but also through different configurations of software, digital services and platforms. This extension of a physical system is called cyber-physical system (CPS). The development towards 'System of Systems' creates products, which fulfil required functions through the interaction of systems. The combination of physical and virtual objects makes cyber-physical systems significantly more complex than purely mechatronic products. The overall goal is to improve the cost-benefit ratio of cyber-physical systems. Therefore, system complexity is examined, in order to improve the design of CPS. In this paper a first step towards this goal is presented, the analysis and synthesis of description features of cyber-physical systems, which are capable to describe any CPS in a standardized logic and include all relevant dimensions that cause complexity.

Keywords: Cyber-physical systems · Complexity management · System of systems

1 Introduction

1.1 Motivation

Given the global trend towards more individual customer requirements combined with the desire to address more and more market segments, companies increase their number of product variants and expand their portfolio to the marginal areas of market demand [1]. Customers not only require individual products but also highly innovative and functional products which leads to further integration of mechanical, electronic and information-technology (IT) based components [2]. Furthermore, product variants are no longer solely created by different physical components but also by different software features and through individual communication between systems [3]. This expansion of physical products is called cyber-physical system (CPS). CPS are characterized by the fact that they connect physical objects and processes with information-processing objects and processes using IT networks [4]. This

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 602–610, 2021.
https://doi.org/10.1007/978-3-662-62138-7_60

combination results in a higher fulfilment of required functions but it also leads to significantly more complex systems. In order to capture the complexity of CPS, traditional complexity drivers like product range, component variety or interface density need to be extended. These new complexity drivers have to consider characteristics of CPSs like real-time capabilities, self-optimizing capabilities and the possibility that different parts of the system may be spread across different physical locations, hardware platforms and communication networks [1, 5, 6]. Current design approaches do not consider these aspects sufficiently. Given the increased complexity of CPS, which has not yet been adequately addressed in previous design approaches, the overall goal is to develop a design approach that enables a cost-benefit-optimized design of cyber-physical systems. To achieve this goal, the system complexity, which has a significant influence on the costs and benefits of a system, is used as a central measured and regulating variable. The first step towards the overall goal is to develop a standardized description of CPS under consideration of complexity causing dimensions.

The paper is structured as follows. After the introduction, Sect. 1 continues with a description of the relevant terminology. Section 2 analyzes related research of descriptions of CPS and models to capture the complexity of CPS. Section 3 provides an overview of the research approach. Afterwards the results of the research are presented in Sect. 4. The last section draws a conclusion and gives an impulse for further research.

1.2 Definitions

In this section, a short definition of the terms “cyber-physical system” and “system complexity” is given to establish a common understanding.

Cyber-physical system. Cyber-physical systems combine systems that interact with other systems in a limited environment, so-called “embedded systems”, with the ubiquitous networking of objects through global IT networks, often referred to as the “Internet of Things” [3]. While purely mechatronic systems focus on the combination of mechanical systems and its embedded control systems in a fixed centralized structure [7]. CPS can be considered as the next evolutionary step as they combine elements of virtual and physical dimensions to a network of decentralized adaptive systems [7]. A mechatronic system can thereby be a subsystem of a cyber-physical system. CPS are an important driver of progress towards “Industrie 4.0”, since they allow higher degrees of production process automation and increase the exchange of digital information [8]. Individual machines communicate with each other in real time enabling a constant flow of information about orders, production rates, materials and other resources [9]. The increasing spread of CPS increases their share of the expenses in companies. An optimized design of CPS can help to keep the costs low and increase the benefits. Complexity plays a significant role in such systems, which is why an introduction to system complexity is given below.

System complexity. The complexity of a system in general depends on the variety and number of system elements as well as the variety and number of connections between these elements [10]. In addition, the complexity of a system is determined by the number of different system states that the system can have over time [11]. Vemuri

specifies the concept of the complex system with the following four characteristic properties [12]:

- Attributes necessary to describe the system are so numerous that a complete or exact description is not possible
- Relations in these systems are of a statistical nature
- Complex systems are not static, but change over time
- The decision-making and action behavior is characterized by contradictory goals

Schuh and Riesener summarize the complexity of a system by the number and variety of elements and relation that occur in this system as well as the variability over time, which itself is determined by the variety of possible states of the elements and relations between elements [1].

After defining the most relevant terms the next section of the paper will analyze related work in the field of complexity-oriented description of CPS.

2 Related Work

The analysis of related work has shown that there is numerous scientific work describing features of CPS, each dealing with different questions. An important differentiating feature is, besides the concrete area of consideration, the level of abstraction of the observation. In the context of this paper, those approaches are particularly relevant that deal with overarching questions regarding CPS as a whole and the interactions of the individual elements.

"Model for the description of cyber-physical systems" by Döbrich and Heidel (2013) [4]: In this approach a system is represented through assets and every asset is represented through a list of property types without specific values assigned to it. The method is called property principle and it classifies property types into the five basic elements business, construction, function, location, and performance. Even though this approach is able to describe a whole system in a detailed manner it is insufficient in mapping the dynamic behavior of CPS and the interaction with other systems.

"Review on Cyber-physical Systems" by Liu et al. (2017) [13]: The work of Liu et al. analyzes the current state of research in the field of CPS. Among other topics, the focus is on the architecture of CPS, which consists of six main elements within three layers. The three layers are the user layer, the information layer and the physical system layer. The six elements include sensor networks, next generation network systems, data center, control center, actuator networks, and system user. Every of these elements is further described by its functions and characteristics. The presented architecture addresses many complexity driving aspects like system interactions or real-time capabilities. However, it only presents an ideal solution of a CPS architecture and does not include a method to represent an actual CPS within the logic of that architecture.

Only very few publications specifically address the dimensions of complexity mentioned in Sect. 1.2. Two examples are presented hereafter:

"Structural complexity and its implications for design of cyber-physical systems" by Sinha (2014) [14]: Sinha measures complexity of cyber-physical systems with a mathematical model consisting of the three factors C1, C2 and C3. The factor C1 evaluates the complexity of each individual component of the system or product. C2 describes the number and complexity of the pairwise interactions between the components. The total complexity of the system topology is represented by C3. The factors are calculated on the basis of algorithms used for design structure matrices. In principle, the formula can be applied to all elements and relations in CPS. The specific properties of CPS and representation of CPS in a standardized description are not addressed.

"Complexity Challenges in Development of Cyber-Physical Systems" by Törngren and Sellgren (2018) [15]: The goal of this work is to investigate the suitability of different development approaches for the development of complex CPS. In order to do so, a detailed description of facets of complexity in CPS is given (for example heterogeneity or uncertainty and change). In addition, there is an in-depth analysis of the interrelations of components of CPS. This work shows the different forms of complexity in CPS very well and describes them by the aid of examples. However, a standardized description for the representation of different CPS is not created.

The four exemplary approaches presented above as well as further analyzed approaches show, that the focus lies either on the detailed description of CPS or on the investigation of complexity in CPS. An approach, which examines both in depth was not identified. Therefore, this paper will combine both and derive characteristics of CPS, which have to be considered when assessing its complexity.

3 Research Approach

The goal of this paper is to derive description features of CPS with particular attention to the system's complexity. According to Kubicek [16], it is recommended to define a core research question for a target-oriented research. Based on the challenges presented, the core research question is the following:s

"How can cyber-physical systems be systematically described under consideration of the complexity causing dimensions?"

Requirements for the description features are derived from the research question and the overall objectives. The first requirement is that the description features must be able to represent every actual CPS, which has the consequence that a holistic approach must be chosen. This must be combined with the objective that the dimensions causing complexity must be adequately represented. According to systems theory, a model is defined as a simplified and abstracted representation of a real situation, whose primary goal is the prediction of the behavior of the real system [10]. In order to develop a model as an abstract representation of reality, it is essential to determine a suitable degree of abstraction in which the essential elements and relationships can be described [17]. In addition, the structure and level of detail of the description features should be suitable for a subsequent quantification of the complexity of CPS.

In order to fulfil these requirements and to ensure the widest possible scope of analysis a combination of bottom-up and top-down approach is used. The main steps of the approach are presented in Fig. 1.

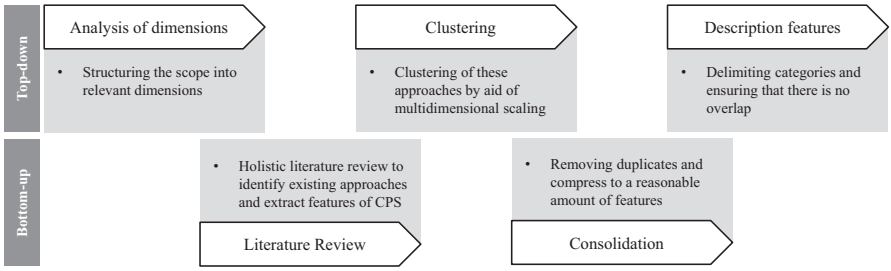


Fig. 1. Research approach of this paper

In the first step the scope will be structured, afterwards a systematic literature review is conducted. The results of the literature review are clustered by aid of multidimensional scaling. In the next step, the features which are extracted from literature are analyzed, filtered and consolidated. In the last step, the features are combined to form the final set of description features for CPS. After presenting the research approach, the next section presents the results of every step as well as the overall result.

4 Derivation of the Description Features

4.1 Dimensions of Cyber-Physical Systems and System Complexity

The essential characteristic of cyber-physical systems is the connection between the virtual (cyber) world and the physical world [4]. For a holistic view of all properties of a CPS it is necessary to examine the elements of the physical and virtual dimensions in detail as well as the relationship of the elements of both dimensions to each other. Additionally CPS are considered as socio-technical systems and include dimensions of social interaction involving users, customers and manufacturers [15]. Therefore, the complexity of a CPS can be understood as the combination of the complexity of the physical components, the virtual elements and the interconnectivity within the system itself as well as the connectivity with the environment of the CPS. These aspects will be considered by aid of the three dimensions: physical world, virtual world and the connectivity of the entire system.

According to Meier, as already presented in Sect. 1, the complexity of a system depends on the number and variety of elements, the relationships between the elements and the number of possible system states [11]. Reiss distinguishes four main factors of complexity: multiplicity, variety, ambiguity and changeability [18]. Schuh and Riesener analyze multiple sources in regard to system complexity and summarize them in the two dimensions of diversity and dynamics [1]. For the purpose of this paper the dimensions of complexity which are considered are number and variety of elements, (inter-)dependency and dynamics.

In the next step, a literature review is conducted to identify existing descriptions of CPS.

4.2 Existing descriptions of cyber-physical systems

In order to identify relevant literature and existing descriptions, a keyword search was used as a literature review technique. To identify a higher amount of relevant literature different keywords and synonyms were applied. The investigation covers various sources such as Google Scholar, SpringerLink or Scopus. The search results are filtered and narrowed down to the most relevant approaches. A similar procedure has been used before in engineering research, for example in [19, 20].

Figure 2 shows several models and descriptions of CPS and their evaluation in regard to the previously defined dimensions of CPS and system complexity.

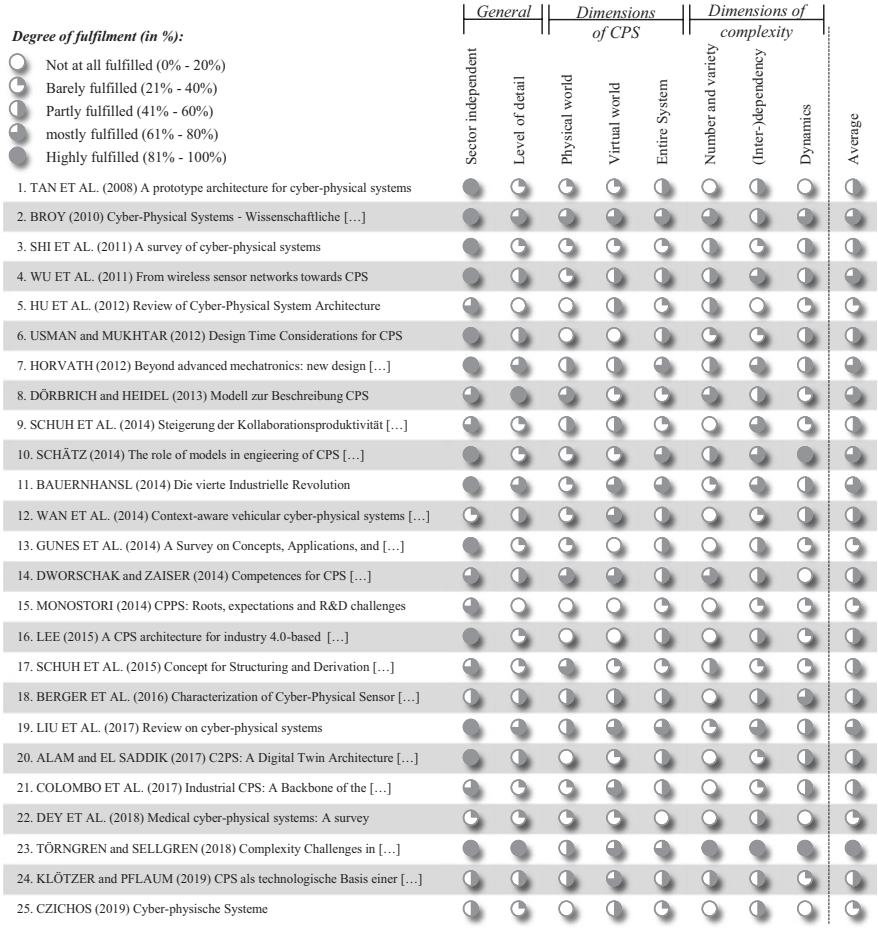


Fig. 2. Overview of approaches that describe CPS or present a description model

To be better able to interpret the evaluation results of the descriptions, they are visualized and clustered in the next section using multidimensional scaling.

4.3 Multidimensional Scaling of Existing Descriptions and Clustering of Features

Multidimensional scaling is a commonly used tool for dimension reduction. Pairwise distances are calculated for the high-dimensional data and a two-dimensional representation reflects the higher dimensional distances as good as possible in the two-dimensional space [21]. Figure 3 shows the result of the multidimensional scaling as well as three exemplary clusters of similar evaluated models. An ideal model with 100% fulfilment rate of every evaluation dimension is also included in the figure.

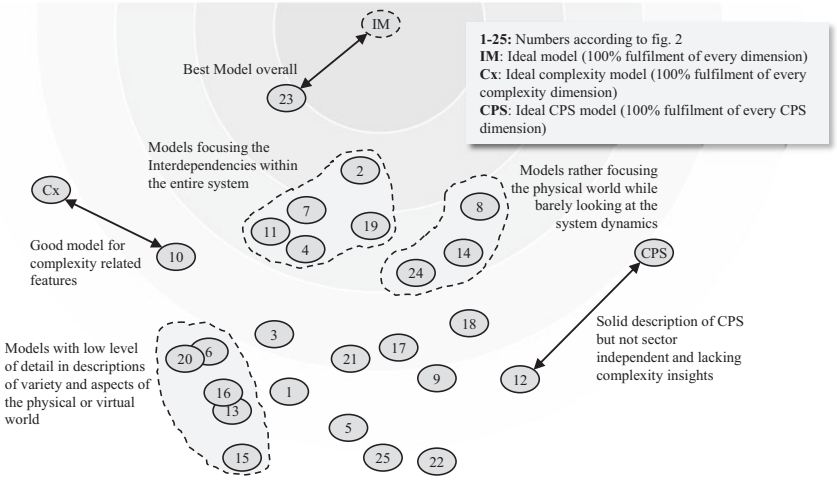


Fig. 3. Result of multidimensional scaling and insights

The distances between the existing approaches and the ideal model are used as an indicator of the relevance of the features derived from each existing approach.

4.4 Consolidation and Deduction of the Complexity-Oriented Description

In every investigated approach, a search for description of features of CPS is conducted. All found features are collected resulting in a list of 97 features of CPS. Duplicates and similar features are combined into one. The remaining features are prioritized and reduced to those with the highest relevance according to the distance of their source to the ideal model (Sect. 4.3). The result is shown in Fig. 4 structured according to the three complexity dimensions defined in Sect. 4.1.

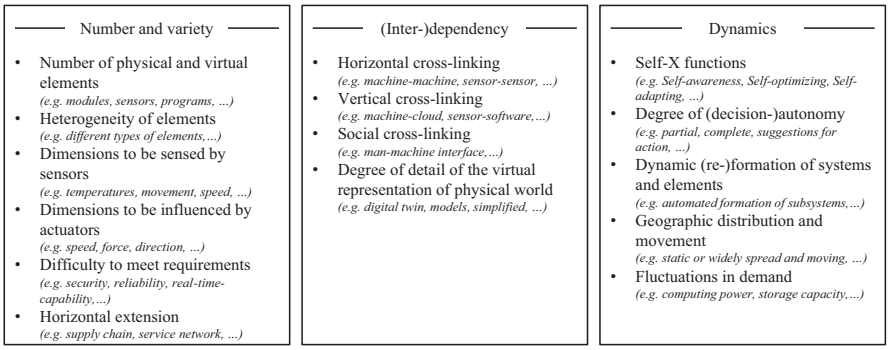


Fig. 4. Features for the complexity-oriented description of cyber-physical systems

Since current design approaches do not consider the complexity of CPS sufficiently, the previously derived description features will be the basis for the evaluation of the systems complexity within CPS. To do this, characteristic values will be assigned to these features and enable a standardized description of CPS. Further next steps and the conclusion are shown in the following section.

5 Conclusion and Further Research

The combination of embedded systems and the "Internet of Things" lead to the concept of CPS. Even though CPS are very complex systems there is no approach addressing the complexity-oriented design of these systems. Therefore, this paper presents complexity-oriented description features of CPS in order to describe them in a standardized logic under consideration of complexity causing dimensions. To identify relevant features that significantly influence the complexity of CPS, a systematic literature analysis was conducted. The identified approaches were evaluated with respect to their broadly transferable description of CPS and their suitability for quantifying complexity. The most suitable features were combined to a set of description features.

The validation and further research on the steering of complexity in CPS is currently part of the authors' research activities in the department of Innovation Management at the Laboratory for Machine Tools and Production Engineering WZL. Upcoming research will be focused on the quantification of complexity in CPS and the derivation of design recommendations.

Acknowledgement. The presented results have been developed within the research project "Complexity oriented design of cyber-physical systems" [GZ: SCHU 1495/146-1] funded by the Deutsche Forschungsgemeinschaft (DFG).

References

1. Schuh, G., Riesener, M.: Produktkomplexität managen. Strategien - Methoden - Tools, 3rd ed. Hanser, München (2017)
2. Schuh, G., Rudolf, S., Riesener, M., Kantelberg, J.: Application of highly-iterative product development in automotive and manufacturing industry. In: ISPIM Innovation Symposium, p. 1 (2016)
3. Schuh, G., Potente, T., Thomas, C., Hauptvogel, A.: Steigerung der Kollaborationsproduktivität durch cyber-physische Systeme. In: Industrie 4.0 in Produktion, Automatisierung und Logistik, pp. 277–295. Springer (2014)
4. Döbrich, U., Heidel, R.: Modell zur Beschreibung cyber-physischer Systeme. atp edition-Automatisierungstechnische Praxis **55**, 38–45 (2013)
5. Schätz, B.: The role of models in engineering of cyber-physical systems-challenges and possibilities. In: Tagungsband des Dagstuhl-Workshops, p. 91 (2014)
6. Akkaya, I., Derler, P., Emoto, S., Lee, E.A.: Systems engineering for industrial cyber–physical systems using aspects. Proc. IEEE **104**, 997–1012 (2016)
7. Plateaux, R., Penas, O., Choley, J.-Y., Mhenni, F., Hammadi, M., Louni, F.: Evolution from mechatronics to cyber physical systems: an educational point of view. In: 2016 11th France-Japan & 9th Europe-Asia Congress on Mechatronics (MECATRONICS)/17th International Conference on Research and Education in Mechatronics (REM). IEEE, pp. 360–366 (2016)
8. Zehbold, C.: Product Lifecycle Management (PLM) im Kontext von Industrie 4.0. In: Fend, L., Hofmann, J. (eds.) Digitalisierung in Industrie-, Handels- und Dienstleistungsunternehmen: Konzepte - Lösungen - Beispiele, pp. 69–89. Springer Fachmedien Wiesbaden, Wiesbaden (2018)
9. Schuh, G., Pitsch, M., Rudolf, S., Karmann, W., Sommer, M.: Modular sensor platform for service-oriented cyber-physical systems in the European tool making industry. Procedia CIRP **17**, 374–379 (2014)
10. Patzak, G.: Systemtechnik – Planung komplexer innovativer Systeme. Grundlagen, Methoden, Techniken, Springer (1982)
11. Meier, J.: Produktarchitekturtypen globalisierter Unternehmen, Shaker (2007)
12. Vemuri, V.: Modelling of Complex Systems. Academic Press, New York (1978)
13. Liu, Y., Peng, Y., Wang, B., Yao, S., Liu, Z.: Review on cyber-physical systems. IEEE/CAA J. Autom. Sin. **4**, 27–40 (2017)
14. Sinha, K., others: Structural complexity and its implications for design of cyber-physical systems. Dissertation (2014)
15. Törngren, M., Sellgren, U.: Complexity challenges in development of cyber-physical systems. In: Lohstroh, M., Derler, P., Sirjani, M. (eds.) Principles of Modeling: Essays Dedicated to Edward A. Lee, pp. 478–503. Springer International Publishing, Cham (2018)
16. Kubicek, H.: Heuristische Bezugsrahmen. Poeschel, Stuttgart (1977)
17. Stachowiak, H.: Allgemeine Modelltheorie. Springer, Wien (1973)
18. Reiß, M.: Komplexitätsmanagement. T. 1. Universität Stuttgart, Stuttgart (1993)
19. Rebentisch, E., Schuh, G., Dölle, C., Mattern, C., Abel, H.: Defining agile culture using topic modelling. In: Proceedings of the DESIGN 2018 15th International Design Conference, pp. 2111–2120. University of Zagreb, Croatia; The Design Society, Glasgow, UK (2018)
20. Riesener, M., Doelle, C., Mendl-Heinisch, M., Schuh, G.: Literature based derivation of a framework to evaluate engineering change requests. In: 2019 Portland International Conference on Management of Engineering and Technology (PICMET), pp. 1–6. IEEE (2019)
21. Buja, A., Swayne, D. F., Littman, M.L., Dean, N., Hofmann, H., Chen, L.: Data visualization with multidimensional scaling. J. Comput. Graph. Stat. **17**, 444–472 (2008)

Adapted Process Model for Manufacturing Within Production Networks

M. Reimche[1(✉)], S. Berghof[2], and J. P. Bergmann[1]

[1] Technische Universität Ilmenau, 98693 Ilmenau, Germany
maxim.reimche@tu-ilmenau.de
[2] Berghof Group GmbH, 07426 Königsee, Germany

Abstract. The paper concentrates on the external elimination of bottlenecks by using congruent shared capacities within a non-hierarchical organized network. The study assumes that the capacity of assembly can be adapted by increasing the resources of the company. Conversely, the capacity of production is restricted by technological requirements and regulated by machine resources.

To relieve capacity by external measures within a non-hierarchical organized network, an adapted process model is required. For this purpose, critical issues for the process flow within a network are identified and necessary regulations suggested.

The paper presents the possibility of integrating the adapted approach into the operations management process as well as into systems of production planning and control and deduces the necessary information that is required to ensure the economic viability of bottleneck removal.

Keywords: Capacity sharing · Non-hierarchical organized production network · Process model

1 Introduction

In comparison with west Germany states, a higher number of small and medium-sized enterprises (SME) in the manufacturing sector distinguishes the in eastern Germany states [1]. In Thuringia, component suppliers and contract manufacturers mainly characterize the manufacturing industry. For example, the number of contract manufacturers in Thuringia is about 1.5 higher than the national average [2].

The efficient management of fluctuations in customer demand is a major challenge for the supply industry. It often reacts by outsourcing to contract manufacturers at short term, which makes harmonized long-term planning very difficult for them. Companies that offer contract manufacturing in addition to their own product portfolio are required to use suitable solutions in order to guarantee high delivery reliability to all their customers [3].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 611–620, 2021.
https://doi.org/10.1007/978-3-662-62138-7_61

Cross-company collaboration is an important part of the value creation strategy of today's companies. It can be found in various forms in development and procurement, also in logistics and production. The collaboration is usually motivated by the saving of time and financial resources and enables companies to react to market or customer requirements with increased flexibility [4].

In the field of production, the collaboration is used to supplement own production resources with complementary technologies, to reduce and relocate activities as well as for a temporary, i.e. order- or product-related, use of congruent resources. A common form of order-related collaboration is cooperation within a network. Such forms of collaboration simplify the compilation of individual cooperation partners who have previously overcome an entry barrier for participation in such networks. Such barriers may concern, for example, the quality to be provided, the fulfilment of standards or requirements for information technology.

A particular challenge is the selection of a suitable organizational form and a suitable control mechanism for such cooperation. Due to the short reaction times required by the customer, companies that procure congruent and complementary resources on an order-related basis are dependent on permanent information availability and low-cost resource initiation.

Current processes in planning systems do not completely consider all of the requirements that SMEs have to fulfill to produce in cooperation with each other. This article systematizes these requirements and transfers them into a proposal for the process model.

2 State of the Art

The advantage of including external resources caused the formation of different organizational types of networking. The scientific results show that a partial sharing of capacity or demand within a network is advantageous. The existing process models simplify the decision process to share resources and demand from economics site and are presented by [5–7]. Furthermore methods to minimize the costs of cooperation [8], to accelerate capacity or demand sharing [9] or increase fairness in the sharing processes [10] have already been developed. In addition, models used the game approach and cloud technology for sharing of capacity have been investigated [11]. All listed works were done without focusing special requirements of an industry sector type.

Vertical cooperation within a network is characterized by a clear differentiation of value-added stages among these participating companies. The participants of a vertical cooperation resemble a customer-supplier relationship [12]. A horizontal cooperation is characterized by congruence of the service offered. The companies are on the same level of the value chain and aim to strengthen their competitive position [13]. (e.g. short delivery time with a high delivery reliability)

Focusing on horizontal cooperation, this can take place in a hierarchically organized network and in a network organized according to heterarchic principles [14]. While a central organization takes over the allocation and distribution of individual

orders or work processes in hierarchically organized structures, heterarchically organized networks make use of free available capacities, which can be found, for example, on stock exchanges [15] or platforms [16–18]. This form of cooperation implies a limited disclosure of company workloads to all network participants or an allocation of tasks by a central organization and can lead to a reluctance to participate in networks.

A look at automation technology shows which forms of organization are theoretically feasible for the separate participants (devices) in a network. In the network topology, a basic differentiation is made between ring, star, line, bus, tree and mesh structures as shown in Fig. 1.

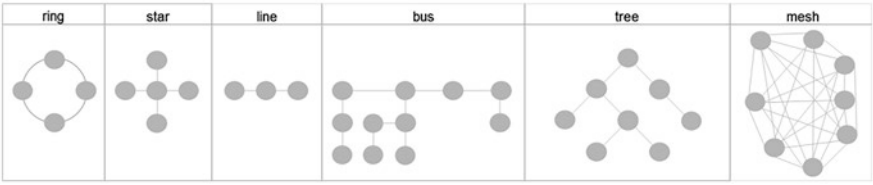


Fig. 1. Topology of a network

A bus structure requires that all participants communicate via a common line, only one participant can transmit at the same time and requires an access procedure to the bus network. In a ring structure, the participant has direct contact with only two of its direct neighbors. For communication with indirect neighbors, the information has to be passed on. The redundant design of the ring enables communication in both directions. A line structure is a special form of ring structure in which two participants have only one direct contact. A star structure is characterized by a central station. While a passive central station is mainly responsible for routing, an active central station first processes the information. The failure of an active central station causes the inability to work of the entire network. An initial node characterizes a tree structure as a further development of the line structure, from which individual edges lead to further accounts or participants. This type is characterized primarily by its hierarchical structure [19].

In comparison to the structures listed above, a mesh structure represents the simplest, least error-prone and fastest communication way, because by rerouting to a central point, the transfer of information and access regulations are omitted. Each participant has direct contact to all network participants and is not subject to a hierarchical structure [20].

Transferred to the network typology, peer-to-peer networks represent a mesh form of organization that is suitable for vertical and horizontal cooperation. They are organized according to a principle by which all nodes (participants) act independently, have a connection to all network peers and maintain their business relationships as their individual responsibility. While a traditional division of labor is based on fixed roles, the peer-to-peer network works on the principle of self-assigned tasks. In contrast to the other forms of organization, the peers act without approval and are independent of a central authority. The accent is strongly on the distribution of tasks on a voluntary

basis, which allows for independent control of individual production performance and self-aggregation [21].

The principle of situational, local formation of relationships requires a real-time and reliable information base. An automated exchange between the companies' planning systems is an important condition for the required speed of reaction [22]. A lack of interoperability of the planning systems used by the network participants leads to the necessity for a central controlling instance or the use of communication middleware.

A middleware, which extends the Aachen PPS model, acts as a connecting element of the planning system to the external world, e.g. to calculate an optimal route for the orders in a network [23]. Another approach for linking planning systems and a fast exchange of information shows its connection to a cloud-based virtual marketplace [24].

In this paper a process model is proposed that allows interoperability of systems in horizontal cooperation as well as fast interaction and does not require disclosure of the capacity utilization. The process model is created for network types applying a peer-to-peer structure as a basis for the communication. In addition, the regional characteristics mentioned above, such as a high number of contract manufacturers, who normally compete with each other, are taken into account. It is based on the example of the metalworking industry.

3 Development of the Process Model

A process model brings together all activities and their dependencies in an orderly sequence. Different techniques are used to create it. The event-driven process chain (EPC) is a possible modelling technique, which is often used to describe business processes and allows a simple representation [25].

In the first step of this technique, the requirements for the process of capacity sharing are compiled, which are set by SMEs in the metalworking industry in the context of a number of workshops. Then the requirements are transferred into a proposal for a process, which should enable their implementation.

3.1 Compilation of Requirements

SMEs in the metalworking sector act on the market as manufacturers of series products or as contract manufacturers. SMEs are organized into departments (e.g. mechanical processing, surface finishing, assembly) that are often functionally separated from each other. In these areas, the order flows of different business fields cross each other, which leads to a high complexity in planning. Independently of the business field, customers expect the shortest delivery times and high delivery reliability.

Workshop production is a common process principle according to which the mechanical processing of these companies is organized. A nonhomogeneous composition of resources regarding their technological characteristics is a frequent consequence of the step-by-step purchase of new machines, e.g. milling centers. For economic reasons, such as reducing machining time, additional work plans and

related CNC programs for the resources purchased are created. The creation was also completed for some articles for which work plans and CNC programs were already available.

Independently of the number of work plans of the article, in most cases there are about one to three work plan positions that use the capacity of a resource in workshop production. In 90% of cases, the maximum duration is eight hours (including set-up and processing time). The duration is the result of analysis of circa 11000 operations that were carried out in the period June to December 2019 in three regional companies in the area of mechanical processing. The evaluation was based on reported order data. According to the initial situation mentioned above, the following requirements can be determined:

1. Companies aim to keep the level of capacity utilization high in order to spread the fixed costs of the machines over as many cost units as possible. Generally, the aim of this procedure is to fill the available capacities with their own and any possible orders from network partners in order to avoid bringing forward orders from the future. The first requirement is that any spare capacity that cannot be filled without orders from the future, should first be announced to the network partners. The announcement of the relevant capacity should (a) be made by specifying the time period and the amount of time available, (b) be kept up-to-date and (c) be capable of immediate elimination in the case of use by an order. Regarding the duration of operations, the notification may concern a capacity gap in the next three working days at the earliest, in order to take into account two days for transport.
2. Depending on the customer delivery date and the throughput time of the product, its realization may involve units utilized to differing degrees in the throughput process (assembly workstations or processing machines). For SMEs in the metalworking sector, mechanical processing represents an initial value-adding process. Delay of the internal delivery date often leads to a delay of the customer delivery date or to a need to reschedule subsequent processes. Compared to large companies, SMEs do not have the possibility to produce future orders within their own organization, e.g. at a different location, and eliminate bottlenecks in this way. Consequently, it is important to avoid the delay from the beginning by identifying overloaded machines. Orders that are assigned to these machines are subject to an analysis process. This initially includes a feasibility check to determine whether it is possible to move the order to an alternative machine, i.e. a machine that has at least the same or higher technological features. At this point, higher technological features mean, for example, higher precision, larger working space or a higher number of machining axes. Another possibility to increase free capacity is the splitting of the working position for processing on several machines. If the bottleneck cannot be resolved by internal optimizations or an increase of working time, a simulation can be done of the available network capacity. The second requirement is that the capacity used for the time of the simulation must be eliminated from the total available network capacity and must be clearly pre-reserved for the user company.
3. The simulation as well as the maintenance of available capacity should be done outside the planning system and the data used should represent the real order situation as precisely as possible. SMEs use production simulation solutions for

planning their resources less often than large companies. One of the reasons for this is the self-dependent transfer of existing planning logic to the simulation. Often the time and the necessary knowledge are not available, so the advantages of simulations are neglected [26]. The third requirement is that the simulation has to be carried out based on the digital shadow of the systems. A digital shadow is a sufficiently precise representation of all relevant processes and data of the planning system of a company [27, 28].

4. The information disclosed on available capacity should not be used for the purpose of providing economic or other benefits. If a request for a capacity quantity takes place in the network, this request should be able to show a clear reference to the capacity requirement item. The requirement item can be a single operation of an order or a sum of several operations of different orders that have to be produced in the same time period on the same machine. The fourth requirement is that the existence of a control mechanism at each partner that only allows a new capacity request from the same source of requirement items after the previous capacity request has been rejected.
5. Binding use of the capacity occurs in accordance with each company's own procedure for outsourcing. The fifth requirement is to link the information from the capacity purchase process with the external outsourcing operation.
6. The capacities offered in the network are generally used to cover fixed costs and to purchase capacity in order to ensure on-time delivery. SMEs often act as suppliers to large companies and are located at the end of the value chain. Large companies tend to pass on the demand fluctuations to their suppliers and force their suppliers to react appropriately. The adjustment of the price of a capacity request for the same item after its previous rejection can lead to a dishonest exploitation of the price mechanism. The sixth requirement is - to link the requested price information to the requirement item and block the change.

3.2 Development of a model for cooperation in networks

The process model represents the part of order processing that potentially belongs to the inter-company order processing. From a time perspective, the focus is primarily on order processing in the short and medium term. In order to simplify matters, the model does not consider the loss of the orders that cause demand for capacity, nor the machine failure of the company offering the capacity. For a better understanding, the principle of communication between the partners in the network is explained here.

Figure 2 shows a schematic structure of the communication. It can be seen that each company has the component "Controller Module", which connects it with the network partners according to a peer-to-peer structure mentioned above. The component receives and processes the information messages from the network, forwards them to the "Simulation Module" if necessary and takes over the control function. The component "Simulation Module" is used for capacity calculation. It receives the necessary information (e.g. transaction data) from the "Digital Shadow" component and the information about additional resources from "Controller Module". In addition, the component is used to transfer free capacities to the "Controller Module". The "Digital Shadow" component functions as a database for selected ERP information.

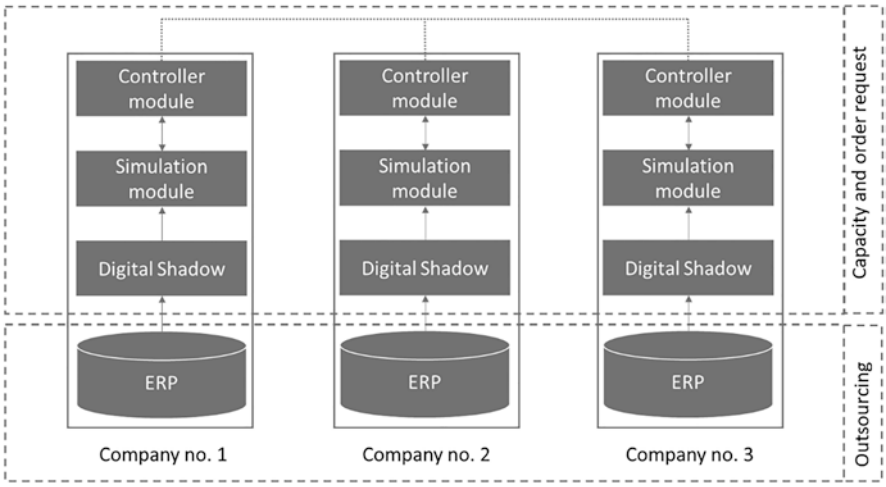


Fig. 2. Communication principle within the network

The process starts with the sub-process of capacity unblocking at the capacity provider. (Fig. 3) Depending on its internal strategy (percentage, staged or complete release), the free capacities for machines that are known in the network are transferred to the "Controller Module" by the "Simulation Module". If the "Controller Module" registers a message with a capacity request from the network, it checks in the matching process for the availability of the resource in the requested period by using technological criteria to allocate it to its own resources providing these are technologically equal or superior. If it is a request with a split quantity (e.g. 100 min, 300 min, 450 min), the maximum combination will be determined. If the capacity is not available, the "Controller Module" sends a rejection with assigned ID of the request. If the capacity is available, the requested quantity or the maximum combination is signed in to the "Controller Module" as pre–reservation (provisional reservation) and the information about pre-reservation is sent in message form with an ID assignment. The available capacity in the "Controller Module" is reduced by pre-reservation quantity. If the capacity offering company needs the signed quantity for a newly arrived customer order, the capacity announced in the network is reduced irrespective of the pre-reservation. The pre-reservation has validity. The determination of its duration will not be discussed at this point. If no order request has been received before the end of the validity period, the pre-reservation will be cancelled with correction of the capacity quantity. If the "Controller Module" registers an order request, the duration of the pre-reservation will be extended until the decision of acceptance or rejection. The verification process is an independent process outside the "Controller Module" and, in case of a positive decision, includes the blocking of the capacity (e.g. the creation of a pseudo order with the associated work plan and a clear reference to the request) in the ERP system. In case of rejection (e.g. for technological or economic reasons) a message is sent with the ID and the decision, pre-reservation is cancelled and the available quantity corrected. If the order request is accepted, the binding

transmission of the message of acceptance takes place after a final reduction of the available capacity in the “Controller Module”. The sending of a rejection with a price proposal is not possible in the “Controller Module”.

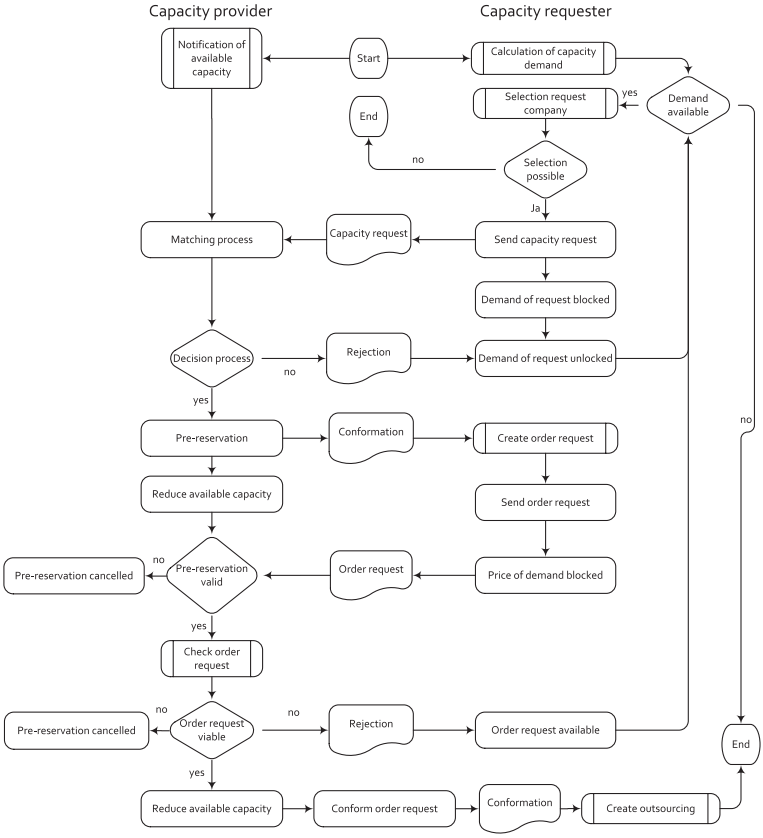


Fig. 3. Process model for production within the network

On the side of the capacity requester, the process begins with the sub-process of calculating capacity requirements. This takes place in the “Simulation Module” and follows the company’s own strategies. As mentioned above, companies in the industry often have several work plans for the same part number. If a capacity requirement is determined for an article number, a capacity request is made for the machine with the weakest technology, which is the article number that has been authorized. The selection of the network partner, as a further sub-process, takes place in the “Controller Module” and follows different rules (e.g. the article number has already been outsourced to a network partner, number of requests and rejections of order request to a network partner, required certification, etc.) When a network partner is selected, the message with the demand is transmitted and the source of the demand is blocked for further capacity requests. A rejection unlocks the demand item and capacity request

can be transmitted to another network partner. In case of a confirmation, an order request can be sent for each individual requirement item within the extension of the pre-reservation. In addition to the transmission of the information for checking the feasibility (e.g. drawing, quantity, etc.), the specification of the maximum price P_{max} including information for its assessment (e.g. delivery conditions, material provision, packaging requirements) is explicitly specified. The communicated P_{max} price may vary depending on basic situation in the company and the timeline of ordering and is currently a topic of the on-going project.

The production costs of the operation include all fixed and variable costs incurred for the execution of the operation. If accepted, a communicated order request is binding for the capacity buyer and requires the process of outsourcing. The communicated price is fixed for each user and cannot be changed. If the order request is rejected or not answered within a specified period, the item is unlocked for further capacity requests.

4 Conclusion

This article presents a process model for production order processing for SMEs in the metalworking industry, including free capacity from peer-to-peer production networks. In the first step, the essential relevant requirements for the capacity planning process were defined, which were elaborated in workshops with representatives of the target industries. Following this, a process model was presented, which enables a quick exchange of information about the free resources of the participants in a network and prevents the exploitation of this additional information for extraneous economic purposes.

Future investigations will focus on the procedure for the calculation of the price that is the maximum possible for the contracting company and profitable to accept. These studies will be used to determine whether and how the procedure for price determination changes over time and what parameter influences this change.

Acknowledgement. The authors sincerely thank Thüringer Aufbaubank and Thuringian Ministry for Economic Affairs, Science and Digital Society for supporting the research project "AgiPro – Agile contribution margin oriented production within networked small and medium sized enterprises" (2018 VF 0020).

References

1. Schiersch, A., Mattes, A., Fritsch, M., Eickelpasch, A., Arnold, M.: Kleinteiligkeit der ostdeutschen Wirtschaft, p. 12. DIW-ECON, Berlin (2015)
2. Thüringer Mittelstandsbericht, pp. 79–80 (2015)
3. Freitag, M., Veigt, M., Schukraft, S.: Adaptive Produktionsplanung und -steuerung. ZWF **112** (2017)
4. Schlick, C., Killich, S.: Identifikation von Kooperationspotenzialen in bestehenden Netzwerken. In: Wojda, F., Barth, A.: Innovative Kooperationsnetzwerke, pp. 159–162. Deutscher Universitäts-Verlag, Wiesbaden (2006)

5. Chen, J.-C., Wang, K.-J., Wang, S.-M., Yang S.M.: Price negotiation for capacity sharing in a two-factory environment using genetic algorithm. Prod. Res. **46**, 1847–1868 (2008)
6. Yoon, S.W., Nof, S.Y.: Demand and capacity sharing decisions and protocols in a collaborative network of enterprises. Decis. Support Syst. **49**, 442–450 (2010)
7. Seok, H., Nof, S.Y.: Collaborative capacity sharing among manufacturers on the same supply network horizontal layer for sustainable and balanced returns. Prod. Res. **52**, 1622–1643 (2014)
8. Moghaddam, M., Nof, S.Y.: Combined demand and capacity sharing with best matching decisions in enterprise collaboration. Prod. Econ. **148**, 93–109 (2014)
9. Moghaddam, M., Nof, S.Y.: Real-time optimization and control mechanisms for collaborative demand and capacity sharing. Prod. Econ. **171**, 495–506 (2016)
10. Yilmaz, I., Yoon, S.W.: Dynamic-distributed decisions and sharing protocol for fair resource sharing in collaborative network. Prod. Econ. **226** (2020). https://doi.org/10.1016/j.ijpe.2020.107644
11. Argoneto, P., Renna, P.: Supporting capacity sharing in the cloud manufacturing environment based on game theory and fuzzy logic. Enterp. Inf. Syst. **10**(2), 193-210 (2014)
12. Picot, A., Reichwald, R., Wigand, R.: Die grenzenlose Unternehmung, p. 281. Springer, Wiesbaden (1996)
13. Hungenberg, H.: Bildung und Entwicklung von strategischen Allianzen. In: Engelhard, J., Sinz, E.J. (eds.) Kooperation im Wettbewerb, pp. 6. Gabler, Bamberg (1999)
14. Schiffenbacher, K.U.: Konfiguration virtueller Wertschöpfungsketten in dynamischen, heterarchischen Kompetenznetzwerken, pp. 36–37. Utz, München (2000)
15. Uygun, Y.: Integrierte Kapazitätsbörse – Entwicklung eines Instrumentariums für den Handel mit Maschinenkapazitäten in regional-lateralen Unternehmensnetzen. Praxiswissen Service Verlag, Dortmund (2012)
16. Meiser, S.: Entwicklung eines Kollaborationsnetzwerkes zur Optimierung von Produktionskapazitäten durch Organisation dynamischer Kooperationen, mit prototypischer Umsetzung für die Druckindustrie. Universitätsbibliothek Wuppertal, Wuppertal (2013)
17. Stonis, M., Fritzsch, B., Schüler, F., Müller, M.: Sharing von Produktionskapazitäten mittels digitaler Plattformen. ZWF **115**, 257–261 (2020)
18. http://www.kreatize.com/de, http://www.v-industry.com. Accessed 6 May 2020
19. Schnell, G., Wiedemann, B.: Bussysteme in der Automatisierungs- und Prozesstechnik, pp. 1–8. Springer, Wiesbaden (2019)
20. Bunke, M.: Einführung und Überblick Baumtransformationen, pp. 1–2. http://www.informatik.uni-bremen.de/st/lehre/Arte-fakt/Seminar/Ausarbeitungen/02_Einfuehrung_Baumtransformation.pdf. Accessed 7 May 2020
21. Bauwens, M.: Class and capital in peer production. Cap. Cl. **33**(1), 121–141 (2009)
22. Schuh, G., Stich, V.: Produktionsplanung und -steuerung 2, pp. 67–69, Springer, Berlin (2012)
23. Ellwein, C., Elser, A.: Vernetzte Produktionsplanung. ZWF **114**, 807-810 (2019)
24. Schmitt, R., Ellerich, M., Groggert, S.: Kollaborative Produktionsplanung und Steuerung in der Cloud. ZWF **110**, 399-402 (2015)
25. Becker, J., Probandt, W., Vering, O.: Grundsätze ordnungsmäßiger Modellierung, p. 15. Springer, Berlin (2012)
26. Schulz, A., Züfle, E., Sommer, L., Haug, Manuel.: Simulation in der operativen Produktionsplanung - Erfolgsfaktoren für KMU. ZWF **102**, 35 (2007)
27. Linnartz, M., Fischer, M., Pause, D.: Assistenzsysteme zur Entscheidungsunterstützung in der Supply Chain. ZWF **114**, 592–595 (2019)
28. Bauernhansl, T., Krüger, J., Reinhart, G., Schuh, G.: WGP-Standpunkt Industrie 4.0, p. 23. WGP, Darmstadt (2016)

Organizational Agility in Development Networks

M. Kuhn(✉), C. Dölle, M. Riesener, and G. Schuh

Laboratory for Machine Tools and Production Engineering WZL, RWTH Aachen University, Campus Boulevard 30, 52074 Aachen, Germany
m.kuhn@wzl.rwth-aachen.de

Abstract. The shortening of innovation and product life cycles induces a development environment that is characterized by high dynamics and uncertainty. Furthermore, the increasing integration of functions from the domains of mechanics, electronics as well as software into complex systems leads to the need for extensive competencies and resources. Traditional development and organization patterns are increasingly unsuitable for successfully meeting these challenges. A solution approach that addresses the described challenges is the interorganizational integration of economically autonomous companies into industrial development networks. Development networks offer fast and cost-effective access to required competencies and technologies. This paper presents a methodology for the systematic configuration of the organizational structure of a development network according to the respective situation of relevant requirements in order to increase organizational agility. Therefore, design recommendations for the organizational structure of the development network are derived based on a typology of requirements.

Keywords: Organizational Agility · Development Networks · Typology

1 Introduction

The product development of manufacturing companies in nearly every industry has faced multiple challenges in recent years. The continuing trend towards reducing trade barriers leads to the establishment of global markets with increased competition [1]. Furthermore, shortening innovation and product lifecycles and growing product complexity require greater efficiency and effectiveness of the product development of manufacturing companies in order to maintain competitiveness [2]. Companies need to be more flexible and have to adjust to changing environmental conditions.

In order to address these challenges manufacturing companies intensify collaborations with external partners in so-called product development networks. The adaption to unforeseen requirements through a flexible configuration of necessary capabilities within the product development is the main objective of a collaboration in development networks [3]. The approach seeks to enhance the overall organizational agility

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 621–630, 2021.
https://doi.org/10.1007/978-3-662-62138-7_62

and therefore addresses the core challenges currently faced by the manufacturing industry. Nevertheless, within actual research insufficient importance is assigned to the systematic design of the organizational structure of the development network [4]. This deficit results in productivity losses [5]. Therefore, this paper presents a methodology for the configuration of relevant elements of the organizational structure of development networks in accordance with the internal and external requirement situation. The methodology is based on the contingency theory as the relevant theory of organization [6]. In order to derive design recommendations quickly and purposefully, a topology of requirements that influence the design of the organizational structure is developed. The objective of the methodology is to systematically support the focal network management in the design and situational configuration of a product development network in order to enhance organizational agility. Organizational agility in this context describes the ability to quickly adapt the organizational structure to changing requirements.

The following sections are structured as follows: After the short introduction, the second section presents and evaluates related scientific literature in the context of the objectives of this paper. The third section presents the several steps of the methodology in detail. The fourth section concludes with a summary and a reflection of proposed further research.

2 Related Work

In the following section, relevant approaches within the context of organizational agility in development networks are examined. The analysis includes approaches, which focus on a configuration of development networks and emphasize to enhance organizational agility. Therefore the following approaches contribute to the objective of this paper.

"Transition between temporary organizations: Dimensions enabling economies of recombination" by Benedittis (2019) [7]: Benedittis contributes to temporary organization theory by identifying influencing factors to the process of transition between one form of an organization to another, which helps manufacturing companies to benefit from economies of recombination. The influencing factors are clustered in a structural, a conjunctural and an interstitial dimension. Although the approach provides starting points for the configuration of elements of the organizational structure, it does not specifically focus on development networks.

"Projects as temporary organizations" by Sydow et al. (2018) [8]: Sydow et al. conceptualize the theory of projects as temporary organizations in accordance with recent developments of managing interorganizational relations. Therefore, the authors introduce three dimensions for the analysis of interorganizational projects: the multi-level perspective, the processual understanding of relationships and modes of interorganizational governance. The approach does not provide any design recommendations depending on situational requirements and does not focus on organizational agility.

"Configurations for interorganizational project networks" by Braun (2018) [9]: Braun analyzes the role of a project management office within a single

corporation and reflects this concept on a network level in terms of the establishment of a network administrative organization. Thereupon, four types of organizational entities are developed and underlined with exemplary organizational configurations. The approach provides a valid basis for this paper, yet the elements of the organizational structure are not detailed and no situational design recommendations are derived.

"Agile network businesses" by Kale (2018) [10]: Kale considers networks from different perspectives according to the system theory. The author also provides basic information on the management of virtual enterprises and cooperative networks. Moreover, there are examples from several industries regarding different types of networks. Kale does not explicitly consider the correlation of organizational agility and the organizational structure of the network. Concrete suggestions for the design of the organizational structure are not presented.

"A typology of interorganizational project networks" by Ahola (2018) [11]: Ahola emphasizes that large projects are predominantly carried out in interorganizational networks that temporarily unify capacities of multiple firms. The author proposes a typology of interorganizational projects with three ideal network types: a market-based network, a dyad-driven network and an integrated core network. Structural differences between these types of networks are discussed. Nevertheless, a recommendation for the best-fitting type based on the respective requirement situation is not given.

"Dynamic capabilities and organizational agility" by Teece et al. (2016) [12]: Teece et al. develop ten principles that help network managers to handle trade-off decisions between agility and efficiency. The authors explain that organizational agility in companies is achieved through so-called dynamic capabilities, which by definition can be distinguished from ordinary capabilities. As dynamic capabilities stand for agility, key characteristics for example are the detection of market trends ahead of the competition, the creation of flexible supply chains and the incremental development of products. Beyond the derived principles, no design recommendations for the structural organization are given. Applicability in development networks is not discussed.

The presented approaches mainly focus on concepts for the enhancement of organizational agility and the configuration of organizational structure of manufacturing companies. Although some approaches examine an integrated analysis and identify relationships between the configuration of organizational elements and the effects on organizational agility, a distinct research gap can be identified in the systematic derivation of design recommendations for the organizational structure in accordance with the requirement situation that affects a development network. Therefore, the following section addresses this deficit and presents a methodology that supports the design of development networks.

3 Methodology

As described in the introduction, organizational agility within network-based development can help manufacturing companies to achieve competitive advantage in

recently volatile markets and dynamic environments. The following section of this paper presents a methodology to systematically support the focal network management in the design and situational configuration of the product development network in order to enhance its organizational agility. According to the guiding principle of agility, the organizational structure has to be reconfigured iteratively throughout different phases of a product development project [13]. The following methodology is divided into four steps. First, relevant elements of the organizational structure within a development network are derived. Therefore, a classification and description logic is presented. Second, relevant requirements from internal and external perspective are derived. Third, a typology for these identified requirements is developed in order to derive specific types of requirement, which significantly reduces the solution space for the design of the development network. Fourth, the relations between elements of the organizational structure and derived types of requirements are analyzed in order to give generic design recommendations. The four steps are described in detail in the following sections.

3.1 Description of Organizational Structure

In order to develop a model for the configuration of a development network, it is necessary to elaborate corresponding design parameters of the organizational structure. The organizational structure can be defined through formal and informal rules. In the context of this paper, the organizational structure mainly focuses on structural design parameters. Process organizational aspects, as the design of time sequences or workflows, are not focused. Based on the popular publication of Kieser and Kubicek [14], the following four structural dimensions provide a framework for the classification of organizational elements of development networks within this paper: *formalization*, *delegation*, *coordination* and *topology* (see Fig. 1).

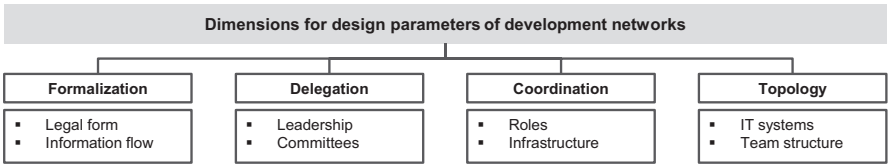


Fig. 1. Overview of design parameters of organizational structure.

By describing and combining distinctive characteristics of each organizational element, a configuration can be derived implying the individual character of a development network according to its specific requirements. *Formalization* defines the use of written rules and principles within the development network [14]. Relevant design parameters are the legal form of the network or the information flows within the network. The collaboration and therefore legal form of a development network can either be established on informal agreements or concrete contracts or ultimately result in a company participation [15]. Information flows can be formally defined or distributed

informally among the partners in the network [10]. *Delegation* describes the distribution of decision-making competencies within the network [14]. This applies in particular to the form of management, which can be located focally at a company or distributed polycentrically at self-organized partners within the network [16]. The form of management is also linked to the establishment of committee structures. Thus, concrete project decisions within the network can either be made centrally in regular committees or discussed in an informative, advisory committee. In the latter case, the implementation is carried out decentrally. *Coordination* includes the cooperative alignment of the individual partners and employees within the network [14]. A key design parameter in this context are the roles established within the network [17]. Roles can be explicitly described in defined role profiles or result implicitly from the cooperation. A detailed description of explicitly defined roles in a managed network, such as the performance manager, order manager, broker, auditor and network coach, are presented within the publication of the virtual factory [18]. In addition to the definition of roles, the infrastructure, i.e. the spatial location as well as the equipment and layout of the premises, e.g. in the form of ideation labs or prototype garages, is an important design parameter [19]. The cooperation of partners can take place locally in explicit premises or can be distributed via virtual project rooms and tools. *Topology* describes the hierarchical structure of the development network. Subsequently, an important design element of the structure of a development network is the anchoring of an IT system [20]. This can be set up as a general system and therefore be used consistently by all network partners (single-source). However, often existing and partly different IT systems of individual partners are used for short to medium-term cooperation (multi-source). Therefore, it is important to create suitable system interfaces to avoid information asymmetries [21]. A further design parameter is the team structure, which is determined by the form of cooperation and the division of tasks within the development team. Depending on the issue at hand, the team structure can be either disciplinary or interdisciplinary [8].

In addition to the design parameters presented, which are considered essentially important, there are possible further design parameters for the description of a development network. However, with the aim of a reduced minimum set of design parameters, a focus on the presented criteria is recommended and ensures applicability of the methodology. These design parameters should be configured according to existing internal and external requirements. In the following section a description logic for relevant requirements, which should be taken into account, is presented.

3.2 Identification of Requirements for the Design of the Development Network

In order to ensure a successful design of the development network, which is optimally adapted to the respective situation of requirements, it is necessary to identify and analyze the influences of requirements on the development network. Therefore, a suitable framework is helpful in order to categorize relevant requirements [22]. Due to the complexity of network organizations, a system-theoretical consideration is suitable. Accordingly, a development network (system), consisting of various partner

companies (system elements) and their relations to each other, can be distinguished from the corresponding environment of the development network (environment) [23]. Therefore, a distinction of requirements can be made between internal requirements from the actual development network and external requirements from the development environment. The description of requirements can be done solution-neutrally by morphologies.

Accordingly, requirements are described by a feature and the corresponding characteristics. This generic description logic enables a quick and situational evaluation of requirements. The requirement characteristics are determined specifically in the context of the individual and situational design of the development network. Figure 2 shows the generic model that categorizes individual requirements within the internal and external dimension.

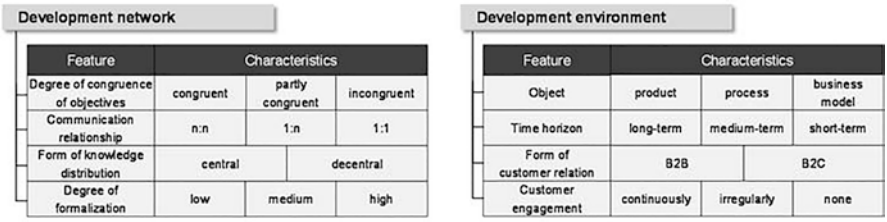


Fig. 2. Morphological description of requirements for the design of development networks.

Thus, from the internal perspective of the development network, it is necessary to assess the degree of congruence of objectives of the individual network partners and the communication relationship between these partners, as both requirements have implications for the design of the development network [24]. Further internal requirements are the form of knowledge distribution within the network, which can be centralized at a focal company or decentralized at the participating partners, and the degree of formalization within the network [15]. In this context, a high degree of formalization has a direct influence on the contracts to be concluded within the development network. In addition, from an external perspective it is necessary to evaluate, which content-related object and time horizon the network is pursuing [5]. The object to be developed within the network can be a product, a process or a business model. In addition, the cooperation between the partners can be short, medium or long-term. The customer relationship and the way in which the customer is involved in the development process also has an impact on the selection of the optimal organizational design parameters of the development network [25]. The derived features and characteristics of internal and external requirements should be considered initially in the design of the organizational structure of a development network. In order to derive recommendations for the design of a development network quickly and therefore enhance organizational agility through a purposeful configuration, a topology of requirements is presented in the following section.

3.3 Typology of Requirement for the Design of Development Networks

After the description of situational internal and external requirements on the basis of features and characteristics in the previous section, a reduction of the requirement space by means of a typology is conducted. In a typological reduction, the original characteristic space is reduced by simplifying characteristics, combining characteristic values or excluding characteristic values [26]. This reduction to specific requirement types is reasonable to ensure a simple and target-oriented applicability of the methodology. The development of a typology within an existing morphological requirement space consists of two main steps [27]. First, dependencies between requirement features are analyzed. This can be conducted, for instance, through an impact matrix. Those requirement features that are highly interdepend to other features can be identified as relevant requirement features for the derivation of requirement types. Second, representative requirement types are derived by a pairwise evaluation of the dependencies between requirement characteristics. The combination of highly interdependent characteristics leads to specific requirement types. Figure 3 shows four representative requirement types, which were identified through the previously described procedure: *Joint venture network, collaboration network*, *orchestrator network* and *virtual network*.

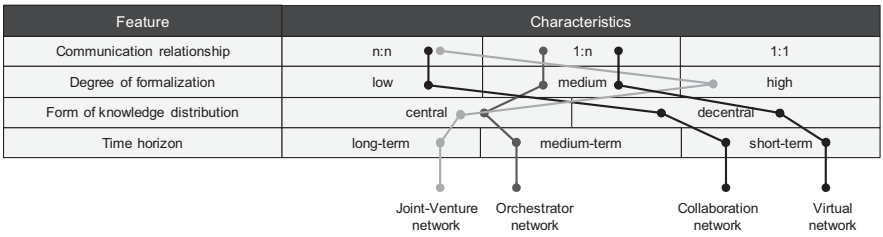


Fig. 3. Types of requirements for development networks.

Taking into account these four derived requirement types as well as the design parameters of the organizational structure derived within the first section of the methodology, design recommendations are given subsequently in the following section.

3.4 Design Recommendations for the Configuration of Development Networks

The selection of the best-fitting organizational structure of the development network is subject to the contingency theory, which constitutes that there is no universally efficient organizational structure, but rather a situational ideal expression [28]. Consequently, Fig. 4 gives an overview of the described steps of the methodology.

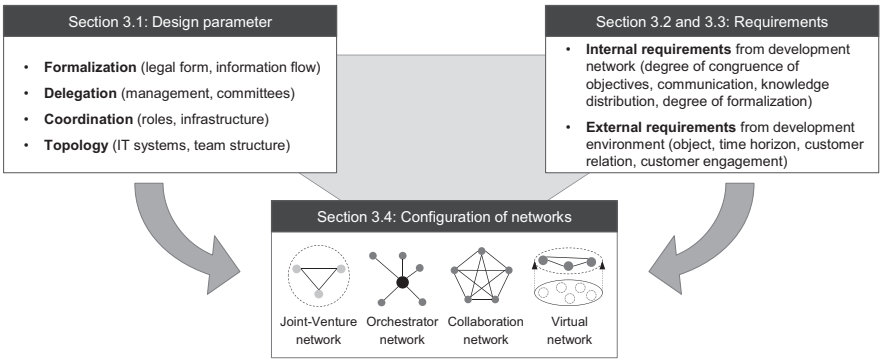


Fig. 4. Configuration of development networks.

In the following, design recommendations for the organizational structure and the respective design parameters are given in accordance to the derived requirement types.

The requirements of an "n:n" communication relationship, a high degree of formalization with central knowledge distribution and a long-term time horizon result in a requirement type that is addressed by a *joint venture network*. It is characterized by a horizontal cooperation of at least three companies that establish a joint venture to achieve a common objective. The cooperating companies often have an equal share within the equity of the established joint venture. There are informal flows of information within the system boundary, management is polycentric. Roles are explicitly formulated and separate premises are usually set up in order to realize joint success potentials in interdisciplinary teams.

The requirement characteristics of a "1:n" communication relationship, a medium degree of formalization with central knowledge distribution and a medium-term time horizon result in a requirement type that is addressed by an *orchestrator network*. The orchestrator network is characterized by a focal company and the synergetic, complementary complementation of partners. A mostly bilateral communication and a medium degree of formalization is realized through contractual agreements between the development partners and the focal company. The management is characterized by situational instructions and there are explicit roles. IT interfaces are usually harmonized in accordance with the focal company.

The requirement characteristics of a "n:n" communication relationship, a low degree of formalization with decentralized knowledge distribution and short-term development focus result in a requirement type that is addressed by a *collaboration network*. The collaboration network is characterized by a decentralized structure in which the configuration process of ready-to-use solutions of the development partners is focused. Communication between all development partners is crucial to ensure clear responsibilities and decision-making processes despite a low degree of formalization and is supported by informal information flows. The management of the network is characterized by decentralized, self-organized units. The roles in the network are not formulated explicitly, also due to the short-term nature of the time horizon. Therefore, defined contact persons are needed in all companies and for each function. A multi-source strategy of IT systems is pursued and an interdisciplinary team composition is targeted.

The requirement characteristics of an "1:n" communication relationship, a medium degree of formalization with decentralized knowledge distribution and short-term time horizon result in a requirement type that is addressed by a *virtual network*. The virtual network is characterized by the temporary linking of core competencies of independent companies. The aim is to establish a stable platform of partner companies that can be activated dynamically. A medium degree of formalization is realized through general agreements between the partners. Information flows to customers are formally defined, whereas there is an informal exchange of information between the partner companies. The network is focused on self-organized, decentralized management and excludes a focal role. Decisions are made decentrally and are coordinated by a temporary project manager. Due to the regularly recurring, order-related activation of the network, a special infrastructure in the sense of locally based premises and laboratories as well as a single source of IT infrastructure is appropriate.

4 Summary and Conclusion

The shortening of development lifecycles and growing complexity of technical systems lead to a demand for the establishment of development networks that offer fast and access to required competencies and therefore enhance organizational agility. This paper presents a methodology for the design of development networks with the objective of an enhanced organizational agility. The methodological steps were introduced to logically illustrate the approach. Further research has to focus on the specification of each step and the validation of the methodology within several industry case studies. The validation and further investigation of the presented methodology is currently part of the research activities of the authors in the department of Innovation Management at the Laboratory for Machine Tools and Production Engineering WZL. Especially the application of the concept in further companies will be a main activity, in order to gain data for the validation of the methodology. The authors would like to thank the German Research Foundation DFG for the kind support within the Cluster of Excellence "Internet of Production" – Project-ID: 390621612.

References

1. Schuh, G.: Innovationsmanagement. Springer, Berlin (2012)
2. Schuh, G.: Produktkomplexität managen. Hanser, München (2018)
3. Walters, D., Helman, D.: Partnerships: managing intra- and interorganizational relationships. In: Walters, D., Helman, D. (eds.) Strategic Capability Response Analysis, pp. 195–219. Springer, Cham (2020)
4. Schrempf, F., Schwaiger, M.: Survival of the Quickest – Agilität als organisationale Ressource in der digitalen Transformation. In: Obermaier, R. (eds.) Handbuch Industrie 4.0 und Digitale Transformation, pp. 623–648. Springer Gabler, Wiesbaden (2019)
5. Rebentisch, E., Riesener, M., Dölle, C., Kuhn, M., Brockmann, S.: Methodology for the design of agile product development networks. Procedia CIRP **84**, 1029–1034 (2019)

6. Miller, D.: Towards a new contingency approach: the search for organizational gestalts. J. Manag. Stud. **18**(1), 1–26 (1981)
7. De Benedittis, J.: Transition between temporary organizations: dimensions enabling economies of recombination. J. Proj. Manag. **37**(7), 912–929 (2019)
8. Sydow, J., Braun, T.: Projects as temporary organizations: an agenda for further theorizing the interorganizational dimension. Int. J. Proj. Manag. **36**(1), 4–11 (2018)
9. Braun, T.: Configurations for interorganizational project networks. Proj. Manag. J. **49**(4), 53–61 (2018)
10. Kale, V.: Agile Network Businesses. Collaboration, Coordination, and Competitive Advantage. An Auerbach Book. CRC Press Taylor & Francis Group, Bica Raton, London (2018)
11. Ahola, T.: So alike yet so different: a typology of interorganisational projects. Int. J. Proj. Manag. **36**(8), 1007–1018 (2018)
12. Teece, D., Peteraf, M., Leih, S.: Dynamic capabilities and organizational agility. Risk, uncertainty, and strategy in the innovation economy. Calif. Manag. Rev. **58**(4), 13–35 (2016)
13. Kuhn, M., Dölle, C., Riesener, M., Schuh, G.: Concept for organizational structures of agile development networks. In: Behrens, B.-A., Brosius, A (eds.) Production at the Leading Edge of Technology, pp. 653–662. Springer Vieweg, Berlin (2019)
14. Kieser, A., Kubicek, H.: Organisation. De Gruyter, Berlin (1992)
15. Salavisa, I., Sousa, C., Fontes, M.: Topologies of innovation networks in knowledge-intensive sectors. Technovation **32**(6), 380–399 (2012)
16. Wegehaupt, P.: Führung von Produktionsnetzwerken. Dissertation, RWTH Aachen (2004)
17. Schuh, G.: Management von Unternehmensnetzwerken – Konzepte zur Gestaltung, Lenkung und Entwicklung. In: Spath, D., Westkämper, E., Bullinger, H.-J., Warnecke, H.-J. (eds.) Neue Entwicklungen in der Unternehmensorganisation, pp. 265–282. Springer Vieweg, Berlin (2017)
18. Schuh, G., Millarg, K., Göransson, A.: Virtuelle Fabrik: Neue Marktchancen durch dynamische Netzwerke. Hanser, München (1998)
19. van Holm, E.J.: Makerspaces and contributions to entrepreneurship. Procedia Soc. Behav. Sci. **195**, 24–31 (2015)
20. Friedli, T.: Die Architektur von Kooperationen. Difo Druck, St. Gallen (2000)
21. Mattern, C.: Steigerung der Kollaborationsproduktivität in der interdisziplinären Produktentwicklung. Apprimus, Aachen (2018)
22. Ropohl, G.: Allgemeine Systemtheorie – Einführung in transdisziplinäres Denken. Edition Sigma, Berlin (2012)
23. Sydow, J.: Der soziotechnische Ansatz der Arbeits- und Organisationsgestaltung: Darstellung, Kritik, Weiterentwicklung. Campus, Fankfurt (1985)
24. Hagenhoff, S.: Innovationsmanagement für Kooperationen. Habilitation. Universitätsverlag Göttingen, Göttingen (2008)
25. Piller, F. T., Ihl, C., & Vossen, A.: A typology of customer co-creation in the innovation process (2010)
26. Kluge, S.: Empirisch begründete Typenbildung – Zur Konstruktion von Typen und Typologien in der qualitativen Sozialforschung. VS Verlag, Wiesbaden (1999)
27. Paulukuhn, L.: Typologisierung von Entwicklungsprojekten im Maschinenbau. Shaker, Aachen (2005)
28. Miller, D.: Towards a new contingency approach – the search for organizational gestalts. J. Manag. Stud. **18**(1), 1–26 (1981)

Towards a Concept for an Employee-Specific Retention Strategy in Value-Adding Areas

S. Korder(✉) and G. Reinhart

Institute for Machine Tools and Industrial Management (iwb), Technical University of Munich, 85748 Garching, Germany
svenja.korder@iwb.tum.de

Abstract. Demographic change is leading to a decline in population in Germany. As a result, fewer skilled workers are available and many jobs remain vacant in various sectors, including the production industry. The turnover of employees favors the problem. In order to reduce employee turnover and to retain skilled workers, different general retention strategies are available for companies. However, a strategy adapted to the needs and requirements of workers at value-adding work stations is missing.

This paper will present a research plan for the development of an employee-specific retention strategy as well as a first draft applicable to value-adding workplaces.

Keywords: Employee retention · Employee turnover · Value-adding workplaces

1 Introduction

Demographic change is currently a major challenge for many EU countries. The interaction between an ageing society, rising life expectancy and low birth rates is putting pressure on social and economic systems. While the proportion of 65 aged and older people was still 13% of the total population in the EU in 2018, this number is expected to rise up to around one third of the population (32%) in 2060 [1]. The figures are also comparable in Germany. In 2018, 19% of the German population was 67 years and older. By 2060 this proportion is expected to rise to 24% or even 30% [2].

Population trends are a key determinant of labour supply [3]. As a result, the developments described above are reflected in labour shortages. If the trend remains constant, in particular the groups with a low or medium level of education are expected to decline within the EU. A reduction from 50.7 to 14.0 million is expected for workers with a low level of education and from 108.2 to 74.2 million for labour force with a medium level of education [1].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 631–639, 2021.
https://doi.org/10.1007/978-3-662-62138-7_63

Germany is already affected by a shortage of skilled workers in many regions and sectors. In the value-adding areas of manufacturing companies, people with a medium level of education (people who have completed a vocational training) make up about two thirds and thus the largest proportion of the workforce [4]. As shown in Fig. 1, for example, the situation in the metal and engineering industry has drastically worsened in recent years [5]. As a result, there are fewer unemployed skilled workers in relation to available jobs.

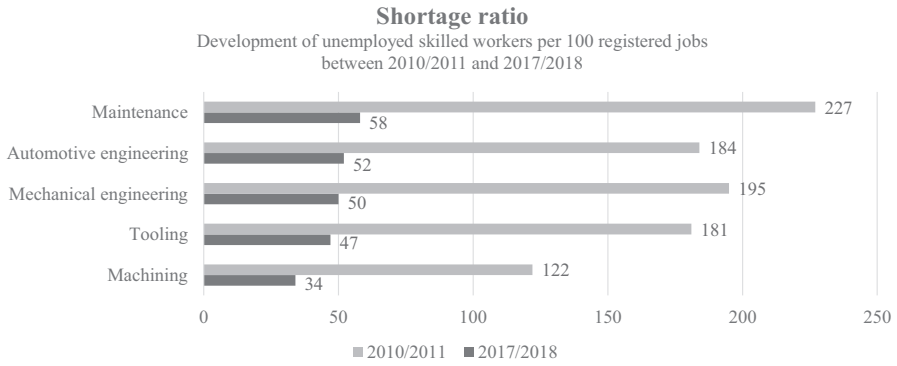


Fig. 1. Skilled labour shortage in the metal and engineering industry (after [5])

A large supply of jobs and a good economic situation also increase the rate of companies unwanted turnover, that means employee turnover based on a voluntary decision to leave the organization [6]. Due to the loss of know-how and productivity as well as the high costs of finding a suitable new worker, this type of turnover is often challenging for companies [7]. It is especially challenging in jobs with high labour shortages, because of the longer replacement time. With this in mind, methods for employee retention are gaining in importance [8].

Many methods aim at general approaches, such as work-life balance, different types of non-monetary benefits, financial incentives, development and training as well as job enrichment [9, 10]. However, it is crucial to adapt the methods to the needs of the individual employees [11].

Within the large number of methods and approaches for employee retention, especially production employees in value-adding areas have not been considered so far. Therefore, this publication presents a research plan for the development of an employee-specific retention strategy as well as a first draft of the application in value-adding areas.

This paper is structured as follows: Chapter two gives an overview of existing approaches in this field of research and outlines the research needs. In the following chapter three the research plan for the employee-specific retention strategy will be explained. The practical application of the draft will be presented in section four. Chapter five concludes the paper and gives an outlook.

2 Literature Review

In the literature, the topic is examined from two perspectives; turnover predictors and retention measures. The former investigates the causes of turnover. Numerous studies have been published in this research area. The latter concentrates on motives to stay with a company. Here, operational measures and longer-term strategies for retaining employees are presented. The status of both perspectives will be examined below.

2.1 Existing Approaches

Turnover Predictors. Many empirical studies examine various factors and their influence on turnover. One of the most frequently considered factors is the construct of (general) job satisfaction. Although different statements about the strength of the influence are given from strong to low. [12–16] The working environment, working climate, stress, working content and wages were also confirmed by studies as factors influencing the decision to turnover [14–19].

In order to provide a summary and overview, the Society for Human Resource Management published the previously examined factors and ordered them by their strength of influence [16]. These turnover predictors are shown in Table 1.

Table 1. Turnover predictors after [15] and [16]

Predictors of turnover after Griffeth et al. (2000) and Allen (2008)	
Decreasing relationship to turnover, strongest on the top	• *Turnover Intentions (+)* • *Thoughts of Quitting (+)* • *Search Intentions/Behaviors (+)* • *Weighted Application Blank (+)* • *Organizational Commitment (-)* • *Relationship with Supervisor (-)* • *Role Clarity (-)* • *Tenure (-)* • *Job Satisfaction (-)* • *Role Conflict (+)* • *Absenteeism (+)* • *Work Satisfaction (-)* • *Alternatives to Present Job or Company (+)* • *Satisfaction of Expectations of Job or Company (-)* • *Job Performance (-)* • *Stress (+)* • *Promotion Opportunities (-)* • *Children (-)* • *Alternative Job Opportunities (+)* • *Job Scope (-)* • *Quality of Communication in Organization (-)* • *Work-Group Cohesion (-)* • *Co-worker Satisfaction (-)* • *Participation in Decision Making (-)* • *Satisfaction with Supervisor (-)* • *Role Overload (+)* • *Job Involvement (-)* • *Age (-)* • *Pay (-)* • *Outcome Fairness (-)* • *Degree of Routinization of Job Responsibilities (+)* • *Family Responsibilities (-)* • *Training (-)* • *Pay Satisfaction (-)* • *Lateness (+)* • *Education (+)* • *Marital Status (-)* • *Sex* • *Cognitive Ability (+)* • *Race*	(+) Positive correlation to turnover | (-) Negative correlation to turnover

Retention Measures. Two types of retention measures can be distinguished. On the one hand, operational measures are directly applicable and affect the whole workforce. On the other hand, strategic or selective measures require an upstream process to classify the employees and apply specific measures to selected groups in order to retain them. In the following sections both will be explained in detail.

Operational Measures. Within this kind of measures, companies can implement material and immaterial measures. Wolf [20] considers material measures to be company cars or company housing, competitive salary, meal vouchers, performance-related bonuses and social benefits. Immaterial measures can include activities to promote the working atmosphere, flexible working hours, employee

interviews, career development, team-building measures, health management measures and so on. In a recent study from 2019, 868 people from the Human Resources (HR) department were asked about different important HR topics including employee retention. The companies had different industry classifications (service sector, industry, public sector) and were based in Germany, Switzerland and Austria. 43% of the participants selected employee retention as one of the most important HR topics. Furthermore, in order to increase employee retention, 44% of the companies stated that they already implemented measures for a good working atmosphere. That was followed by implementing interesting tasks (39%), flexible working hours (38%), employment security (38%), competitive salary (36%), and additional social benefits (35%). These results show that some operational measures are already being applied in companies. [21]

Strategic Measures. Strategic measures give companies a basic orientation to retain employees in the long term, e.g. the personnel portfolio or the function portfolio. The personnel portfolio is used to divide employees into high performance and high potential employees. They are then retained according to the classification. For example, high performers are more likely to receive incentives and high potentials are more likely to receive qualification measures. If employees cannot be assigned to the given categories, no measures are applied. The function portfolio classifies employees according to their relevance for the company. For example, employees whose qualification level is not available on the labour market are more likely to undergo retention measures [20].

2.2 Research Demand

Research Gap. The literature review shows that various turnover predictors matter. Whether job satisfaction, working content or qualification possibilities, the reasons for employee turnover can be very individual and also vary from different work conditions. It is therefore of great importance for companies to choose a strategy that is adaptable to different employees and conditions. Within the operational and strategic measures existing in the literature, however, it has been shown that up to now, there are only isolated solutions to pursue single goals or solutions for a special group of employees.

Industrial Need. As described in the introduction, the scope of this publication is on value-adding areas of manufacturing companies. The focus here is on stationary assembly workstations, which are located at the end of the product creation process. The assembly process is often characterized by complex activities and a high proportion of manual work [22].

The strategies described in the literature review are, however, unspecific with regard to their areas of application and do not take into account the requirements and circumstances of manufacturing.

In order to close the identified gaps, the aim of this paper is to develop a concept of a retention strategy. This concept follows a holistic approach to employee retention (meeting different kind of predictors). It is individually adaptable to the needs of

the employees and the respective working conditions and specifically addresses the requirements of assembly. The strategy thus describes the path the company takes to reach the goal of taking specific and targeted measures that fit the needs of the individual employee.

3 Research Plan Towards an Employee-Specific Retention Strategy

As shown in Fig. 2, the strategy is developed in five steps. The perspective in this section is that of the developer of the strategy.

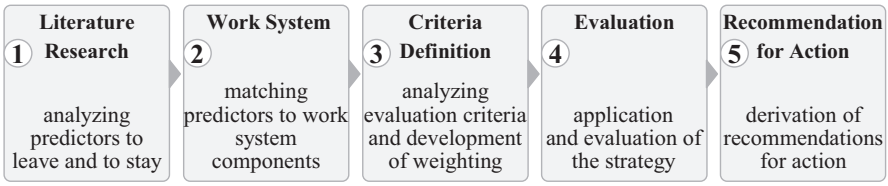


Fig. 2. The 5 steps for strategy development

Step 1. In the first step, a detailed quantitative literature research will be carried out. The reasons that induce an employee to leave the company (e.g. working contents) as well as those that induce him to stay (e.g. colleagues) will be analyzed. Afterwards, the predictors will be filtered with regard to their relevance for the assembly as well as their assessability. The filtered factors are the base for the following step.

Step 2. The identified predictors are now assigned to the components of the work system. According to REFA [23], these are the working environment, input, output, work task, work person, work equipment and work object. To stay with the above given example, the predictor working contents can be assigned to the work task and the colleagues to the working environment. All components of the work system, which have not been assigned to any predictors, can be ignored for the further procedure. The remaining components create superior assessment areas for the strategy.

Step 3. The aim of the third step is to define individual sub-criteria within each component or assessment area. In addition, these criteria should be made quantifiable by the development of a rating system. The sub-criteria help to specify the assessment areas and to derive concrete actions when applying the strategy (see Chapt. 4). They will be derived based on existing work analysis procedures and a literature search. Furthermore, a system for weighting individual criteria will be developed. In this way, areas with urgent need for action can be distinguished from others. The assessment area work task would have for example the criteria integrity of the work task, variety of requirements, time constraint.

Step 4. Within the scope of this step, an extensive application and evaluation of the procedure and the specific criteria at different assembly workplaces will take place. The aim of the evaluation is on the one hand to check the criteria for clarity, holistic and assessability and the weighting system for plausibility. On the other hand, the results obtained will be used for the development of recommendations for action for the companies.

Step 5. Finally, in this step, a procedure for companies to deal with the results is developed. The weighting of the evaluated criteria is used to estimate the urgency and thus the prioritization of further actions. The recommendation for actions will be a generic step by step guidance, since the evaluation results will be very individual depending on the different companies, workplaces and requirements.

4 Procedure for the Application of the Retention Strategy

While the research plan for developing the strategy was described in the previous section, this section focuses on the application of the strategy. As shown in Fig. 3, the strategy can be carried out in three phases: preparation, application and actions. These three phases build on the five steps described above. The results of steps 1–3 are mainly used in the application. Steps 4 and 5 form the basis for the actions.

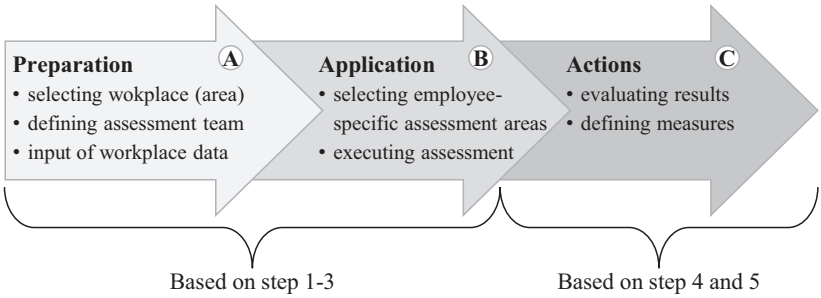


Fig. 3. Procedure for applying the strategy

Before the application at the workplace can take place, some preparations must be made (**Preparation**). First, a suitable work station or work area should be selected. Once this has been determined, the evaluation team has to be defined. At least three people should carry out the evaluation. A representative of the works council, the team leader responsible for the chosen area and finally the person working at the workplace. In order to have references for the evaluation and the single criteria (e.g. variety of requirements), workplace-specific key figures and information are collected (e.g. number of product variants to be processed, quantities of the product, cycle times, etc.).

Once all preparations are completed, the evaluation team can go to the work place for the assessment (**Application**). The employee involved can now choose the assessment areas that appear relevant to him or her or in which he or she sees potential for improvement (e.g. work environment, work task, work equipment, etc.). If, for example, the employee cannot see any potential in the areas of input and output of the workplace, these areas can be neglected when carrying out the evaluation. Within the selected areas, the weighting of the individual criteria is then carried out by the employee.

Once the evaluation is completed, the evaluation team can analyze the weightings of the individual criteria (**Actions**). Criteria with a high weighting require rapid action, because these indicate a turnover tendency. Based on the input of workplace parameters in the preparation phase, the measures can be derived very concretely. Possibilities and responsibilities can be discussed and assigned within the evaluation team.

5 Summary and Outlook

The literature review and the research demand showed the need for an employee-specific retention strategy. Moreover, a strategy that can be applied to the value-adding sector of manufacturing companies is not yet available. The research plan presented in this paper provides an approach to address both gaps. The strategy will be developed in five steps.

Step one lays the foundation with the analysis of turnover predictors. In step two, these predictors are to be assigned to the work system. Specific evaluation criteria for the strategy will be developed in the third step. In addition, a quantitative evaluation system will be developed here. The fourth step provides a major evaluation of the evaluation criteria in various companies. Finally, recommendations for action measures for companies are derived.

The presented concept will be carried out within the scope of a research project at the Institute for Machine Tools and Industrial Management (*iwb*) of the Technical University of Munich.

Acknowledgement. The authors thank the German Federal Ministry for Economic Affairs and Energy (BMWi) for its financial and organizational support of the “Mittelstand 4.0-Kompetenzzentrum Augsburg”.

References

1. European Commission: Demographic Scenarios for the EU. Migration, Population and Education. European Union, Luxembourg (2019)
2. Destatis: Bevölkerung im Wandel. Annahmen und Ergebnisse der 14. koordinierten Bevölkerungsvorausberechnung. Destatis, Wiesbaden (2019)

3. Zika, G., Schneemann, C., Hummel, M., Maier, T., Kalinowski, M., Bernardt, F., Mönnig, A., Parton, F., Sonnenburg, A., Ulrich, P., Wolter, M.: Langfristige Folgen von Demografie und Strukturwandel für regionale Arbeitsmärkte. IAB Forschungsbericht (2020)
4. Dettmann, E., Fackler, D., Müller, S., Neuschäffer, G., Slavtchev, V., Leber, U., Schwengler, B.: Fehlende Fachkräfte in Deutschland – Unterschiede in den Betrieben und mögliche Erklärungsfaktoren: Ergebnisse aus dem IAB-Betriebspanel 2018. IAB Forschungsbericht (2019)
5. Malin, L., Risius, P., Jansen, A., Schirner, S., Werner, D.: Fachkraeftecheck Metall- und Elektroberufe. Analyse der Fachkräftesituation (2018)
6. Armknecht, P., Early, J.: Quits in manufacturing: a study of their causes. Mon. Labor Rev. **95**(11), 31–37 (1972)
7. Anvari, R., JianFu, Z., Chermahini, S.H.: Effective strategy for solving voluntary turnover problem among employees. Procedia – Soc. Behav. Sci. **129**, 186–190 (2014)
8. Cardy, R.L., Lengnick-Hall, M.L.: Will they stay or will they go? Exploring a customer-oriented approach to employee retention. J. Bus. Psychol. **26**(2), 213–217 (2011)
9. Ghosh, P., Satyawadi, R., Prasad Joshi, J., Shadman, M.: Who stays with you? Factors predicting employees' intention to stay. Int. J. Org. Anal. **21**(3), 288–312 (2013)
10. Allen, D., Shore, L., Griffeth, R.W.: The role of perceived organizational support and supportive human resource practices in the turnover process. J. Manag. **29**(1), 99–118 (2003)
11. Allen, D.G., Bryant, P.C.: Managing Employee Turnover. Dispelling Myths and Fostering Evidence-Based Retention Strategies. Business Expert Press, New York (2012)
12. Hom, P.W., Mitchell, T.R., Lee, T.W., Griffeth, R.W.: Reviewing employee turnover: focusing on proximal withdrawal states and an expanded criterion. Psychol. Bull. **138**(5), 831–858 (2012)
13. Zimmerman, R.D.: Understanding the impact of personality traits on individuals' turnover decisions: a meta-analytic path model. Pers. Psychol. **61**, 309–348 (2008)
14. Ayodele, O.A., Chang-Richards, A., González, V.: Factors affecting workforce turnover in the construction sector: a systematic review. J. Constr. Eng. Manag. **146**(2), 3119010 (2020)
15. Griffeth, R., Hom, P.W., Gaertner, S.: A meta-analysis of antecedents and correlates of employee turnover: update, moderator tests, and research implications for the next millennium. J. Manag. **26**(3), 463–488 (2000)
16. Allen, D.G.: Retaining Talent: A Guide to Analyzing and Managing Employee Turnover (2008). available: https://www.shrm.org/hr-today/trends-and-forecasting/special-reports-and-expert-views/Documents/Retaining-Talent.pdf
17. Lo, J.: The information technology workforce: a review and assessment of voluntary turnover research. Inf. Syst. Front. **17**(2), 387–411 (2013). https://doi.org/10.1007/s10796-013-9408-y
18. Bridger, R.S., Day, A.J., Morton, K.: Occupational stress and employee turnover. Ergonomics **56**(11), 1629–1639 (2013)
19. Krill, M.: Mitarbeiterbindung als Umkehrung von Fluktuation: Implikationen der Fluktuationsdeterminantenforschung. Z. Manag. **6**(4), 401–425 (2011)
20. Wolf, G.: Mitarbeiterbindung. Strategie und Umsetzung im Unternehmen, 3rd edn. Haufe Gruppe, Freiburg (2018)
21. Eilers, S., Möckel, K., Rump, J., Schnabel, F.: Schwerpunkt Beschäftigungseffekte der Digitalisierung. HR-Report (2019)
22. Lotter, B., Wiendahl, H.-P.: Montage in der industriellen Produktion. Ein Handbuch für die Praxis, 2nd edn. VDI-Buch, Springer (2012)
23. REFA: Methodenlehre des Arbeitsstudiums, 7th edn. Hanser, München (1984)

Identification of Project-Related Context Factors for the Tailored Design of Hybrid Development Processes

J. Ays(✉), C. Dölle, M. Riesener, and G. Schuh

Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
johanna.ays@wzl.rwth-aachen.de

Abstract. To fulfil the strategic importance of product development, manufacturing companies need to deal with the underlying processes. As plan-driven approaches have proven to be too inflexible for today's market dynamics, development projects are alternatively carried out by using agile process models. However, while a too high degree of agility leads to efficiency problems, a too low degree brings major challenges in market competition. As a result, a multitude of hybrid approaches as an integration of agile and plan-driven processes arise. When defining those hybrid processes, the project context should be taken into account. The consideration of the context is decisive for an optimal process performance. Therefore, the focus of this research lies on the identification of relevant context factors. The knowledge of those factors enables the characterisation of each project individually and subsequently supports determining the appropriate level of agility as a prerequisite for adjusting the optimal hybrid development approach.

Keywords: Agile · Development · Context Factor

1 Introduction

Today's manufacturing companies have to deal with a highly dynamic and turbulent environment [1]. Due to the resulting market-based volatility, long-term planning is gradually becoming more uncertain and complicated [2], which has a particular impact on the development of physical products. Thus, development projects are confronted with shortened product life cycles and heterogeneous customer requirements [3]. For a business environment characterised by discontinuity rather than stability, the use of traditional plan-driven processes in product development appears inadequate [4].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 640–649, 2021.
https://doi.org/10.1007/978-3-662-62138-7_64

1.1 The Need of Hybrid Development Processes

Bearing this inadequacy in mind, it seems unavoidable for the manufacturing industry to strive for an optimization of the development processes with the aim of both shortening the time to market and ensuring the degree of innovation for every individual product [5]. According to Farnbach, the implementation of a detailed planned Stage-Gate process, as an example for a plan-driven process model, is not only ineffective but counterproductive and has long since ceased to be able to fully address today's market dynamics [6]. Those mainly sequential approaches are further criticised as being too invariable, linear and rigid [7]. In line with so-called agile approaches already successfully established in software industry, manufacturing companies have been trying to apply agile process models in physical product development for some time now [5]. A study conducted by the Laboratory for Machine Tools and Production Engineering (WZL) and the Fraunhofer IPT at RWTH Aachen University from 2018 shows that only a very small percentage of companies prefer a purely agile development. Alternatively, the majority of the 138 companies surveyed use agile process models selectively, i.e. alternately or in an individually defined mixed form [8]. Although the practical implementation within individual companies varies greatly, the motivation to combine the stability of structured process models with the flexibility and adaptability of agile approaches is uniformly in the foreground. Up to now, however, such integration has only been based on subjective assessments by individuals, which means that the use of so-called hybrid process models is predominantly unsystematic and undirected [9]. So far, in theory and practice, there exists no established concept for a systematic adaptation of a hybrid process model [10]. One of the main challenges in adjusting such a model is the diversity of development processes to be considered. Since development projects vary greatly, so do the individual processes, which means no "one-size-fits-all" general approach can simply be implemented [11].

1.2 The Importance of the Project Context for Hybrid Processes

To address the above-mentioned deficits, Cooper et al, among other researchers, recommend considering the optimal balance between agile and plan-driven elements when building a hybrid approach taking into account the specific and individual context of a development project. In other words, they advise adjusting hybrid process models individually according to the context factors that have a concrete impact on the development process [12]. In regard to the central importance of the context-term in this paper, a brief definition is given in the following.

Etymologically, the word *context* originates from Latin, where it means as much as "connection". Thus, the Merriam-Webster dictionary defines context as "the interrelated conditions in which something exists or occurs" [13]. Meißner et al. apply this understanding to product development and describe the context as the environmental circumstances in which a product development process takes place [14]. The term context factor refers to this definition and describes a factor that has an influence on the course of a development project and thus to the underlying processes [15].

Following this definition, Feldmüller summarises the current state of research in one precise sentence: “There is consensus that a hybrid approach has to be tailored to the product, project and environment context, but actually no consensus how to do that” [10].

1.3 Objective of the Paper

From the authors’ perspective, an important issue that hinders manufacturing companies from an optimal derivation of a context-oriented hybrid process is the lack of an empirically based selection of those context factors, which are determining for the process adjustment in practice. Not all context factors are important for every project or company. For a certain group of context factors, there may be several suitable development processes; conversely, a concrete development process may be suitable for different contexts. [16] The knowledge of the relevant context factors is the fundament for an effective and efficient hybrid process model oriented to the specific operating conditions of the development project [17]. Therefore, the focus of this research lies on the identification of project-related context factors. It is assumed that the provision of those context factors significantly supports the tailored adjustment of hybrid process models. By characterising a specific development project, the factors enable the realisation of the appropriate level of agility for the underlying hybrid process.

In line with the presented objective of this paper, the next sections deal with the consideration of related work and the description of the methodology for determining project-related context factors including the presentation of the research findings.

2 Related Work

Prior scientific approaches are evaluated below to outline the need for further research. For this purpose, the second chapter is divided into two parts. The first part concentrates on hybrid research approaches addressing the integration of plan-driven and agile processes while taking into account the project context. The second part deals with the specific research on context factors influencing product development.

2.1 Approaches Focussing Hybrid Product Development

Current literature on hybrid product development has been expanding especially in recent years. Even Cooper [18], who is the creator of the most known plan-driven process model, focuses in his latest publications on the integration of agile elements in the traditional Stage-Gate process. Although he does not name specific context factors that influence this integration, he emphasizes the need to be able to guarantee a situation-specific and context-based reaction capability for the resulting process. In accordance, Ahmed-Kristensen & Daalhuizen [19] published an empirically study making the practical implementation of hybrid processes a subject of discussion. The

authors describe in detail the characteristics relevant for the design of a hybrid process model, but only with regard to the four case studies, they conducted. General valid design implications, such as context factors, cannot be deduced from this research either. In contrast, Brehm et al. [20] build on the established, software-oriented approach of Boehm and Turner [21], who advocate combining agile and plan-driven process models depending on the situation and project, and include the influence of the product and its life cycle on a development process. For this purpose, they extend the criteria of personnel, dynamism, culture, size and criticality established by Boehm and Turner and add a product perspective. However, the impact of those factors on the adjustment of a hybrid process remains too superficial and generic for concrete design recommendations.

Conclusively, this compact overview stresses that there is a common understanding that the project context plays an essential role in product development. Nevertheless, current researches lack the concrete identification of those relevant context factors.

2.2 Approaches Focussing Context Factors for Product Development

There is a multitude of approaches in literature that deal with the analysis of influences on product development. Three of these publications are described in the following. Starting with Hales & Gooch [22], who stress the project context as one of the main parts to be considered when managing the engineering design, the effect of influences is described as the range "from being strongly positive towards the attainment of project goals, through neutral, to strongly negative". The author, therefore, provide a list of 41 context factors categorized in three levels: Macroeconomic, microeconomic and corporate. Ponn [23], in turn, presents a collection of context factors from different authors and sorts them into the categories development task/product, developer/team and general conditions. The presented catalogue shows many factors, which are named differently but are quite similar in content. However, Ponn does not develop his own redundancy-free list. In contrast, Gericke et al. [24] provide a scheme that consolidates 26 different approaches proposed by other researchers in the years from 1961 to 2013. The identified context factors are hierarchically organized and with regard to their congruence, merged in case of duplications.

Even though all three mentioned approaches offer a detailed description of potential context factors, none of the authors considers the criterion of process relevance or temporal constancy when determining those factors. Referring to the objective of this paper, however, this evaluation is essential when adjusting a hybrid process.

3 Methodology

Regarding the described initial situation and the identified need for research, the intention of the following methodology (see Fig. 1) is to enable manufacturing companies to adopt the potentials of hybrid product development processes.

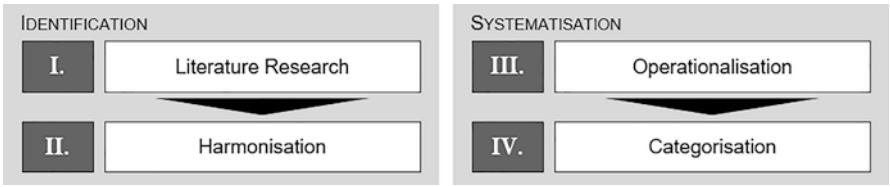


Fig. 1. Four-Step-Methodology for determining project-related context factors

Referring to PONN [23], there are many context factors to describe the situation around a development process, but not all of them are pertinent. Therefore, the consideration of the relevant context factors for an optimal project-oriented adjustment of hybrid development processes is presented as an important prerequisite for the successful use of such processes. Accordingly, the methodology aims at the definition of context factors that have a significant influence on hybrid development processes by presenting a four-step-approach for their determination.

3.1 Literature Research for the Identification of Context Factors

In a comprehensive literature analysis, context factors influencing product development from 20 different approaches proposed by other researchers were identified. By consolidating these empirical studies, 522 context factors could be compiled (see Table 1).

Table 1. Approaches enclosed in the consolidated catalogue of context factors

Albers et al. (2019)	PM Institute (2017)	Spundak (2014)	Ponn (2007)
Hüsselmann et al. (2019)	Klein (2016)	Wysocki (2014)	Paulukuhn (2005)
VDI 2221 (2019)	Rigby et al. (2016)	Gericke et al. (2013)	Hales & Gooch (2004)
Brehm et al. (2017)	Schoeneberg (2014)	Kuster (2011)	VDI 2223 (2004)
Ehrlenspiel et al. (2017)	Shenhar et al. (2014)	Lindemann (2009)	Specht et al. (2002).

By including the context factors defined in the work of PONN and GERICKE ET AL., approaches prior to 2002 were not considered further, since these authors again considered research results up to 1961 in their lists. Nevertheless, no claim is made to completeness, as the factors depend not at least on the uniqueness of the individual project situation. In a next step, the identified context factors were subsequently revised according to their congruence and merged in case of content-related overlapping. At the same time, care was taken to ensure that the factors are as independent of each other as possible. However, such dependencies cannot be avoided entirely, as there are always interdependencies between the factors [25]. The removing of duplicates lead to a number of 353 context factors remaining.

3.2 Harmonisation of Context Factors

The aim of the harmonisation step is to assess each context factor with regard to its relevance for the development process. Although the consolidated factors of the 20 approaches have a verifiable influence on product development, they are not necessarily relevant to the underlying process. Whether a context factor is relevant for the development process is defined by certain characteristics. Therefore, the harmonisation involves three criteria. If a context factor does not meet one of these criteria, it is excluded from further consideration.

Following the example of Hollauer [26], simplicity is the first criterion that is decisive. Whilst this criterion has no direct influence on the process relevance itself, it does play an important role, as each of the context factors must be evaluable. If the factor is too complex and lacks comprehensibility, there is a risk that the classification will not be valid. In the worst case, this would have negative effects on the design of the hybrid development process. This leads to the second criterion, which deals with the measurability and observability of a context factor. Thus, the context factors need to be as assessable as possible. The more objectively a context factor can be defined according to an individual project, the more accuracy can be ensured and vagueness reduced. Measurability allows avoiding estimations based on personal experience and subjectivity. However, the most important criterion for the assessment of process relevance is traceability. Traceability means the impact of a context factor is traceable to a concrete process element, either an agile or a plan-driven element. In this context, an element describes a constituent characteristic of a process model. Conversely, a concrete consequence for the adjustment of a hybrid development process can be derived from a traceable context factor.

As a result, the remaining 353 context factors could be further reduced to a number of 203 factors by considering the three defined criteria. Figure 2 summarises the reduction steps taken up to that point.

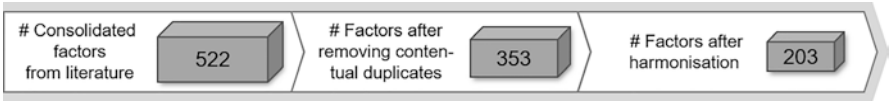


Fig. 2. Overview of the reduction steps

Within the framework of the initially presented four methodology steps, the section of the harmonisation completes the identification of relevant context factors. In order to bring the 203 factors into a manageable structure, the next step is the systematisation of the context factors. This is firstly carried out through operationalising each factor.

3.3 Description of Context Factors Through Operationalisation

In order to consider the defined context factors for the project-related adjustment of a hybrid process, one challenge is the valid and reliable evaluation of each factor. So far, this evaluation has been mainly carried out based on the personal knowledge and the realistic estimation of the project members. In contrast, the operationalisation enables the decomposition of a merely theoretical term into dimensions. Therewith, this step supports the possibility of classifying a context factor objectively in order to describe a specific development project. [27] With regard to the process relevance of the context factors, each dimension has a specific influence on a hybrid process, advocating either an agile or plan-driven process element.

For example, the context factor "Technological Novelty" can be split up into the four dimensions "Low", "Multi", "Interdependent" and "High" Technology. The higher the degree of novelty, the more agile process elements should preferably be selected when adjusting the hybrid process. Following this example, each identified context factor was divided into a number of three to five descriptive dimensions.

3.4 Categorisation of Context Factors

The categorisation represents the last step of the methodology. It concludes the systematisation by categorising the identified 203 context factors on three levels. In the here proposed categorisation, each context factor is described in terms of content, its significance for management and its temporal constancy. This differentiation allows a project-oriented consideration or exclusion of specific context factor groups. Thus, context factors e.g. with a high dynamic can be excluded for the adjustment of a hybrid process, if necessary. Figure 3 shows an extract of the factor catalogue.

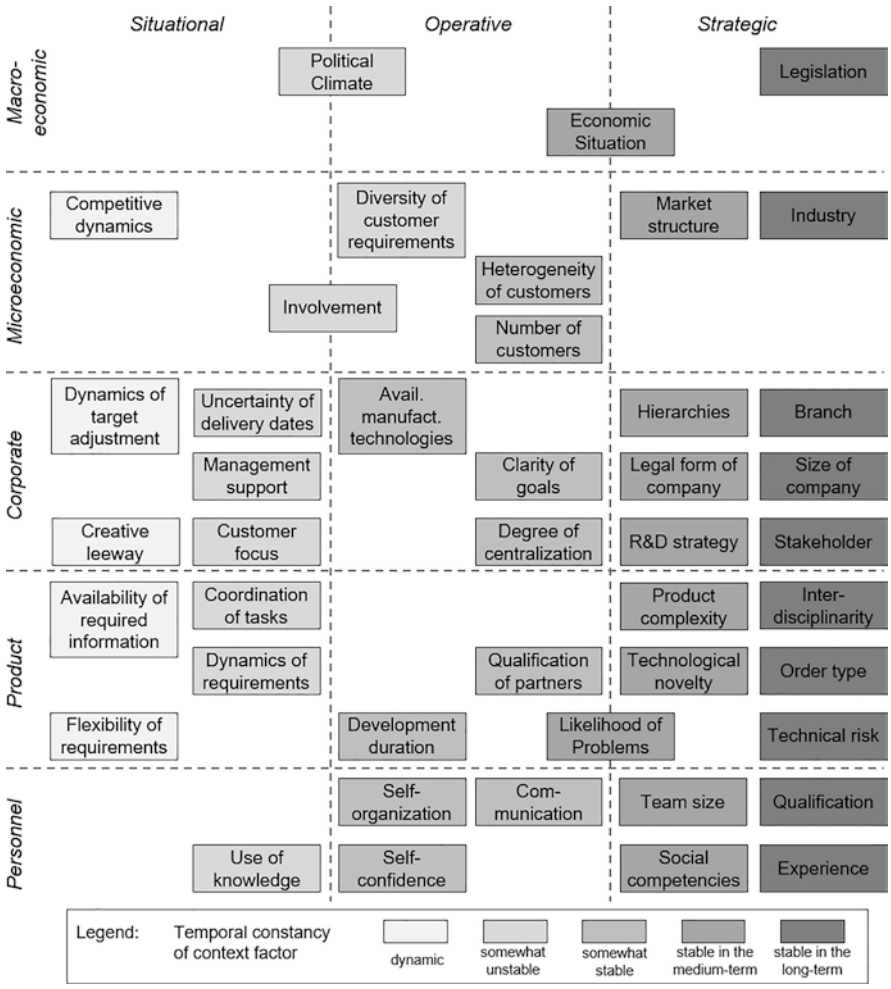


Fig. 3. Extract of the categorised factor catalogue

Next to the content and management level, the temporal constancy plays an important role for the adjustment of a hybrid process [16]. Not all context factors are equally dynamically affected by change. Depending on the stability of a context factor, the complexity of a hybrid process can be varied. A stable context factor remains fixed from the project starting point, a dynamic context factor instead might change frequently or at least occasionally during the development. The lower the stability of a factor, the more difficult it is to address the impact of this factor on the hybrid process.

4 Conclusion and Further Research

The presented catalogue of context factors is the result of a comprehensive literature analysis. 203 process-relevant context factors were derived, specified and systematised. It is assumed that the determination of project-related context factors will contribute to an improved tailored adjustment of hybrid development processes. Nevertheless, further research is needed to enable the adjustment realisation in practice. Above all, there is a research need to evaluate the context factors identified in terms of their tendency to agility. This step is necessary to facilitate companies to transfer the project characterisation according to the assessed context factors to the process adjustment itself. Therefore, future research will be concentrated on the classification of each factor dimension in terms of the appropriate agility level.

Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC-2023 Internet of Production - 390621612.

References

1. Mack, O., Khare, A., Krämer, A., Burgartz, T.: Managing in a VUCA World. Springer, Cham, Heidelberg (2016)
2. Mikusz, M., Heber, D., Katzfuß, C., Monauni, M., Tauterat, T.: Changeable manufacturing on the network level. Procedia CIRP **41**, 27–32 (2016)
3. Schuh, G., Dölle, C., Kantelberg, J., Menges, A.: Identification of agile mechanisms of action as basis for agile product development. Procedia CIRP **70**, 19–24 (2018)
4. Kienbaum Management Consultants GmbH: Agility - überlebensnotwendig für Unternehmen in unsicheren und dynamischen Zeiten. Change-Management-Studie. Kienbaum Management Consultants GmbH, Berlin (2015)
5. Schuh, G., Riesener, M., Ortlieb, C., Diels, F., Schröder, S.: Agile Produktentwicklung. In: Schuh, G. (ed.) AWK Aachener Werkzeugmaschinen-Kolloquium 2017 Internet of Production für agile Unternehmen, pp. 29–51. Apprimus Verlag, Aachen (2017)
6. Farnbach, J.S.: Why companies aren't jumping on the 'Agile' bandwagon. Visions (Product Development & Management Association) **35**(3), 36–39 (2011)
7. Cooper, R.G.: What's next?: after stage-gate. Res. Technol. Manage. **57**(1), 20–31 (2014)
8. Schuh, G., Kreutzer, R., Vogt, F., Schwarberg, F.: Agile Invention. Wie agile Arbeitsweisen erfolgreich in die frühen Phasen des Innovationsprozesses integriert werden. ZWF Zeitschrift für Wirtschaftlichen Fabrikbetrieb **114**(5), 284–287 (2019)
9. Nuhn, H.F., Martini, J.-P., Kostron, A.: Hybride Strukturen in der Automobilindustrie. Studie zu Agilen Praktiken in Forschungs- und Entwicklungsprozessen. In: Engstler, M., Fazal-Baqaie, M., Hanser, E., Linssen, O., Mikusz, M., Volland, A. (eds.) Projektmanagement und Vorgehensmodelle 2016, pp. 29–36. Gesellschaft für Informatik e.V., Bonn (2016)
10. Feldmüller, D.: Usage of agile practices in mechatronics system design. In: 19th International Conference on Research and Education in Mechatronics (REM), pp. 30–35. Delft (2018)

11. Shenhar, A., Dvir, D., Lechler, T., Poli, M.: One size does not fit all. True for projects, true for frameworks. In: Proceedings of PMI Research Conference 2002, pp. 99–106. Project Management Institute (2002)
12. Cooper, R.G.: Agile-Stage-Gate Hybrids. Res. Technol. Manage. **59**(1), 21–29 (2016)
13. Merriam-Websters (Definition "context"). https://www.merriam-webster.com/dictionary/context. Accessed 27 March 2020
14. Meißner, M., Gericke, K., Gries, B., Blessing, L.: Eine adaptive Produktentwicklungsmethodik als Beitrag zur Prozessgestaltung in der Produktentwicklung. In: Proceedings of 16th Symposium Design for X, pp. 67–76. Neukirchen (2005)
15. Gericke, K., Meißner, M., Paetzold, K.: Understanding the context of product development. In: 19th International Conference on Research and Education in Mechatronics (REM), pp. 191–200. Seoul (2013)
16. VDI Verein Deutscher Ingenieure e.V.: VDI 2221 - Blatt 1 - Entwicklung technischer Produkte und Systeme. Beuth, Düsseldorf (2019)
17. Riesener, M., Dölle, C., Ays, J., Ays, J.L.: Hybridization of development projects through process-related combination of agile and plan-driven approaches. In: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), pp. 602–606. IEEE (2018)
18. Cooper, R.: Idea-to-launch gating systems. Better, faster, and more agile. Res. Technol. Manag. **60**(1), 48–52 (2017)
19. Ahmed-Kristensen, S., Daalhuizen, J.: Pioneering the Combined Use of Agile and Stage-Gate Models in New Product Development - cases from the manufacturing industry. In: Proceedings of the 22nd Innovation and Product Development Management Conference. European Institute for Advanced Studies in Management (2015)
20. Brehm, L., Feldmüller, D., Riebke, T.: Konfiguration des hybriden Projektmanagements für die Entwicklung technischer, physischer Produkte. In: Barton, T., Herrmann, F., Meister, V., Müller, C., Seel, C. (eds.) Prozesse, Technologie, Anwendungen, Systeme und Management 2017. Tagungsband zur 30. AKWI-Jahrestagung, pp. 30–39. Aschaffenburg (2017)
21. Boehm, B., Turner, R.: Balancing Agility and Discipline. A guide for the perplexed. Addison-Wesley, Boston (2004)
22. Hales, C., Gooch, S.: Managing engineering design, 2nd edn. Springer, New York (2004)
23. Ponn, J.: Situative Unterstützung der methodischen Konzeptentwicklung technischer Produkte. Hut, München (2007)
24. Gericke, K.; Meißner, M.; Paetzold, K.: Understanding the Context of Product Development. In: Proceedings of the 19th International Conference on Engineering Design (ICED 13), pp. 191–200. Design Organisation and Management, Seoul (2013)
25. Kalus, G.: Projektspezifische Anpassung von Vorgehensmodellen. Feature-basiertes Tailoring. Technical University of Munich, Munich (2013)
26. Hollauer, C.: Workshop-based tailoring of interdisciplinary product development processes by means of structural analysis. Technical University of Munich, Munich (2019)
27. Hennen, C., Kalenborn, A., Stadlbauer, S., Timm, I.: Systematisierung der Auswahl von Vorgehensmodellen durch Kennzahlen. In: Engstler, M., Fazal-Baqaie, M., Baqaie, M.F., Hanser, E. (eds.) Projektmanagement und Vorgehensmodelle 2015. Hybride Projektstrukturen erfolgreich umsetzen. 22th/23th October in Elmshorn, pp. 55–65, Gesellschaft für Informatik e.V. (GI), Bonn (2015)

Systematization of Adaptation Needs in the Design of Global Production Networks

N. Rodemann[1(✉)], M. Niederau[2], K. Thomas[1], A. Gützlaff[1], and G. Schuh[1]

[1] Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
n.rodemann@wzl.rwth-aachen.de

[2] Forschungsinstitut für Rationalisierung (FIR) e. V., RWTH Aachen University, Campus-Boulevard 55, 52074 Aachen, Germany

Abstract. In order to ensure long-term competitiveness, production networks need to be designed in a continuous way that allows a quick adaptation to changing conditions. Today, due to high complexity in production networks adaptation needs are identified too late and the selection of the appropriate adaptation reaction is a time-intensive process. This turns resilience into one of the main challenges in creating robust global production networks. Therefore, in order to reduce the complexity in the design of production networks, this paper presents a systematization of adaptation needs in network configuration. The systematization can be used to facilitate the identification of adaptation needs by structuring adaptation options and reducing them to a finite number of options. The structured approach to derive the adaptation cases in response to changes serves as a first step to identify adaptation needs at an early stage.

Keywords: Global production networks · Network design · Network adaptation

1 Introduction

The supply chains of many manufacturing companies are increasingly globally distributed [1]. The global production networks created in this way are exposed to a large number of influencing factors [2]. The multitude of these factors is characterized by a high degree of uncertainty and volatility, which prevents the prediction of future events [3]. In order to be competitive, however, companies must react as quickly as possible to changes and adapt the network configuration [4]. In this context, the term network resilience is used, which describes the capability and ability of the footprint to return to a stable state after a change or disruption [5]. Resilience of the global production network can be achieved by faster identification of adaptation needs and immediate initiation of countermeasures [6].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 650–659, 2021.
https://doi.org/10.1007/978-3-662-62138-7_65

The resilience of the network requires a continuous planning process that includes the proactive identification of adaptation needs and the derivation of measures at the tactical level [7]. However, today's production networks are complex in their structure [2]. The current design approaches are mostly very complicated and are no practical means to quickly adapt to changes [8]. In industrial practice, adjustments to the production network are made either organizationally through strategic corporate decisions or through one-off optimizations [9].

Thus, companies lack an adequate approach for the structured design and continuous control of the network [10]. As a result, the effort of applying strategic approaches systematically is considered too high for tactical use. The extent to which the effort of adaptation is appropriate must be given greater consideration. This can be achieved by introducing a way of systematization that is based upon possible adjustments.

Accordingly, the focus of this paper consists of a systematization of adaptation cases in the configuration of global production networks, which is of great necessity in the context of tactical management of production networks. The containment and systematization of design options in the network configuration enables a reduction of complexity and justifies the first step towards a data-based identification of adaptation needs.

2 State of the Art

The described necessity of rapid adaptation of the network design of manufacturing companies is a topic that has been dealt with by various authors in literature. Ferdows points out that the main problem is the long adaptation time from the perception of the change to the implementation of the measure [4].

Regarding to Anthony, within planning problems, a distinction is made between three basic levels, which namely are strategic planning, tactical planning and operational planning [11]. Based on this classification Schuh et al. developed a continuous design process for global production networks, while taking into account all of the levels [7]. Since the operational level represents the implementation of the change, the strategic and tactical level was emphasized in the literature research.

Some authors focus on the strategic level. Christodoulou et al. have recognized in practice the deficit of a robust, comprehensive and repeatable process for the configuration of global production networks and developed a strategically designed business process to anchor the configuration [12]. A reference framework for continuous adaptation is likewise presented by Neuner. Unlike other approaches, the frame of reference can be adapted relatively easily to the question and the effort involved, while its continuous character is realized by the control loop of the management. However, a tactical orientation is neglected [13]. In further elaborations, Ferdows et al. try to reduce the complexity of production networks by dividing them into subnetworks and to identify deviations from the corporate strategy by regular audits [8]. Strategic factors are also discussed within the approach of Wiezorrek, whose model is based on the idea of a continuous decision-making process in order to identify changes at an early stage and to proactively determine the necessary need for

adaptation [14]. Other authors, such as Sager et al., propose a continuous monitoring of the forecast for the key performance indicators that are used to evaluate an existing manufacturing network [15].

At the tactical level Gölzer et al. aim at an increase in flexibility and changeability in production structures that should be made possible by shifting strategic management tasks to the tactical level. In order to set focus on planning and evaluating adjustments to the production network during the operational phase they propose generic planning cases that can be used to plan, execute and validate the necessary adjustments [16]. Additionally, the use of big-data techniques is suggested, while the basis of the procedure is formed by a detailed and up-to-date digital shadow of the production [17]. Wittek also provides an example of increasing flexibility in production networks on a tactical level. Thereby, the goal is an optimal allocation of products to the locations, whereby in this model the contribution margin is maximized [18]. Moser et al. present a method for risk-efficient migration strategies for global production networks by combining a stochastic, dynamic optimization model with a portfolio theory approach [5]. At the plant level, Hartkopf deals with the topic of how adaptability can be planned and used in a structured way. To this end, he assigns conversion drivers to the content categories of products, quantities and technology [19].

In total, it can be stated that some authors deal with the adaptation of network configuration and give continuous strategic as well as tactical approaches. However, most of the approaches barely consider a systematization of the adaption with regard to the tactical background. The identified approaches concentrate on the solution of the adaptation of the network configuration but do not limit the solution space, which results in a high degree of complexity. A structured description of different adaption cases has not yet been covered in literature and a systematic generation or identification of adaptation needs is missing.

3 Conception of the Approach

The suggested approach offers a consistent systemization of adaptation needs regarding the configuration of global production networks. It aims to improve the structuring of the problem by systematizing the description of the adaptation cases that can occur in the design of global production networks. Due to the multitude of options in network configuration, this is a necessary first step to enable a data-based identification of adaptation needs. Based on this first step, in future research work the described adaptation cases in this paper will be linked to concrete causes and measurable key figures. Using a data-based analysis, correlations between the causes of adaptation needs are identified, which is followed by the monitoring of the influencing factors to finally derive the required adaptation need. This paper focuses on the first step to systematize the adaptation needs in order to enable a further data-based analysis as described.

The approach includes four steps which focus on the description of actuating variables in global production networks, the discretization of degrees of freedom and their possible variations. The first step defines and examines relevant actuating variables

which characterize an adaptation. This is done in order to analyze the degrees of freedom of each variable in the next step. On the basis of the determined degrees of freedom, the modelling of complex configuration measures is examined, considering all relevant objects in the network. In the last step, possible contradictions are ruled out by a pairwise comparison.

The systematic procedure allows to describe every possible adaptation need in the network configuration and enables decisions to be structured at a later point. To reduce complexity, this approach focuses on the description of production problems. Logistics problems are outside the scope of this paper and have to be considered within supply chain management in a separate step.

3.1 Description of the Actuating Variables

In the first step, object types are examined, which means that all objects occurring in real production networks are assigned to classes of objects, the object types. These are formally defined. This approach allows a generic consideration of production networks. The object types are considered as actuating variables for the production network planner. When identifying the relevant actuating variables for the tactical configuration, the right level of detail is crucial, which is examined in the following.

The entire production network can be represented by superimposing several node/edge models, each representing the subnetworks of the product families. Accordingly, a continuous edge in a subnetwork can be understood as the flow of a product or product component between locations through the network. This flow must be controlled and the necessary resources for the transformation of the object streams must be provided. Such a group of edges is referred to as a production chain in the following. A definition is introduced for production chains, which will be used as a basis for further considerations:

Process chain: *A continuous flow of repeat factors through the system, represented by edges. An edge always has a node as its target and a type of transport as its mode.*

Besides repeat factors, which are represented by production chains, potential factors are primarily relevant in the configuration. If a global production network is disaggregated, different locations can be identified in a first step; these are characteristic for a production network. According to graph theory, the disaggregated model can be described as a network of factories with matrix connections, in which each node (location) influences the other nodes and therefore cannot be managed independently [20].

Location: *A non-empty quantity of operating resources contained in a geographical location.*

This level of detail, however, is insufficient for modelling the processes involved in network configuration. For this reason, the existing model is disaggregated again and operating equipment is considered. These are decisive for the flow through the sites

and at the same time the reason for the flows between the sites. In order to achieve an appropriate level of detail, the operating equipment is combined into resources. A resource represents a set of operating, storage and handling resources that are constantly passed through together, so it is a substitute object with the same characteristics. The production manager is responsible for making this distinction; the level of detail is decisive. The flow of other production chains is also relevant for this differentiation. Therefore, these should be taken into account when delimiting the resources. An important criterion for the delimitation of resources is the typology of the material flow [21].

Resource: *A non-empty set of operating, conveying and storage equipment that performs manufacturing operations. The characteristics of a resource are unique and may be aggregated.*

In the current state of aggregation, resources represent black box models. The processes in the resources must be described to ensure the continuity of the edges in the production system. Therefore, production activities are used in this case to describe the input/output relationships of the black box models. In this paper, a qualitative description of the production activities is sufficient. In the following, the production activities are referred to as production processes.

Production process: *A process of transformation that produces economic or consumer goods from raw materials using energy and certain means of production or through further processing steps. The process is carried out within a resource and describes the operations within a resource.*

At this level of detail, disaggregation is discontinued because the model is considered sufficiently accurate for the tactical management of production networks. The actuating variables identified for tactical management are the object types *production chain, location* and *resource* as well as the production activity *production process*, which are shown as an overview in Fig. 1. Comparably defined structural levels of production networks can be found in literature [22].

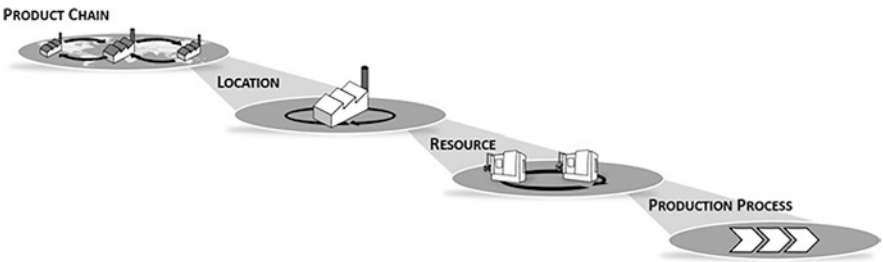


Fig. 1. Overview of the object types and their aggregation

3.2 Discretization of the Degrees of Freedom

The aim of the approach is to improve the structuring of the problem by systematizing the description of adaptation reactions. The improved structuring of the description supports the modelling of the decision-problem and thus creates the basis for an informed decision. The model developed in the following sections can be used to investigate possible degrees of freedom and their variations. In this case, an appropriate level of detail must be chosen, since in reality the adjustments are manifold and complex. Therefore, adjustments at the elementary level are considered for the identified actuating variables. When deriving possible adaptations, discrete states are defined which are called adaptation reaction. The individual adaptation reactions are unambiguous and do not overlap.

The adaptation reactions describe in sum all states that a manipulated variable can pass through during its existence. In order to define them formally, this principle can be concretized by transferring the life cycle concept. In this context, the distinction between individual product and product type is relevant. In the context of the product life cycle, the term ecological product life cycle, which is applied to an individual product, is usually used, which is contrasted with the economic product life cycle, which is defined for an entire product type. The economic product life cycle is also referred to as the marketing life cycle [23]. The economic product life cycle consisting of the phases introduction, growth, maturity, saturation and degeneration is applied to the four identified actuating variables and serves as a support for the definition of the adaptation reactions. In this way, for example, the degrees of freedom ramp-up, adaptation and ramp-down can be identified for the object type production chain. This orientation is also maintained for the remaining actuating variables to define all possible states of an object type during its existence.

By following this principle, a framework can be created that can be seen as a descriptive model for adaptation reactions in the configuration of global production networks. The full factorial variation creates a framework whose elements allow a comprehensive assessment of the effects on the production network on a diminutive level. At the same time, the space contains all possible manifestations of these elements. The spanned space is shown in a compact representation in Fig. 2. Each object type has the additional value “no adaptation” to extend the possible adjustment reactions by this case. To define an adaptation strategy, an adaptation reaction must be selected for each object type.

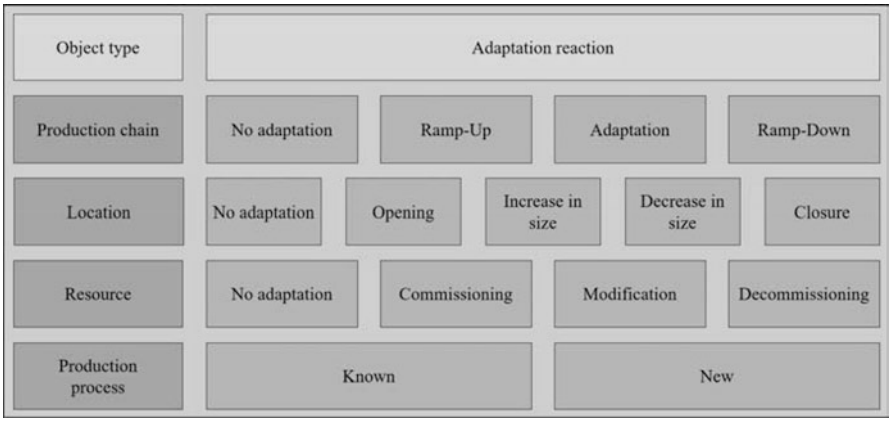


Fig. 2. Configuration framework of the considered object types

3.3 Variation of the degrees of freedom

In the description model, the activity space can be spanned by a full factorial variation of the adaptation reactions. By extending the framework with a list of possible combinations of adaptation reactions for each object type and assigning explicit ordinals to them, a systematic structure for describing the possible combinations is created (see Fig. 3).

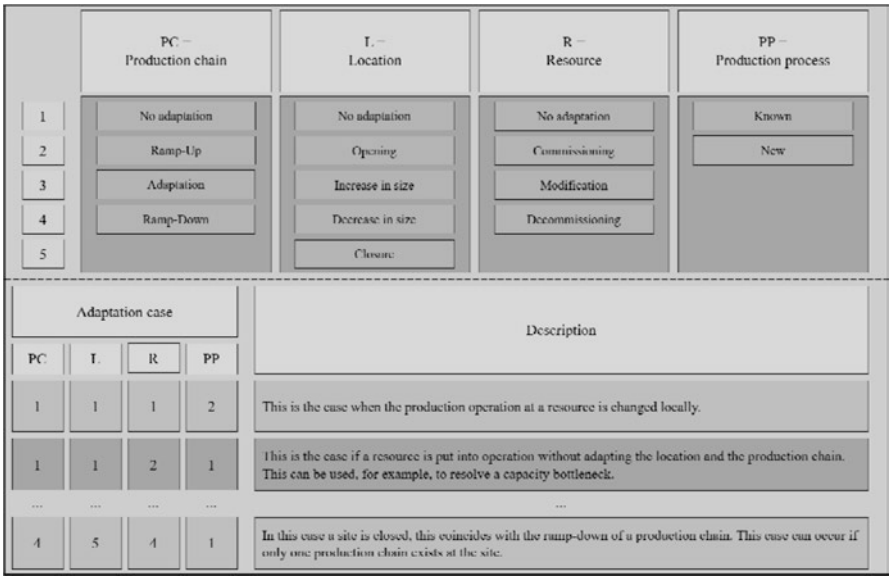


Fig. 3. List of possible combinations of adaptation reactions

The concept offers the possibility to describe all adjustment cases and give their possible effects on the production network. However, the absence of further restrictions within the framework leads to possible combinations that may contain contradictions. For this reason, the 160 possible adaption cases need to be checked for internal consistency.

3.4 Reduction of the Adjustment Cases (Consistency Check)

To efficiently reveal all contradictions, all adaptation reactions are compared pairwise. Adaptation reactions of an object type are not compared because they do not occur together in an adaptation case. This procedure allows to identify all contradictions within the cases. It is postulated that within an adaptation case an adaptation reaction can cause or exclude an adaptation reaction of another object type. If, for example, the object type *location* takes on the adaptation reaction *opening*, this requires a new resource to be put into operation and thus the object type *resource* takes on the corresponding characteristic value (*commissioning*). Otherwise, a contradiction would occur.

As a result, all adaptation cases that contain the specific combination of adaptation reactions are excluded. This procedure reduces the number of cases from 160 to 62.

For a more extensive reduction, a different procedure is chosen, since otherwise, the effort would increase exponentially. Accordingly, the resulting 62 adaptation cases are checked individually for consistency, whereby only one further contradiction is detected. As a result, a set of 61 possible adjustment cases improves the structuring of the decision problem, supports its modelling and thus creates the basis for a solid decision.

4 Conclusion

It is essential for companies to react to changes at an early stage and build up resilience in their network. The high complexity of the network configuration presents a main challenge to recognize adaptation needs quickly. Therefore, in this paper, a practical approach is presented to systemize adaptation needs regarding the design of global production networks in order to reduce complexity.

In the future, the presented model can be further developed to enable a data-based identification of adaptation needs in network configuration. For this purpose, in future research work, the adaptation cases determined by the approach presented will be examined for their causes by identifying relevant influencing factors. To this end, a data-based approach will be chosen to examine historical data for correlations and proactively identify adaptation needs through trend analyses.

Since real production networks are very complex and the described adaptation needs for complexity reduction are simplified, validation must also be carried out on the basis of a real use case. The possible adaptation reactions are to be derived as described in a data-based way and presented as proposals to the network planner and do not replace a detailed analysis by experts. The adaptation reactions shown in

this paper are limited to the variables production chain, location, resources and production process. Identified adaptation needs and possible adaptation reactions have to be examined and evaluated to determine whether the cause of the adaptation can be countered by other measures, for example by measures that are easier to implement. However, to keep the complexity of the model low, such measures should only be validated by experts in a second step, which is currently being developed in further research work.

Acknowledgements. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

1. Ferdows, K.: Keeping up with growing complexity of managing global operations. Int. J. Oper. Prod. Manage. **38**, 2 (2018)
2. Váncza, J.: Production networks. In: Laperrière, L., Reinhart, G., (eds.) CIRP Encyclopedia of Production Engineering. Springer, Berlin (2016)
3. Schuh, G.; Prote, J.-P.; Fränken, B.; Ays, J.; Cremer, S.: Dedicated agility: a new approach for designing production networks. In: IEEE International Conference on Industrial Engineering and Engineering Management, Bangkok (2018)
4. Ferdows, K.: Relating the firm's global production network to its strategy. In: Johansen, J., Farooq, S., Cheng, Y. (eds.) International Operations Networks. Springer, London (2014)
5. Moser, E., Stricker, N., Lanza, G.: Risk efficient migration strategies for global production networks. Procedia CIRP **57**, 104–109 (2016)
6. Lanza, G., Ferdows, K., Kara, S., Mourtzis, D., Schuh, G., Váncza, J., Wang, L., Wiendahl, H.-P.: Global production networks: design and operation. CIRP Ann. **68**(2), 823–841 (2019)
7. Schuh, G., Prote, J.-P., Dany, S.: Reference process for the continuous design of production networks. In: IEEE International Conference on Industrial Engineering and Engineering Management, Singapore (2017)
8. Ferdows, K., Vereecke, A., Meyer, A.: Delayering the global production network into congruent subnetworks. J. Oper. Manage. **41**, 63–74 (2016)
9. Friedli, T., Mundt, A., Thomas, S.: Strategic management of global manufacturing networks. Aligning strategy, configuration, and coordination. Springer, Berlin (2014)
10. Friedli, T., Thomas, S., Mundt, A.: Management globaler Produktionsnetzwerke: – Strategie – Konfiguration – Koordination. Hanser, München (2013)
11. Anthony, R.N.: Planning and control systems. A framework for analysis. Harvard Business School, Boston (1965)
12. Christodoulou, P., Fleet, D., Hanson, P., Phaal, R., Probert, D., Shi, Y.: Making the right things in the right places. A structured approach to developing and exploiting 'manufacturing footprint' strategy. Cambridge (2007)
13. Neuner, C.: Konfiguration internationaler Produktionsnetzwerke unter Berücksichtigung von Unsicherheit. Dissertation, Universität Bayreuth, Bayreuth (2009)
14. Wiezorrek, A.: Beitrag zur Konfiguration von globalen Wertschöpfungsnetzwerken. Dissertation, Technische Universität Dortmund, Dortmund (2017)
15. Sager, B., Hawer, S., Reinhart, G.: A performance measurement system for global manufacturing networks. In: 49th CIRP Conference on Manufacturing Systems, pp. 61–66 (2016)

16. Gölzer, P., Arnhold, D., Macke, N., Amberg, M.: Taktische Planung in Produktionsnetzwerken. Herausforderungen für Zulieferer mit großer Produktvielfalt und hoher Wertschöpfung. ZWF Zeitschrift für wirtschaftlichen Fabrikbetrieb **108**(5), 300–304 (2013)
17. Gölzer, P., Simon, L., Cato, P., Amberg, M.: Designing global manufacturing networks using big data. Procedia CIRP **33**, 191–196 (2015)
18. Wittek, K.: Standortübergreifende Programmplanung in flexiblen Produktionsnetzwerken der Automobilindustrie. Dissertation, Braunschweig (2013)
19. Hartkopf, M.: Systematik für eine kontinuierliche und langfristig ausgerichtete Planung technologischer und kapazitiver Werksentwicklungen. Bd. 18: Stuttgarter Beiträge zur Produktionsforschung. Fraunhofer, Stuttgart (2013)
20. Shi, Y., Gregory, M.J.: International manufacturing networks - to develop global competitive capabilities. J. Oper. Manage **16**(2–3), 195–214 (1998)
21. Dyckhoff, H., Spengler, T.: Produktionswirtschaft. Eine Einführung. Springer, Berlin (2010)
22. Krebs, P.: Bewertung vernetzter Produktionsstandorte unter Berücksichtigung multidimensionaler Unsicherheiten. Herbert Utz, München (2012)
23. Walther, G.: Nachhaltige Wertschöpfungsnetzwerke. Überbetriebliche Planung und Steuerung von Stoffströmen entlang des Produktlebenszyklus. Gabler, Wiesbaden (2010)

Data-Assisted Value Stream Method

C. Urnauer[1(✉)], V. Gräff[1], C. Tauchert[2], and J. Metternich[1]

[1] Institute for Production Management, Technology and Machine Tools, Technical University of Darmstadt, Otto-Berndt-Str. 2, 64287 Darmstadt, Germany
c.urnauer@ptw.tu-darmstadt.de
[2] Technical University of Darmstadt, Software & Digital Business Group, Hochschulstr. 1, 64289 Darmstadt, Germany

Abstract. The value stream method is widely used in the manufacturing industry to analyze and redesign value streams. However, with the increasing complexity of modern production systems, conducting a value stream analysis (VSA) and extracting reliable information for an accurate value stream design (VSD) becomes a challenging task for practitioners. Utilizing data from production-related IT systems offers the potential to support the value stream method with target-oriented analyses. Process mining (PM) supports the VSA by deriving process flows from production data as well as by analyzing process performances. Focused analyses of master data and transactional data enable reliable VSD activities without having to assume an oversimplified current state. This paper provides a framework for a continuously integrated data assistance within the value stream method, presenting a team structure, best practice procedures, and requirements for the application of the data assisted value stream method supported by examples from industry projects.

Keywords: Data Analytics · Process Mining · Value Stream Analysis · Value Stream Design

1 Introduction and State of the Art

The procedure of the value stream method starts with a product family selection, followed by VSA and VSD [1]. The method is typically carried out in a team consisting of a value stream manager, a facilitator, a coordinator, and a lean specialist and it is widely spread in the manufacturing industry to analyze and improve material and information flows [2]. In the VSA, a current state value stream map is derived in a shop floor inspection (gemba walk) and improvement potentials (kaizen) are identified. In the VSD, a waste-reduced future state is developed [3]. Following this approach makes a company's material and information flows transparent and plannable with comparatively little effort in a uniform and easily understandable modeling language [1].

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 660–669, 2021.
https://doi.org/10.1007/978-3-662-62138-7_66

However, these advantages are offset by deficits that need to be addressed: The modeling language is designed for simple, linear material flows but is less suitable to represent complex value streams with many branches or even several product families [4, 5]. Identifying the process steps, recording the correct data, and creating a target future state are not trivial and depend on the experience of the practitioners [6]. Another frequent criticism is that the value stream map is only a static representation based on a snapshot in which no dynamic or variability is recorded [4, 5]. Handling the captured key figures as deterministic values, however, is a huge simplification of the real situation [7]. Furthermore, recording times, stocks and key figures manually makes the VSA very time consuming and therefore costly to repeat [6]. There is a broad demand to further develop the value stream method given the possibilities that the progressive digitalization in productions provides [6].

Serrano et al. give the advice that data from Enterprise Resource Planning (ERP) systems should be used to facilitate the data collection in the VSA [2]. Recent literature suggests using PM - a discipline of data science – to map the current state value stream. PM transforms collected data into accessible information about processes. Discovery algorithms derive processes and their connections from event data while performance analyses give insights on process performance indicators. These techniques are carried out by data scientists and require event log data in sufficient quality [8]. A widely used methodology to follow in data mining projects is CRISP-DM [9]. Particularly for PM projects, the methodology PM^2 has been developed [10]. They both have an iterative procedure in common, including a view on the business context to understand and evaluate the input and output of the analysis. The application of PM in production and logistics is receiving more and more attention [11]. Since 2019, different publications discuss the application of PM in the context of the VSA. Ziegler at al. compare PM to the traditional VSA with the conclusion that it is the better way to depict a value stream map for the current state [12]. Klenk proposes to derive value stream maps via PM based on system data and to enlarge the focus of the analysis to IT processes [13]. Besides discovery, also performance analysis and conformance checking are useful to support the VSA, which is shown in a case study by Knoll et al. [14]. This study analyzes the spread of key figures for individual parts in an internal logistics value stream and checks the conformance of the event logs with given reference processes for a waste analysis. Urnauer and Metternich suggest to extend the VSA by PM to include an additional, neutral perspective into the team discussion [15]. However, Urnauer and Metternich as well as Klenk claim that the on-site inspection is an essential part of the VSA and must not be replaced in the data-assisted method [13, 15]. While all mentioned publications consider important aspects for data assistance in the value stream method, a higher-level framework is missing to collate the existing approaches in one method. Since the presented approaches are all either based on single use cases or theoretical concepts, broader practical experience needs to be gathered and included. Furthermore, requirements for the application need to be discussed.

Data assistance for the value stream method is enabled by the collection of data in production-related IT systems. According to Seitz et al., 75% of companies are already collecting feedback and operating data but only 52% are using it for analyses. This data has a huge potential to provide deeper insights into a company's operations. The same

applies to order data (73% collected, 51% analyzed) and stock data (61% collected, 33% analyzed) [16]. Since the field of data science is comparably new, not every company has data competences built up yet. However, much more training in this field is expected within the upcoming years. In 2018, 65% of manufacturing companies had competences in data analytics either already available or planned to build up [16].

2 The Data-Assisted Value Stream Method

Different publications have suggested the use of data at different points within the value stream method. It is the target of this paper to set a framework for a continuously integrated data assistance. Therefore, the data-assisted value stream method (dVSM) is being developed deductively and iteratively on VSA and VSD projects in the industry. This paper collates the existing approaches on data assistance in the value stream method and classifies them into the three fields cluster analysis, process mining and spread sheet calculation. Figure 1 illustrates how these categories (in the right) relate to the major activities (in the left) of the three steps of the value stream method. In the following, the dVSM is embedded into a framework with respect to the team constellation, a structure with best practice procedures, and the requirements that are necessary for a successful application of the method. The focus is on the data assisted current state analysis since the most experience could be gained in this field so far. However, the greatest potential is seen in the application of an entirely data-assisted value stream method.

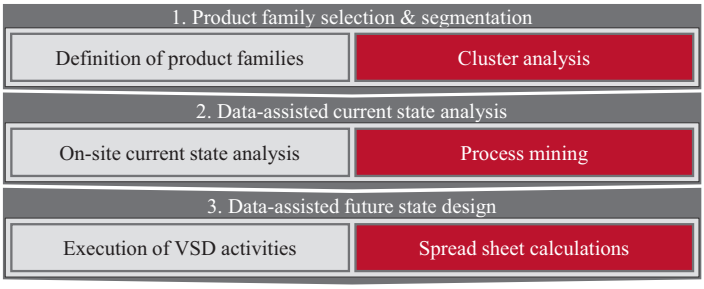


Fig. 1. Categorization of data analyses for the data-assisted value stream method

2.1 Team

Traditionally, the value stream method is carried out in a team of four roles. These are the *value stream manager*, who is responsible for the product family, the *facilitator*, who knows the production process best, a *coordinator*, who assists in the project by collecting the required data and managing the documentation, and a *lean specialist*, who gives guidance and training for lean methodologies [2]. For the execution of the dVSM, the additional role *data scientist* is introduced. The core of the interdisciplinary competence profile of data scientists is the ability to turn data into value [8]. Their job within the dVSM is to extract and transform data from different sources within the company, process it for the cluster analysis, for the application of PM as

well as for calculations supporting VSD activities and to drive the team discussions from a data point of view.

2.2 Procedure

1. Product family selection and segmentation:
For the segmentation of the value stream, product families are identified based on the similarity of their process flows – typically derived from a product-process matrix [3]. In a study on two industry data sets Gaida et al. illustrate that a cluster analysis facilitates and accelerates this step [17]. However, practical experience shows that the generated clusters must be discussed in the team to determine the number of distinct product families and to include additional aspects such as setup families into the decision. Two approaches of using cluster analyses can be distinguished: If the practitioners know how product families can be differentiated in their case, it is only an assignment problem. They name a representative product for each family and the clusters are built around them. If, however, they do not have an idea about how product families might look like in their use case, clustering is done in a more explorative manner with the sub-problem of finding the best number of clusters.

In the subsequent steps, the product family assignment is added to the data as a filter attribute, so that all products are taken into account simultaneously while providing the opportunity to filter for a specific product family. This way, design decisions consider e.g. shared resources correctly and re-assignments of product families can still be made at a more advanced point in the execution of the value stream method.

2. Data-assisted current state analysis:
In the current state analysis, processes and process connections are depicted in a value stream map and then enriched by process key figures and inventories. Figure 2 illustrates, how the PM activities process discovery and performance analysis are embedded in the procedure for the data-assisted current state analysis as tasks of the data scientist. Regular discussions with the entire team play a vital role in the analysis since they assist in creating a common understanding of the material and information flows as well as in identifying problems and possible kaizen in the value stream.

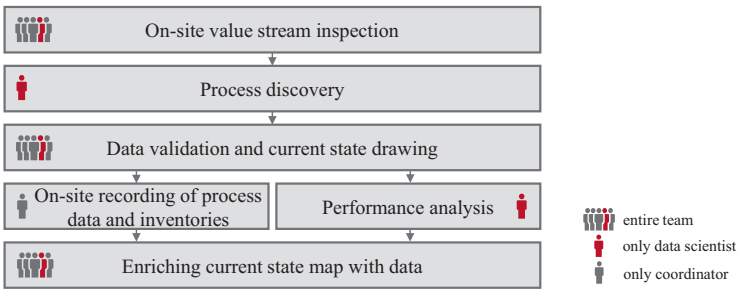


Fig. 2. Procedure of the data-assisted current state analysis

In a first step, the value stream of one representative product family is inspected in a gemba walk by the entire team. Then, the data scientist derives a process map from the transactional data of the production utilizing process discovery algorithms. This requires event log data in which each event represents a uniquely identified product, correlated to an activity and time stamps for start and end of this activity. This can e.g. be customer or production orders that were tracked in an ERP or Manufacturing Execution System (MES) [15].

The results of both, the on-site inspection and the process discovery are discussed in the team to draw a valid value stream map and to check the analyzed data for plausibility [15]. Additional check-ups for other product families can be useful at this point.

In the next step, a performance analysis is carried out on the discovered process model to derive key figures on process performances and inventories. Standard event log data allows for the identification of lead times, waiting times, process times, and synchronization times as well as various quantifications such as the number of processed orders per day [8]. Depending on the use case, the log can be enriched with additional data to fix data quality issues or to conduct more advanced analyses. In the common case that only one timestamp per activity is available, e.g. default process times can be used to heuristically approximate correct time values from the event log. To evaluate the validity of the analyses, a sample of the corresponding data is recorded manually on the shop floor. The resulting data is managed in spread sheets and aggregated figures are entered into the current state map.

Comparing the results of the traditional approach with the achieved results using data assistance, the advantages of the dVSM become clearly visible. Table 1 lists the key benefits of the dVSM in four industry applications. In two cases process steps would have been left out, in one case the sequence of processes was captured incorrectly, and in all cases the quantification could be significantly improved. In company A, the performance analysis uncovered that the customer demand was 43 % lower than initially estimated and that one product family had a 21 % lower share of the overall volume than expected. In company B it was discovered, that 25 % of the products considered were processed on a manual work station, although they should only be allocated to fully automated machines in this step. In companies C and D, the greatest benefit was in the possibility to differentiate the analyses. Additionally, recurring rework loops could be identified in company D with a major impact on the captured key figures. This demonstrates how the dVSM contributes to the avoidance of misjudgments.

Table 1. Key benefits of the data-assisted current state analysis in four industry projects

Company	Product or service	Key benefits of data assistance in VSA
A	Office supplies	• Discovery of additional process step • Quantification of actual customer demand and share of product families in branched value stream
B	Paints and varnishes	• Discovery of two additional process steps • Quantification of actual order-to-machine allocation
C	Logistic services	• Comparison of process performance and utilization for different locations, product families, daytime • Quantification of waiting times and process times
D	Pumps	• Discovery of correct process sequence • Quantification of waiting times on different production paths and for different product groups • Identification of rework loops

3. Data-assisted future state design

With this knowledge on hand, the future state value stream is designed, carrying out activity-specific calculations and analyses for each VSD step based on master data and transactional data from ERP or MES.

Rother and Shook structure the future state design with eight guiding questions [1]. However, as shown by Kaiser et al., VSD activities are highly influenced by each other [18]. A common data basis and interlinked calculation sheets enable an accurate consideration of all interdependencies. In a VSD project at a manufacturer for batteries, the following data assistance has proved valuable. Using the product-process matrix combined with master data for products and machines, load charts can be derived for each process. This supports the value stream segmentation as well as the capacity alignment. Building on this, heijunka sequences are defined and an EPEI (every part every interval) can be calculated which in turn is required for the quantification of supermarkets. For every new calculation, the required input data is first discussed with subject matter experts before the calculations are carried out to ensure its validity and to avoid extra work. The resulting interlinked spread sheets facilitate the assessment of design alternatives or kaizen. By varying the affected input factors, the impact on the entire value stream quantification can be calculated with low effort. In the mentioned project, e.g. the prioritization of changeover improvement projects was assessed by calculating scenarios with a decreased changeover time and an increased overall equipment effectiveness for the considered machines which directly revealed the impact on the available capacity, the achievable EPEI, and the required inventories.

The key benefits of the data assistance in the VSD project were:

- Quantitative assessment of design decisions and kaizen ideas
- Better understanding for interdependencies in the value stream and the impact of design decisions
- Lower implementation effort through reliable quantifications and visual support for VSD activities
- Faster repetition of VSD activities using the created data structure

2.3 Requirements

To point out limitations of the dVSM, requirements for the successful application are identified from industry projects in the four categories data, data preparation, competences, and method.

The *data* must be available in the right quantity (i.e. completeness) and quality. Using PM techniques, e.g. granularity and correctness of time stamps have a major impact on the quality of the outcomes as well as on the effort that needs to be put into the data preparation [11].

The *data preparation* is the most time-consuming part of the data scientist's tasks so that sufficient working time needs to be scheduled. Depending on the use case, it may be necessary to handle large data volumes. In one application of the method at a logistic service provider, e.g. monthly data exports had to be processed separately, which lead to a high time effort. Furthermore, the addressed processes need to be known well to extract the correct data and transform it accordingly. For this reason, it is important that the data scientist participates in the on-site analysis.

To carry out the dVSM most effectively, different *competences* need to be available. Know how on the dVSM itself, the data and the organization is important. Therefore, all introduced roles need to be included. The interdisciplinarity of the team increases its creativity and its problem-solving competence.

To ensure transparency and the acceptance for decisions, stakeholders must be involved on a regular basis. In addition, requirements for the *method* itself can be identified. As described in the procedure, a joint interpretation of the results is necessary for plausibility checks and to derive sound decisions. Therefore, a visual preparation of the data analyses is needed. A close collaboration among the team members enables the integration of different perspectives. To leverage the data perspective, it is especially important to achieve a close collaboration of the data scientist with the rest of the team. For an efficient application, the method must have a clear integration concept for the data assistance, which is given by the role definition and the procedure in Sects. 2.1 and 2.2. Eventually, the mindset of the executing employees has a major impact on the success of the resulting future state design so that their involvement should be of high priority.

Since the integration of data assistance into traditional methods has been the subject of previous research, a systematic literature review is used to reflect the so far identified requirements. According to Schryen et al., this tool has proven its worth in the discipline of information systems [19]. The following research steps are based on a procedure by Kitchenham [20]. A manual search is carried out on the three platforms "IEEE Xplore", "Web of Science" and "Science Direct". Therefore, the following search line, considering technical and sociological aspects, is used:

(Business Analytics) AND (Process Mining) AND (Requirement OR Acceptance OR Decision Process OR Management)

This leads to a set of 394 publications, which is reduced by discarding duplicates (334 papers left) as well as filtering by title (53 papers left), abstract (21 papers left), and availability (18 papers left). Papers that did not apply a scientific research procedure were sorted out as well (16 papers left). The requirements extracted from these papers are summarized in Table 2. The findings show that in existing data-assisted approaches in literature, the same requirements as discussed for the dVSM are found. This confirms the scope and relevance of the identified requirements.

Table 2. Results of the systematic literature review compared to dVSM

		Requirements																		
		Data			Data Preparation			Competences					Method							
Author (year)	**Type**	Data Quality	Accessability	Availability	Handling of big data volume	Focus on decisions	Sufficient working time	Data knowledge	Organizational knowledge	Stakeholder involvment	Interdisciplinarity	Method knowledge	Interpretation of results	Close collaboration	Visualization of results	Integration concept	Open mindset			
Bole et al. (2015) [21]	C					●			●	●	●			●						
Caron et al. (2013) [22]	C					●									●	●				
Jans et al. (2013) [23]	C	●				●			●				●		●					
Lee and Chen (2016) [24]	L					●		●		●			●	●	●	●	●			
Lu (2018) [25]	C	●				●		●		●			●							
Trieu (2017) [26]	L								●											
Kumar and Belwal (2017) [27]	L				●			●		●	●		●	●	●	●				
Groggert et al. (2017) [28]	E	●		●				●	●				●							
Yamada and Peran (2017) [29]	L					●	●			●				●						
Cognini et al. (2014) [30]	L		●			●	●				●	●	●	●	●	●				
Sahu et al. (2017) [31]	C	●		●		●		●					●		●	●				
Sakr et al. (2018) [32]	L					●														
Mishra et al. (2016) [33]	C			●			●						●			●				
Rajpurohit (2013) [34]	L	●			●	●		●	●	●	●			●	●	●	●			
Proctor et al. (2011) [35]	L	●				●		●	●			●	●	●		●				
Sanjay and B. H. (2016) [36]	L	●			●								●			●				
dVSM		●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●			

C = Case Study; L = Literature Review; E = Empirical Study

3 Conclusion and Outlook

This paper introduces the data-assisted value stream method dVSM. Supported by industry case examples, the procedure as well as a team structure are presented and requirements for the successful application are derived and compared with the literature. The exemplary use cases demonstrate how the dVSM overcomes the identified weaknesses of the value stream method with the following strengths:

- Multiple product families can be analyzed and designed simultaneously
- Dynamics and variability over time are captured using historical data
- Depicted value stream map and the associated key figures are of higher quality
- Data analyses enable better quantification and visual support for VSD activities
- Possibility to create scenarios facilitates the prioritization of kaizen activities
- Interlinked calculations increase the understanding for interdependencies and enable an integrated planning of the future state
- Possibility to update the data basis in the created data structure accelerates the repetition of analysis or design activities

The identified requirements indicate limitations of the dVSM. The two major success factors for its application are the availability of data and the competences to analyze it. The depiction of the analyses can only be as good as the underlying data. Therefore, the available data quality constitutes a limitation to the approach. Furthermore, in industry applications of the dVSM it could be observed that practitioners tend to get lost in very specific in-depth analyses, shifting the focus from the rather qualitative and holistic view of the value stream method towards detailed quantifications.

In the next steps, the dVSM will be continuously advanced and validated in industry applications. The focus of further research is on the development and standardization of a continuous data structure and further analysis formats for general applicability.

References

1. Rother, M., Shook, J.: Learning to see: mapping the value stream to add value and eliminate waste. Lean Institute, Cambridge (1999)
2. Serrano, I., Ochoa, C., Castro, R.D.: Evaluation of value stream mapping in manufacturing system redesign. Int. J. Prod. Res. **46**(16), 4409–4430 (2008)
3. Erlach, K.: Value Stream Design. Springer, Berlin (2013)
4. Forno, A.J.D., Pereira, F.A., Forcellini, F.A., Kipper, L.M.: Value stream mapping: a study about the problems and challenges found in the literature from the past 15 years about application of Lean tools. Int. J. Adv. Manuf. Technol. **72**(5–8), 779–790 (2014)
5. Shou, W., Wang, J., Wu, P., Wang, X., Chong, H.-Y.: A cross-sector review on the use of value stream mapping. Int. J. Prod. Res. **55**(13), 3906–3928 (2017)
6. Winkler, H., Lugert, A.: Die Wertstrommethode im Zeitalter von Industrie 4.0 - Studienreport. BTU Cottbus - Senftenberg. Cottbus (2017)
7. Luz, G.P., Tortorella, G.L., Narayanamurthy, G., Gaiardelli, P., Sawhney, R.: A systematic literature review on the stochastic analysis of value streams. Prod Plan Control, pp. 1–11 (2020)
8. van der Aalst, W.: Process Mining: Data Science in Action, 2nd edn. Springer, Heidelberg (2016)
9. Shearer, C.: The CRISP-DM model: the new blueprint for data mining. J. Data Warehouse. **5**(4), 13–22 (2000)
10. van Eck, M.L., Lu, X., Leemans, S.J.J., van der Aalst, W.M.P.: PM2: A Process Mining Project Methodology. Advanced Information Systems Engineering, pp. 297–313. Springer, Cham (2015)
11. Urnauer, C., Gräff, V., Metternich, J.: Process Mining in einer Produktion mit kundenanonymer Bevorratung: Heuristischer Ansatz zur Vergabe von Case IDs. TU Prints, (2020)
12. Ziegler, S., Braunreuther, S., Reinhart, G.: Process Mining zur dynamischen Wertstromaufnahme. ZWF **114**(6), 327–331 (2019)
13. Klenk, E.: Konzept zur systemdatenbasierten Wertstromanalyse: Generierung von Wertströmen mittels Process Mining. ZWF **114**(9), 513–516 (2019)
14. Knoll, D., Reinhart, G., Prüglmeier, M.: Enabling value stream mapping for internal logistics using multidimensional process mining. Expert Syst. Appl. **124**, 130–142 (2019)
15. Urnauer, C., Metternich, J.: Die digitale Wertstrommethode: Process Mining als digitale Stütze der Wertstromanalyse. ZWF **117**(12), 855–858 (2019)

16. Seitz, M., Härtel, L., Hübner, M., Merkel, L., be Isa, J., Egenhausen, F., Schmidhuber, M., Sauermann, F., Hünnekes, P.: PPS-Report 2017/2018. ZWF **113** (12), pp. 840–844 (2018)
17. Gaida, M., Günther, G., Wilsky, P., Riedel, R.: Bildung von Produktfamilien als Planungsgrundlage auf Basis von Clusteralgorithmen. ZWF **115**(3), 111–114 (2020)
18. Kaiser, J., Urnauer, C., Metternich, J.: A framework for planning logistical alternatives in value stream design. Procedia CIRP **81**, 180–185 (2019)
19. Schryen, G., Benlian, A., Rowe, F., Gregor, S.D., Larsen, K.R.: Literature reviews in is research: what can be learnt from the past and other fields? CAIS **41**, 557–569 (2017)
20. Kitchenham, B., Brereton, P.: A systematic review of systematic review process research in software engineering. Inform. Softw. Tech. **55**, 2049–2075 (2013)
21. Bole, U., Popovic, A., Zabkar, J., Papa, G., Jaklic, J.: A case analysis of embryonic data mining success. Int. J. Inf. **35**(2), 253–259 (2015)
22. Caron, F., Vanthienen, J., Baesens, B.: A comprehensive investigation of the applicability of process mining techniques for enterprise risk management. Comput. Ind. **64**(4), 464–475 (2013)
23. Jans, M., Alles, M., Vasarhelyi, M.: The case for process mining in auditing: Sources of value added and areas of application. Int. J. Account Inf. Syst. **14**(1), 1–20 (2013)
24. Lee, R., Chen, I.-Y.: A novel production process modeling for analytics. Int. J. Geomate. **11**(24), 2370–2377 (2016)
25. Lu, J.: A data-driven framework for business analytics in the context of big data. New Trends Databases Inf. Sys. 339–351 (2018)
26. Trieu, V.-H.: Getting value from business intelligence systems: a review and research agenda. Dec. Support Syst. **93**, 1–34 (2017)
27. Kumar, S.M., Belwal, M.: Performance dashboard cutting-edge business intelligence and data visualization. Smart Tech. Con. 1201–1207 (2017)
28. Groggert, S., Wenking, M., Schmitt, R.H., Friedli, T.: Status quo and future potential of manufacturing data analytics. In. C Ind. Eng. Eng. Man. 779–783, (2017)
29. Yamada, A., Peran, M.: Governance framework for enterprise analytics and data. Int. Conf. Big. Data. 3623–3631 (2017)
30. Cognini, R., Corradini, F., Polzonetti, A., Re, B.: Five factors that make pervasive business intelligence a winning wager. In. C Ind. Eng. Eng. Man. 617–621 (2014)
31. Sahu, S.K., Jacintha, M.M., Singh, A.P.: Comparative study of tools for big data analytics: An analytical study. ICCCA 2017, 37–41 (2017)
32. Sakr, S., Maamar, Z., Awad, A., Benatallah, B., van der Aalst, W.: business process analytics and big data systems: a roadmap to bridge the gap. IEEE Access **6**, 77308–77320 (2018)
33. Mishra, B.K., Hazra, D., Tarannum, K., Kumar, M.: Business intelligence using data mining techniques and business analytics. SMART, 84–89 (2016)
34. Rajpurohit, A.: Big data for business managers - Bridging the gap between potential and value. Int. Conf. Big Data, 29–31 (2013)
35. Proctor, L., Kieliszewski, C.A., Hochstein, A., Spangler, S.: Analytical pathway methodology: simplifying business intelligence consulting. SRII Glob. Conf. 495–500 (2011)
36. Sanjay, M., Alamma, A.: An insight into big data analytics - Methods and application. ICICT, 1–5 (2016)

Definition of Process Performance Indicators for the Application of Process Mining in End-to-End Order Processing Processes

S. Schmitz(✉), F. Renneberg, S. Cremer, A. Gützlaff, and G. Schuh

Laboratory for Machine Tools & Production Engineering (WZL), RWTH Aachen University, Aachen, Germany
S.Schmitz@wzl.rwth-aachen.de

Abstract. Todays' increasing market volatility and product variety result in growing business processes complexity. To master the arisen challenges of high process complexity, the process performance description of the end-to-end order processing process (ETEOPP) is crucial. With the trend of digitalization in manufacturing companies, an increasing availability of data is created that can be used to master process complexity by data-based methods. One suitable method is process mining (PM), which offers a continuous analysis of event data from business information systems. This paper aims to describe the minimum viable dataset to thoroughly evaluate the process performance in ETEOPP by PM. Therefore, process performance indicators (PPI) are first scientifically derived by a systematic literature review and afterward defined across the ETEOPP as quantifiable parameters based on processes and data. By doing so, the required event log attributes, as well as corresponding data requirements, are presented and information systems for data extraction are pointed out.

Keywords: Process Performance Indicators · Process Mining · End-To-End Order Processing

1 Introduction

Manufacturing companies are challenged by increasing market dynamics and uncertainties. To withstand the highly competitive pressure from the market, high efficiency across the end-to-end order processing process (ETEOPP) is crucial [1], as the ETEOPP comprises all technical-operative core processes to fulfill customers' orders.

The efficiencies of manufacturing companies are based on processes and are measured by process performance indicators (PPIs) [2]. To evaluate the decisive process efficiency of organizations, PPIs must be defined and measured across the entire ETEOPP [3, 4].

In order to describe processes and their performance, traditional methods, such as workshops or interviews, are often used. However, those methods are subjective and

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 670–679, 2021.
https://doi.org/10.1007/978-3-662-62138-7_67

unsubstantiated, wherefore the described process performance is based on gut feelings, static mean values and low scatter [5, 6]. Since actual process complexity and process variance is intricate, real as-is process behavior is barely reflected by traditional methods. Thereby, process complexity can be defined as the degree of which a process is difficult to analyze, understand and explain due to high uncertainty, activity interfaces and interdependencies [7].

The rapidly growing amount of data in manufacturing companies can be used to gain more transparency in order to simplify process complexity and to describe process variances [8]. Due to the weaknesses of traditional methods, process mining (PM) became a broad research field in recent years [6]. PM represents an objective and effortless method to evaluate process performance based on facts. For this purpose, PM applications generate knowledge from already available event logs that reflect the behavior of real process performance in a scheduled manner [8]. Therefore, the required data for measuring the PPIs across the entire ETEOPP must be included in the event log. However, existing approaches merely consider minimum required data for PM applications (i.e. case-ID, activity and timestamp) or only several indicator-related data [9]. Thus, this paper presents an approach to describe PPIs across the entire ETEOPP for the application of PM.

The remainder of this paper is structured as follows: Sect. 2 highlights the theoretical basics – in detail PM, ETEOPP and process performance indicators. Afterwards, related work is presented to summarize existing approaches. Next, PPIs are consolidated by a systematic literature review and defined methodically. The results are presented in Sect. 5 and finally discussed in Sect. 6.

2 Theoretical Basics

2.1 Process Mining

PM is defined as the discovering, monitoring and improvement of real processes by extracting data from event logs readily available in today's business information systems. The goal of PM is to convert event data into insights and actions and thus describe processes, interrelationships, bottlenecks and process variants. [9] The process visualization (i.e. process models) resulting from PM can be used as an objective starting point or as a feedback mechanism to optimize business processes. The essential requirements for fact-based process descriptions with PM are event logs [6]. Event logs are a collection of event data that is used as input data for PM [10]. The bare minimum of an event log contains case-ID (i.e. orders in the given context) as well as its activities and timestamps. The respective activities are listed chronologically in the event log to identify the causal relationships in the generated process model. Furthermore, depending on the availability of data, event logs can be enriched by additional attributes [9]. The process models are created from the event logs using process discovery algorithms [11].

2.2 End-to-end Order Processing Process

In this paper, the understanding of the ETEOPP is derived from [12] and the German word "Auftragsabwicklung" and defined as *the core process of producing companies that converts incoming inquiries and orders from external customers into internal specifications, orders and courses of action.* Different order types characterize the ETEOPP. These types are make-to-stock, build-to-order, make-to-order and engineer-to-order. The order types are defined according to the customer-order decoupling point, i.e. the change between order-anonymous and order-related processes. [13] Referring to the customer-order decoupling point, the ETEOPP starts when the order-related processes are triggered by external customers and ends when external customer's needs are met. The ETEOPP consists of different sub-processes [14], which often represent particular functional units in industrial practice. The sub-processes subsist of different activities and tasks. In this paper, the ETEOPP is defined according to Luczak et al., which is subdivided into the sub-processes sales, design, procurement, work preparation, production and shipping [15].

2.3 Process Performance Indicators

The performance of sub-processes and its activities is precisely and consistently measured by performance indicators [16]. In technical literature, a distinction is made between key performance indicators (KPI) and PPIs. KPIs pursue goals at the tactical and strategic levels and PPIs at the operational level [17]. Accordingly, KPIs are defined as quantifiable and strategic indicators that reflect the critical success factors of a company [18]. PPIs, on the other hand, are defined as quantifiable indicators that evaluate business process efficiency and effectiveness [18]. Efficiency is defined resource-oriented, i.e. how well resources are used to produce an output. Effectiveness is defined target-oriented, i.e. how well process outputs are met [19]. According to Del-Rio-Ortega, PPIs need to be defined as specific, measurable, achievable, relevant and time-bounded (satisfying the SMART criteria) metrics and can be seen as a particular case of KPI [20]. In this paper, the following meaning is considered: *Specific* makes clear what the PPI exactly describe. *Measurable* provides how the PPI is calculated and what data (i.e. variables and constants) are required for calculation. *Achievable* ensures that a PPI value can be met to pursue a goal. *Relevant* confirms that the PPI affects organizational performance. As relevance strongly depends on organizational goals, it is assumed in the paper that the researched performance indicators are relevant in general. *Time-bounded* specifies that a PPI is directly dependent on time and only has a meaning if the measured time period is known.

3 Related Work

In this chapter, existing approaches for defining PPIs for the application of PM along the ETEOPP are examined in more detail. The existing approaches are evaluated based on requirements that are derived from Sect. 2. First, the PPIs should cover

the ETEOPP, i.e. from the placing of a customer order, through the manufacturing and assembly, until the order is shipped to the customer. Second, the PPIs must be defined process-related and according to the SMART criteria (see Sect. 2.3) [20]. In particular, PPIs are differentiated from KPIs by time dependency and only indicators that directly change over time intervals are defined as PPI. Third, the required data to describe the PPIs by PM must be defined by the minimum event log attributes (i.e. case-ID, activity, timestamp) or as additional event log attributes. The third requirement is completely fulfilled, if the required event log attributes are fully mentioned. Last, the approach should be generically adaptable for practical implementation. In this context, generically adaptable means that the performance indicators are applicable for different use cases (i.e. application-neutral and valuation-neutral) and not related to a specific application case. So far, numerous authors have defined process-related KPIs or PPIs that only fulfill some of the mentioned requirements.

Nenadál lists numerous process-related KPIs that are relevant for producing companies. The KPIs are categorized universally and process-specific [21]. However, only product development and procurement are listed as processes and the ETEOPP is not covered entirely. Therefore, the first requirement is not sufficiently focused. The KPIs are specifically described for the process-specific KPIs (e.g. time to release drawings and specifications). The KPIs are achievable if an upper or lower boundary is defined. Yet, the KPIs are just described literally and not measurably. Furthermore, not all KPIs are time-bounded (e.g. nonconformities ration or input yield) and the authors do not differentiate between KPI and PPI. In consequence, the performance indicators are not defined according to the SMART criteria. No event log data attributes are described wherefore the third requirement is assessed as not covered. As the KPIs are not quantified and the required data to calculate the performance indicators for PM applications are not described precisely, the fourth requirement is partially fulfilled. Wetzstein et al. define KPIs to control achievements of process objectives and to measure process performance [22]. By doing so, the authors give an abstract approach to measure holistic PPIs according to the SMART criteria that are measurable across the entire ETEOPP. Thus, the first and second requirements are covered. Additional data, such as case-ID and activity, are not defined in order to describe the KPIs by PM, which means the third requirement is partially covered. Although the KPIs can be determined for different use cases, it is not specifically described for PM applications, how and from which event log attributes the KPIs are calculated. As a result, the last requirement is just rudimentarily fulfilled. Gladen presents selective KPIs that can be assigned to sub-processes of the ETEOPP (e.g. adherence to delivery dates) but also sub-process-specific KPIs, such as production throughput time, which are only useful for partial processes [23]. Thus, the first requirement is just covered rudimentarily. Like Nenadál, the KPIs are described specifically and achievable but not measurable as well as time-bounded. All KPIs are just described literally and without calculations for quantification. Similar, none event log attributes are described for PM application. The third requirement is not fulfilled. Although the KPIs are initially suitable for many applications due to their literal formulation, their imprecision, lack of quantification and lack of data necessities make them difficulty adoptable for industrial practice. Consequently, the last requirement is only weakly covered. Lanza presents an evaluation framework for operative KPIs [24], which just partially covers

KPIs that are measurable for the entire ETEOPP and therefore fulfills the first requirement just rudimentarily. The SMART criteria are not completely focused, as not all KPIs are time-bounded (e.g. the KPI quality) and it is not specified for all KPIs, how exactly they are calculated and what data are required. No event log data attributes are described and the third requirement is not addressed. Similar to Nenadál and Gladen, the last requirement is not met. In ISO 22400, numerous process-related KPIs are defined and established as a standard for manufacturing companies [16]. Due to its manufacturing focus, most of the KPIs only refer to production processes. Within the framework of ISO 22400, the KPI attributes are described in detail according to the SMART criteria. However, just like Wetzstein et al., additional data are not defined in order to describe the KPIs by PM. Therefore, the third requirement is only partially covered. In addition, the fourth requirement is just rudimentarily covered, as it is not specified how the KPIs can be calculated from event log attributes for PM applications.

	Requirements to describe PPI for the application of PM in ETEOPP			
	The PPIs cover the entire ETEOPP	The PPIs are defined process-related and according to the SMART criteria	The event log attributes to describe the PPIs by PM are defined	The PPIs are generically adaptable for PM applications in practice
Nenadál (2008)	◑	◑	○	◔
Wetzstein et al. (2008)	●	●	◔	◑
Gladen (2014)	◑	◑	○	◔
Lanza (2015)	◑	◑	○	◔
ISO 22400 (2017)	◔	●	◔	◑

○ : Not covered ◔ : partially covered ◑ : rudimentarily covered ◕ : significantly covered ● : fully covered

Fig. 1. Requirement assessment and research potential of related work.

As result, the state of the art shows that none of the authors have systematically derived requirements for PM to measure PPIs across the ETEOPP (see Fig. 1). Deficits lie in the fulfillment of the data needs for PM applications and the clear distinction between PPIs and KPIs since the authors use these terms interchangeably or do not consider the SMART criteria. As result, there is still research potential to define PPIs for the application of PM along the ETEOPP.

4 Research Methodology

In this paper, PPIs are first consolidated through a systematic literature review (SLR) based on guidelines proposed by Kitchenham [25]. Thus, the SLR is divided into three parts: planning the review, conducting the review and analyzing the results. Subsequently, additional methodical steps are defined to holistically derive and describe the PPIS for a PM application along the ETEOPP. Figure 2 shows the structure of the research methodology.

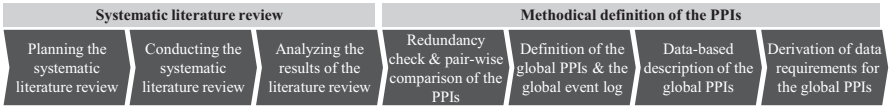


Fig. 2. Structure of the research methodology.

4.1 Conducting the Systematic Literature Review

A designed and detailed protocol for the description of the PPIs was initially defined during SLR-planning. The protocol contained the research questions, the search process, the inclusion criteria (IC) and exclusion criteria (EC), the data extraction and the synthesis of the results. To examine whether existing literature answers the research questions, a search string in English and German was used based on the three steps: (i) identifications of keywords, (ii) definition of synonyms based on relevant studies about PPIs, (iii) use of boolean operator "OR" between synonyms and boolean operator "AND" to connect keywords. The search string consists of the three object areas PPI, ETEOPP and PM. Since authors often do not distinguish between PPI and KPI, KPI was included in the search string as synonym for PPI. In addition, the word "ETEOPP" was replaced in the German search string by the word "Auftragsabwicklung". Based on KITCHENHAM, relevant studies were selected based on IC and EC. Respectively, the inclusion of a study is sequentially determined by the relevance of title, abstract, introduction, summary and complete study. The planning of the SLR for this paper is shown in Fig. 3.

Research Questions	
Q1: Have PPIs already been defined along the ETEOPP for producing companies according to the definition in chapter 2?	
Q2: Are these PPIs applicable for PM?	
Databases	
▪ ScienceDirect (www.sciencedirect.com) ▪ ResearchGate (www.researchgate.net) ▪ Springer (www.link.springer.com)	▪ IEEE Xplore (www.ieeexplore.com) ▪ Google Scholar (www.scholar.google.com)
Search String	
English: (i) Process performance indicator, end-to-end order processing process, process mining German: (ii) Process performance indicator, Auftragsabwicklung, process mining	English and German: (iii) Key performance indicator
English: (iv) (Process performance indicator OR key performance indicator) AND (end-to-end order processing process) AND (process mining) German: (v) (Process performance indicator OR key performance indicator) AND (Auftragsabwicklung) AND (process mining)	
Inclusion Criteria (IC) & Exclusion Criteria (EC)	
▪ IC 1: Primary study that defines PPIs for the ETEOPP ▪ IC 2: Primary study that presents PPIs in the research field PM	▪ EC 1: Primary study that is not available in electronic format ▪ EC 2: Primary study that is repeated ▪ EC 3: Primary study that is incomplete (text, content and incomplete results) ▪ EC 4: Primary study that is not written in English or German ▪ EC 5: Primary study that is not full paper or short paper

Fig. 3. Planning the SLR to consolidate the PPIs.

Overall, 363 primary studies with 60 in ScienceDirect, 312 in ResearchGate (only the first 100 were taken), 2017 in Springer (only the first 100 were taken), 3 in IEEE and 26300 in Google Scholar (only the first 100 were taken) were identified. Conducting IC and EC result in a total of 27 studies. The studies were further reduced by application-specific requirements (see Sect. 3), mutual adaption of PPIs and cross-references. Finally, five studies remain to cover all relevant PPIs.

The results of the SLR are shown in Fig. 4. The PPIs are consolidated from the identified primary studies based on the definition of PPIs according to Sect. 2.3.

4.2 Methodical Definition of the PPIs

As a first methodical step, the preselected PPIs from Fig. 4 were checked for redundancy. One PPI is redundant if it can be replaced or described by another PPI. The redundancy was checked combinatorial in a matrix, with the preselected PPIs as rows and columns. The remaining, non-redundant PPIs are separated in global and sub-process-specific PPIs, using pairwise comparison. By definition, global PPIs can be described for all sub-processes of the ETEOPP, whereas sub-process-specific PPIs are only applicable to individual sub-processes. The global PPIs are defined as quantifiable metrics and the event log attributes for a data-based description of the global PPIs are specified. Finally, the data needs are derived from business information systems. Hereto, the event log attributes were mirrored to the individual sub-processes and the process-related information systems. As result, the required event log data can be extracted from the corresponding information system for a data-based description of PPI across the ETEOPP.

Nenadál (2008)	1. Cycle time (p); 2. Time of response; 3. Time to release drawings and specifications (d); 4. Rework and delay time caused by design failures (d); 5. Production delay caused by late deliveries (pc)
Wetzstein et al. (2008)	6. Process Duration; 7. Deadline Adherence
Gladen (2014)	8. Adherence to delivery dates (r); 9. Lead time procurement process (pc); 10. Delivery service level towards production (r); 11. Delivery time (s); 12. Delivery reliability (s); 13. Throughput performance (p); 14. Capacity utilization of the logistics (pc); 15. Production throughput time (p); 16. Adherence to delivery dates (r); 17. Capacity utilization of the machines (p)
Lanza (2015)	18. Supplier punctuality (r); 19. Lead time (p); 20. Total production time (p); 21. Mean tardiness; 22. Process reliability; 23. Capacity utilization (r); 24. Mean time to repair (p)
ISO 22400 (2017)	25. Worker efficiency (p); 26. Allocation ratio (p); 27. Throughput rate (p); 28. Utilization efficiency (p); 29. Availability (p); 30. Effectiveness (p); 31. Setup ratio (wp); 32. Technical efficiency (p); 33. Production process ratio (p); 34. Mean operating time between failures (p); 35. Mean time to failure (r); 36. Mean time to repair (r)

Legend:
(r): PPI reduced due to redundancy; (d): PPI reduced due to design-specificity; (pc): PPI reduced due to procurement-specificity; (wp): PPI reduced due to work-preparation-specificity; (p): PPI reduced due to production-specificity; (s): PPI reduced due to shipping-specificity

Fig. 4. Consolidated PPIs from SLR [16, 21–24].

5 Results of the Methodical Definition of PPIs

During redundancy check, the PPIs adherence to delivery dates (8th and 16th), delivery service level towards production (10th) and supplier punctuality (18th) were assessed as redundant to deadline adherence (7th) and reduced from the consolidated PPI. Similarly, capacity utilization of the machines (17th) and capacity utilization (23rd), process reliability (22nd) and mean time to failure (35th) as well as mean time to repair (24th) and mean time to repair (36th) were assessed as redundant (see. Fig. 4) that leaves a total of 29 PPIs. Next, the relationship between the PPIs was examined using a pairwise comparison to determine whether the PPIs were globally or sub-process-specifically. By doing so, the PPIs were reduced by design-specific PPIs (3rd and 4th), procurement-specific PPIs (5th, 9th, 11th and 14th), work-preparation-specific PPI (31st), production-specific PPIs (1st, 13th, 15th, 17th, 19th, 20th, 24th, 25th, 26th, 28th, 29th, 30th, 32nd, 33rd, 34th and 35th) and shipping-specific PPIs (12th). Overall, the PPIs were merged to five global PPIs using the redundancy check and the pairwise comparison. These are the process duration (PPI_1), the time of response (PPI_2), the deadline adherence (PPI_3), the mean tardiness (PPI_4) and the process reliability (PPI_5). The event log attributes can be derived from the data needs of the quantifiable global PPIs (see Fig. 5). Accordingly, the minimum required data of an event log are enriched by additional attributes for a data-based description of the PPIs along the ETEOPP (marked with an X in Fig. 5.). Finally, the sub-process-specific business information systems that cover the attribute-specific data needs of the global event log are derived. Since the business information systems are used differently in practice, they cannot be uniquely and clearly selected. Therefore, the assigned business information systems in Fig. 5 can be seen as an exemplary selection.

6 Discussion and Outlook

In this paper, the definition of PPIs across the ETEOPP for the application of PM was presented. The PPIs were consolidated by a SLR and then methodically defined based on the application-specific requirements for PM. After a redundancy check and a pairwise comparison, five global PPIs were identified. The PPIs process duration (PPI_1), time of response (PPI_2), deadline adherence (PPI_3), mean tardiness (PPI_4) and process reliability (PPI_5) led to the definition of global event log attributes. In the last step, the data needs were derived from the business information systems using the event log attributes. Thus, the approach enables a data-based description of process performance across the entire ETEOPP by PM, which offers considerable potential in the process optimization of producing companies.

The approach can be further honed by several improvements. First, the ETEOPP was derived by comparing different reference processes in literature and can be further validated beyond one industrial company. Additionally, ETEOPP and PPIs can vary in practice based on different decoupling points. The approach should be validated to investigate, if in practice, all attributes in the respective sub-processes of the ETEOPP have clear semantics and can be derived without ambiguity from business information systems. Additionally, compared to Fig. 5, the derived PPIs should be specified

to determine, if they should be calculated order-related or activity-related. Finally, although PM was already applied to ETEOPP in previous approaches [1], the presented approach offers the potential to consider additional PPIs. Thereby, existing PM discovery algorithms can be enhanced in information science to evaluate the PPIs. Finally, besides the application of PM for a simplified description of process complexity in ETEOPP, other strategies to reduce or avoid process complexity can be investigated.

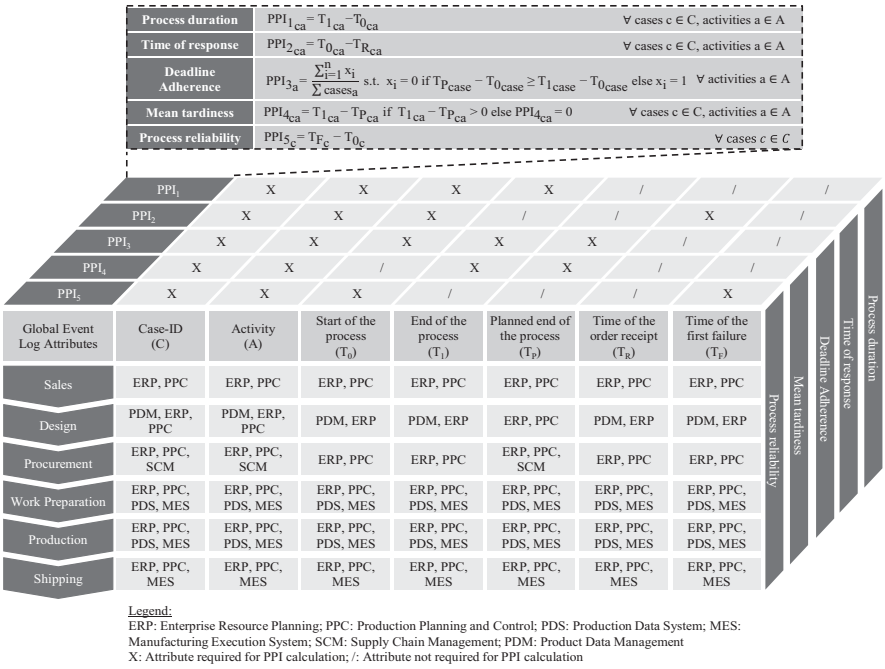


Fig. 5. Global event log attributes and data requirements (X) across the ETEOPP.

Based on this approach, a methodology for the application of PM across the ETEOPP should be further researched, which offers a guideline for producing companies to increase their process performance.

Acknowledgement. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

1. Schuh, G., Gützlaff A., Schmitz S., van der Aalst WMP.: Data-based description of process performance in end-to-end order processing. CIRP Ann. Manuf. Technol. **69**(1), (2020)
2. Schmelzer, H., Sesselmann, W.: Geschäftsprozessmanagement in der Praxis, p. 2. Hanser, München (2013)

3. Tsutsumi, D., Gyulai, D., Kovács, A., Tipary, B., Ueno, Y., Nonaka, Y., Monostori, L.: Towards joint optimization of product design, process planning and production planning in multi-product assembly. CIRP Ann. **67**(1), 441–446 (2018)
4. Schönsleben, P., Weber, S., Koenigs, S., Duchi, A.: Different types of cooperation between the R&D and Engineering departments in companies with a design-to-order production environment. CIRP Ann. **66**(1), 405–408 (2017)
5. Knoll, D., Reinhart, G., Prüglmeier, M.: Enabling value stream mapping for internal logistics using multidimensional process mining. Expert Syst. Appl. **124**, 130–142 (2019)
6. Van der Aalst, W. M. P.: Process Mining Overview and Opportunities. p. 1 (2012)
7. Cardoso, J.: How to measure the control-flow complexity of web processes and workflows. Future Strategies Inc., Lighthouse Point (2005)
8. Reinkemeyer, L.: Process Mining in Action. Principles, Use Cases and Outlook. Springer, Switzerland (2020)
9. Van der Aalst, W.: Process Mining. p. 128. Springer, Berlin (2016)
10. Nguyen, H., Dumas, M., ter Hofstede, H. M., La Rosa, M., Maggi, F. M.: Mining Business Process Stages from Event Logs. p. 3. Queensland (2016)
11. Medeiros, A. K. A. de; Weijters, A. J. M. M.; van der Aalst, W. M. P.: Genetic process mining: an experimental evaluation. Data Min. Knowl. Disc. **14**(2), 247. Eindhoven
12. Wildemann, H.: Auftragsabwicklung in einer computergestützten Fertigung (CIM). In: Albach, H. (eds.) Industrielles Management, pp. 198–223. Gabler, Wiesbaden (1993)
13. Olhager, J.: Strategic positioning of the order penetration point. Int. J. Prod. Econ. **85**, 319–329 (2003)
14. Schuh G., Gützlaff A., Cremer S., Schopen M.: Understanding process mining for data-driven optimization of order processing. In: 10th Conference on Learning factories (2020)
15. Luczak, H., Becker, J.: Workflowmanagement in der Produktionsplanung und -steuerung. Qualität und Effizienz der Auftragsabwicklung steigern. p. 17. Springer, Berlin (2003)
16. Varisco, M., Johnsson, C., Mejvik, J., Schiraldi, M. M., Zhu, L.: KPIs for Manufacturing Operations Management: driving the ISO22400 standard towards practical applicability, p. 1. (2018)
17. Rosenberg, A.: Applying real-world BPM in an SAP environment, 1st edn., p. 100. Galileo Press, Bonn, Bosten (2011)
18. Automation systems and integration - Key performance indicators (KPIs) for manufacturing operations management. BS ISO 22400
19. Harrington, H.J.: Business process improvement. The Breakthrough Strategy for Total Quality, Productivity, and Competitiveness. McGraw-Hill, Inc, New York (1991)
20. del-Río-Ortega, A.: On the Definition and Analysis of the Process Performance Indicators. Sevilla (2012)
21. Nenadál, J.: Process performance measurement in manufacturing organizations. In: International Journal of Productivity and Performance Management. 57th edn. p. 24. (2008)
22. Wetzstein, B., Ma, Z., Leymann, F.: Towards measuring key performance indicators of semantic business processes. In: Abramowicz W., Fensel D. (eds.) BIS 2008, LNBIP 7, pp. 227–238. Springer, Berlin (2008)
23. Gladen, W.: Performance Measurement. Controlling mit Kennzahlen, 6th edn. Springer, Wiesbaden (2014)
24. Lanza, G.: D3.2.: Operational KPI evaluation framework. Customer-driven design of product-services and production networks to adapt to regional market requirements. (2015)
25. Kitchenham, B.: Guidelines for performing Systematic Literature Reviews in Software Engineering. EBSE Technical Report, 2.3 edn. Keele (2007)

Highly Iterative Planning of Mixed-Model Assembly Lines

J. Maetschke[1(✉)], B. Fränken[2], F. Sauermann[1], A. Gützlaff[1], and G. Schuh[1]

[1] Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
J.Maetschke@wzl.rwth-aachen.de
[2] e.GO Mobile AG, Campus-Boulevard 30, 52074 Aachen, Germany

Abstract. Manufacturing companies are confronted with increasing market and innovation dynamics resulting from the demand for differentiated products. This leads to increasing importance of flexibility and versatility as objectives of production, especially in final assembly, since the highest variant-induced complexity occurs there. For this reason and to avoid investment costs for product-dedicated assembly lines, companies strive to assemble several products on one assembly line in a so-called model mix. Highly iterative product development is increasing in order to take the mentioned volatile market environment into account. However, this agile development methodology contradicts the sequential processes of production development and the solution space already defined by the existing production system. In order to address this conflict of objectives, this paper proposes a procedure for a holistic and cross-domain development cooperation for the integration of a new product on an existing assembly line.

Keywords: Highly iterative product development · Process development · Production development · Mixed model assembly line

1 Introduction

Manufacturing companies, especially those in the automotive industry, are confronted with great market and innovation dynamics [1, 2]. This results from the heterogeneity of customer requirements and the individualization urge of customer groups, which is expressed in the segmentation of the customer structure and forces companies to develop differentiated products in small quantities and with a high number of different parts, components and modules [3]. Considering this, higher model-specific fluctuations in demand are to be expected in the future, which can only be predicted with a high degree of uncertainty [4].

This leads to an increase in the importance of flexibility and versatility as objectives of production, especially of assembly-lines, since the highest

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 680–688, 2021.
https://doi.org/10.1007/978-3-662-62138-7_68

variant-induced complexity occurs here [4, 5]. Addressing these objectives, the production of different derivatives of a product is conducted on a single assembly line, a so called mixed-model assembly line. In this way, it is possible to react to shifts in derivative-specific demand and market fluctuation in a demand-oriented manner without having to constantly maintain capital-intensive capacities for specific products. Additionally, derivative-specific bottlenecks of capacity can be avoided with mixed-model assembly lines. [4, 6].

When trying to establish a new product derivative on the market as an innovation, contradictory objectives must be aligned: the shortest possible time-to-market with lowest possible product costs [7]. This leads to the need of reducing the product and production process development times as well. Therefore, sequential, plan-driven product development models get less feasible since changes in advanced planning phases lead to high change efforts and costs [5, 7]. An development approach for physical products that enables taking customer requirements into account even in advanced project phases is the highly iterative product development (HIPD) according to [8]. A final definition of the product in terms of HIPD is not made at the beginning of the development project but is developed successively within iterative short-cycle phases.

The combination of these two perspectives leads to the research question of this paper: How should a methodical support for the highly iterative collaboration of product and production development in the automotive industry be designed when planning the integration of new derivatives on an existing assembly line?

Section 2 outlines the challenges of a highly iterative approach to planning derivative flexible assembly-lines. Subsequently, Chap. 3 gives an overview of existing approaches to the topic, while Chap. 4 presents the approach developed in this paper. Chap. 5 provides the summary and identifies future research needs.

2 Challenges of Integrating a New Derivative on an Existing Assembly Line in a HIPD Approach

Contemplating the fundamental concept of HIPD (respectively parallelized development in general), decisions are generally made under uncertainty and with incomplete information of possibilities and restrictions [9–11]. Consequently, this also applies to the evaluation of the technical feasibility of integrating a new derivative on an existing assembly line of an automotive plant. In this case the assembly line is already physically materialized at the beginning of the product creation process, while the specifications for the new product are mostly insufficient for a detailed technical evaluation at that time [12]. Classical methods of assembly planning, however, require detailed specifications for the assembly line in order to plan its adaptation [4, 13, 14]. The planning process is further complicated by different, often contradictory, perspectives and objectives on the product and the production process in each department of an enterprise [15].

Due to the existing assembly line, there are various restrictions as a result of the layout design, equipment and the precedence sequence of the existing derivatives,

which must be considered when integrating new derivatives. Thus, a non-conformance to production constraints will inevitably lead to an inability of assembling the new derivative on an existing line.

Required changes to the product or the assembly line are often associated with high adaptation costs or the delay of the start of production (SOP) due to adaptation times of the production system [7, 14]. For example, necessary structural adjustments to equipment or production processes may require a specification of certain product characteristics well in advance of the SOP due to the long time required for implementation. For this, even following a highly iterative approach, it is necessary to restrict the degree of freedom in production development at certain milestones and thus aligning product development and production planning. This results in a discontinuous restriction of the level of freedom for product development as shown in Fig. 1 exemplarily.

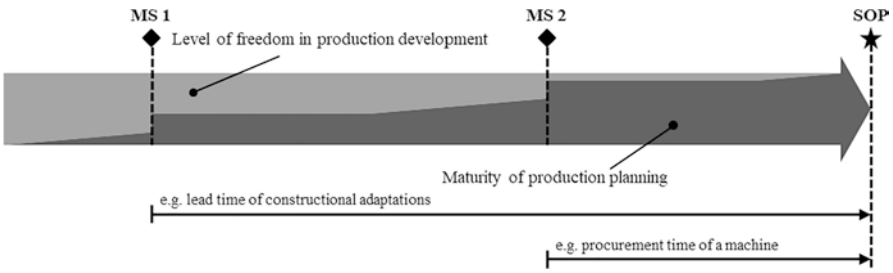


Fig. 1. Discontinuous restriction of the degree of freedom due to delivery times etc.

In order to avoid these cost-intensive adjustments and to ensure economic production, it is therefore essential to integrate knowledge about production constrains into product development from the earliest possible stage [16].

From the point of view of a superordinate project management, late product adaptations to meet customer requirements must be set in relation to the resulting change costs and realization time in an impartial and cross-domain manner.

3 Previous Approaches for Planning Mixed-Model Assembly Lines

In the following, a selection of relevant approaches is presented and discussed regarding the challenges mentioned in Chap. 2.

Planning of product flexible assembly configurations by Küber This approach is based on modularizing assembly processes in order to master the complexity of their dependencies. To facilitate planning, a distinction is first made between strategic and flexible assembly modules. The former cannot be moved in the assembly layout without incurring time and costs, but the latter can initially be neglected in planning. Afterwards, the degrees of freedom for the allocation of product modules to assembly

stations are determined. Finally, the flexible modules are positioned after the strategic ones. In this way, not only the length of the assembly line and the workload of the employees can be optimized, but also derivative-specific deviations in cycle times can be homogenized [4].

As a result, this approach requires complete information on the product from the beginning of the methodology, does not include a highly iterative understanding of planning and does not take temporal assurance of SOP into account. On the other hand, a high-quality solution can be achieved through this detailed analysis.

Complexity-focused planning of product flexible assembly lines by Keckl et al. This methodology involves continuous analysis, evaluation, and adjustment in the form of a five-step control loop. In the beginning, an existing assembly line is characterized and modeled. Subsequently, the system is examined for complexity, meaning variance and change, and indicators are determined that reflect the influence of complexity on efficiency. From this, strategic approaches to achieve optimal efficiency and flexibility are derived. Finally, change processes are initiated that lead to the overall goal of increasing flexibility and efficiency of the assembly line [17].

However, this approach disregards time as a resource as an important basis for decision-making in highly iterative production planning. What is particularly interesting about this approach is the idea of strategic approaches to solving problems from a higher-level perspective.

Integrated product and assembly configuration by Landherr In this approach, the planning and optimization process of factories is considered a permanent task of adapting and coordinating measures. The goal is to ensure the efficiency and versatility of the multi-variant series assembly. This is achieved by the consideration of product and assembly configurations in an integrated product and assembly system (IPS). From that point on, possible needs for change due to adaptions to the product or the assembly are constantly monitored. An evaluation scheme shows the greatest correspondence with previously defined situation-specific objectives [18].

Landherr does not give a dedicated planning method for product-flexible assembly lines in flow production and ignores some of the agile values, but already meets many of the requirements for approaches to agile planning of product-flexible assembly lines.

Agile process planning in the product development process by Schneider This concept of process planning is based on the values of agile software development and the commonly used Scrum methodology [19]. Since this basis, the actual process planning takes place in sprints. At regular intervals, the developed production system increments are evaluated and the prioritized work for the following sprint is determined [20].

This approach does not focus on product flexible assembly lines and hardly takes restrictions of existing assembly lines into account and the time factor is only integrated in form of typical milestones and not as a basis for decision making.

The identified challenges for a method to support agile collaboration between product and production development in the planning of the integration of new derivatives on an existing assembly line are currently not satisfactorily met in the literature.

However, an excellent basis for answering the research question is provided, as particular aspects of the problem are well investigated.

4 Approaches for Highly Iterative Planning of Product Flexible Assembly Lines

Six iterative tasks are proposed which enable the identification of change requirements necessary, if a new derivative is to be integrated on an existing assembly line. The identification of these are based on the comparison of existing derivatives and the existing assembly line on the one hand and the incomplete information about the new derivative to be integrated on the other hand.

Figure 2 below gives an overview of the link between these tasks and shows the iterative character of the methodology.

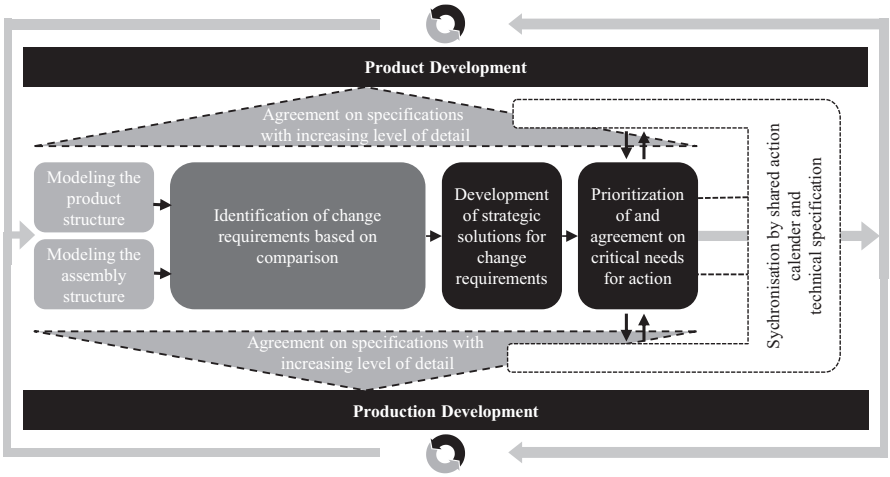


Fig. 2. Approach for iterative planning of mixed-model assembly lines

In the following, the tasks shown in the Fig. 2 above are presented successively. For these tasks, self-organizing, partially autonomous teams are assumed in both product and production development. It is also assumed that the employees have detailed knowledge of the product already assembled on the existing production line. The methodology was developed with the automotive industry in mind, but can also be transferred to other industries that assemble in assembly lines.

Modeling the Product Structure First, the derivatives currently assembled on the assembly line must be represented as product modules to reduce complexity. In principle, only those components may be combined to modules which can be mounted simultaneously on the assembly line in regard to the precedence sequence. This means that only components that require the same assembly progress may be combined. Then the new derivative to be integrated must also be described following the same principles. The information about the new product becomes more reliable and more

detailed as it matures in the product development process, while is vague at first. In contrast, the information on the derivatives already assembled on the line is available in sufficient detail at the start of the project and must therefore be used as a reference point for planning the new derivative. The modules of the new derivative need to be typified by means of communalities (based on a comparison to existing derivatives) in order to estimate the influence of integration on the existing assembly line.

Modeling the Assembly Structure The existing assembly line must be described after modeling the product structure. First, the capabilities for each assembly station are collected. Capabilities enable the execution of certain assembly processes at a station by e.g. assembly equipment or employee qualification. The allocation of assembly modules to stations reduces fixed, limited resources at these stations. Resources are understood to be quantitative units to be consumed, e.g. area or assembly time. An integral part of the method is the detection of production-related restrictions. Restrictions are limitations of the product that result from the realization of the assembly line and that would eliminate the possibility of assembling a derivative even if the line was used as a dedicated line for that derivative, e.g. limited ground bearing capacity of the shop floor. Capabilities can be added to the assembly line, but restrictions can only be overcome by changing existing elements of the production system. The consolidated and aggregated restrictions result in the solution space for the development of new product derivatives. For this reason, the integration of knowledge about the fundamental restrictions in product development is elementary in order to ensure assembly with as few adaptations as possible. The focus must be on the assembly line itself, although upstream and downstream support processes must also be taken into account.

Identification of Change Requirements Based on Comparison Based on the precedence sequence of the existing derivatives and the communality typification of the assembly modules, an initial and simplified allocation of these modules of the new derivative to the assembly stations of the existing assembly line must be conducted. The product and assembly structure are linked to each other with the help of visualizations in varying degrees of detail depending on the project's progress. Initially, the aim is not to show technical details, but rather to offer representations that provide all involved departments with information in early planning phases, combine different perspectives on the product und establish a common understanding for discussion.

By comparing the capabilities, resources and restrictions of the assembly stations with the product requirements, change requirements are identified that prevent the assembly of the new derivative on the existing assembly line.

Development of Strategic Solutions for Requirements for Change In order to overcome the problems identified, strategic approaches to solutions must be developed in various fields of design. In doing so, adaptations of the assembly line, the product as well as the organization must be considered equally. Solutions can only be generated in close coordination between all involved planning departments. Examples of strategic solution approaches are the standardization of process steps, increasing physical commonality, shifting the variance in pre-assemblies, reducing the depth of production, desperation of the line for specialized assembly tasks or just-in-sequence

provision of material [17, 18, 21]. It is decisive that the progress of the project must also be anticipated for every strategic solution in order to estimate the timing of information, decision and action needs for further detailing. Thus, based on the requirements for changes and the estimated implementation time, points in time are identified at which degrees of freedom in product development must be narrowed down in order to ensure compliance with the SOP while taking downstream processes into account. Different departments usually prefer different solutions, as they try to minimize the respective change effort in their own domain. The prioritization of the detailed elaboration will be decided objectively and across domains in the next task.

Prioritization of an Agreement on Critical Needs for Action Considering the strategic positioning of the company, the developed strategic solution approaches must be evaluated and prioritized before they are worked out in detail. The urgency of a decision is derived from the combined assessments of the time required to fulfill the task and the risk of inability to assemble the new derivative on the assembly line. In this way, any need for change needs to be set in relation to the common goal of securing the SOP and the resulting customer benefit regarding the objectives time, costs and quality.

Agreement on Specifications with Increasing Level of Detail The entire process is accompanied by the central idea of successive agreements on specifications with increasing level of detail. This is therefore to be understood less as a task and more as an underlying maxim in each of the tasks described above, which is why this is also shown for all tasks in Fig. 2. Basically, it is necessary to check when decisions can be made at the latest in order to not delay the course of the project. This results in a due date for each information requirement, which is fixed in a shared action calendar. At the latest of these due dates for a specific change, a specification must be that detailed that the next planning phase can be initiated. Qualitative concept agreements, which are sufficient in early stages of the project, need to be detailed to quantitative requirements over time, which are then binding for all teams involved, provided that the intended SOP is to be kept.

Synchronization by Shared Action Calendar and Technical Specification In order to have an overview of the adherence to the schedule (and thus the SOP) and the interdependencies of the requirements of the teams of different planning departments, it is advisable to establish a dedicated role for these tasks, an integration manager. Integration in this context ambivalently describes the integration of a new derivative on an existing line and the integrated product creation in terms of HIDP. If decisions in the project exceed the Integration Manager's decision-making responsibility, these need to be made in cross-departmental escalation committee on short notice. The integration manager is responsible for managing the shared action calendar and the technical specification.

In order to address changing conditions and to consider the constant detailing of the new derivative, the above-mentioned steps must be carried out iteratively.

The presented methodology provides a framework for the cooperation between product and production development and defines a problem-specific interface between these departments by providing a common and simplified view on the project. The

methodology can be combined with - and respectively integrated into - existing agile project management and development procedures as presented by [20] or [14] due to its iterative character.

5 Conclusion and Further Research

Due to the increasingly difficulty in forecasting demand for a growing number of product derivatives, companies in the automotive industry are striving towards the assembly of different derivatives on one assembly line. In order to meet simultaneously decreasing product life cycles, products are being developed in highly iterative, short-cycle development processes. Within the scope of this paper an iterative method was developed which supports the cooperation between product and production development in the integration of a new product derivative on an existing assembly line under incomplete information about the new product derivative. The aim was to create a common view of the integration project, to identify possible requirements for changes to both the product and the assembly line, to make it possible to evaluate implications in terms of customer benefit and to achieve transparency regarding the use of time as a resource in order to secure the date of SOP.

Based on the approach, further research is needed, especially with regard to the gradual definition of specifications between product and production development. Therefore, the development of an agile understanding of requirements is proposed, which supports both the continuous implementation of changes and cost-optimized production planning. For this purpose, it seems necessary to identify and systematically describe information requirements without assuming sequential planning procedures.

The concept presented provides a framework for various to be detailed, methodical modules that enable the integration of innovative products on existing assembly lines which can lead to cost-savings and the reduction of time-to-market.

Acknowledgement. The authors thank the German Research Foundation DFG for the kind support within the Cluster of Excellence “Internet of Production” - Project-ID: 390621612.

References

1. Abele, E., Reinhart, G.: Zukunft der Produktion. Herausforderungen, Chancen. Hanser, München (2011)
2. Rudow, B., Neubauer, W.: Trends in der Automobilindustrie. Entwicklungstendenzen, Betriebsratsarbeit, Steuer- und Fördertechnik, Gießereitechnik, Informationstechnologie und -systeme. Oldenbourg, München (2012)
3. Wiendahl, H.-P., Klepsch, B.: Komplementäre Produkt- und Fabrikmodularisierung als Ansatz zur Komplexitätsbewältigung. ZWF **101**, 367–373 (2006)
4. Küber, C.: Methode zur Planung modularer, produktflexibler Montagekonfigurationen in der variantenreichen Serienmontage. Dissertation, Universität Stuttgart (2017)

5. Feldhusen, J., Grote, K.-H. (eds.): Pahl/Beitz Konstruktionslehre. Methoden und Anwendung erfolgreicher Produktentwicklung, 8th edn. Springer Vieweg, Berlin (2013)
6. Lotter, B., Wiendahl, H.-P.: Montage in der industriellen Produktion. Ein Handbuch für die Praxis, 2nd edn. VDI-Buch. Springer, Berlin (2012)
7. Schuh, G., Prote, J.-P., Dany, S., Molitor, M., Giner, B.: Process model for generative assembly planning in the highly iterative product development. Procedia CIRP **72**, 363–368 (2018)
8. Diels, F., Rudolf, S., Schuh, G.: Highly iterative product development process for engineering projects. AMM **794**, 532–539 (2015)
9. Bochtler, W., Laufenberg, L.: Simultaneous Engineering. Erfahrungen aus der Industrie für die Industrie. Springer, Berlin (1995)
10. Berndes, S., Stanke, A.: A Concept for revitalisation of product development. In: Bullinger, H.-J., Warschat, J. (eds.) Concurrent Simultaneous Engineering Systems, p. 20. Springer, London (1996)
11. Schlick, C., Bruder, R., Luczak, H.: Arbeitswissenschaft, 4th edn. Springer Vieweg, Berlin (2018)
12. Schuh, G., Gartzen, T., Basse, F., Schrey, E.: Enabling radical innovation through highly iterative product expedition in ramp up and demonstration factories. Procedia CIRP **41**, 620–625 (2016)
13. Molitor, M.: Generative Montageablaufplanung in der hochiterativen Produktentwicklung, 1st edn. Produktionssystematik, vol. 17/2019. Apprimus Verlag, Aachen (2019)
14. Wlecke, S., Prote, J.-P., Molitor, M., Müller, C., Schuh, G.: Concept for the industrialization of physical products in the highly iterative product development. In: Wulfsberg, J.P., Hintze, W., Behrens, B.-A. (eds.) Production at the leading edge of technology, pp. 583–592. Springer Berlin Heidelberg, Berlin, Heidelberg (2019)
15. Schuh, G., Stich, V. (eds.): Produktionsplanung und -steuerung. Funktionen. Springer, Berlin (2012)
16. Eversheim, W., Schuh, G.: Integrierte Produkt- und Prozessgestaltung. Springer, Berlin (2005)
17. Keckl, S., Abou-Haydar, A., Westkämper, E.: Complexity-focused planning and operating of mixed-model assembly lines in automotive manufacturing. Procedia CIRP **57**, 333–338 (2016)
18. Landherr, M.H.: Integrierte Produkt- und Montagekonfiguration für die variantenreiche Serienfertigung. Zugl.: Stuttgart, Univ., Diss. Stuttgarter Beiträge zur Produktionsforschung, vol. 39. Fraunhofer, Stuttgart (2014)
19. Schwaber, K., Beedle, M.: Agile software development with Scrum. Series in agile software development. Prentice Hall, Upper Saddle River (2002)
20. Schneider, S.: Agile Prozessplanung im Produktentstehungsprozess am Beispiel der Motorenproduktion. Zugl.: Technische Universität Dortmund, Diss. Schriftenreihe Industrial Engineering, vol. 16. Shaker, Aachen (2015)
21. Löffler, C.: Systematik der strategischen Strukturplanung für eine wandlungsfähige und vernetzte Produktion der variantenreichen Serienfertigung. Zugl.: Stuttgart, Univ., Diss. IPA-IAO Forschung und Praxis, vol. 519. Jost-Jetter, Heimsheim (2011)

Token-Based Blockchain Solutions for Supply Chain Strategies

F. Dietrich[1,2(✉)], A. Turgut[3], D. Palm[1,4], and L. Louw[2]

[1] ESB Business School, Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany
fabian.dietrich@reutlingen-university.de
[2] Department of Industrial Engineering, Stellenbosch University, Stellenbosch, South Africa
[3] Steinbeis Innovation gGmbH, Alteburgstr. 150, 72762 Reutlingen, Germany
[4] Fraunhofer Institute for Manufacturing Engineering and Automation, Alteburgstr. 150, 72762 Reutlingen, Germany

Abstract. Distributed ledger technologies such as the blockchain technology offer an innovative solution to increase visibility and security to reduce supply chain risks. This paper proposes a solution to increase the transparency and auditability of manufactured products in collaborative networks by adopting smart contract-based virtual identities. Compared with existing approaches, this extended smart contract-based solution offers manufacturing networks the possibility of involving privacy, content updating, and portability approaches to smart contracts. As a result, the solution is suitable for the dynamic administration of complex supply chains.

Keywords: Blockchain · Smart contract · Supply chain risk management · Sustainability · Supply chain control

1 Introduction

In increasingly complex supply chains (SCs), companies are faced with numerous challenges [1] and need to respond to these in order to fully identify and manage risks [2]. Innovative technologies can be used to overcome challenges in this context [3], such as the blockchain technology (BCT). The BCT is a distributed and immutable database that allows transactions to be processed directly and transparently between parties [4]. At the same time, it is possible to write and execute algorithms and rules, e.g. to execute automatic transactions [5]. These so-called 'smart contracts' are scripts which are created and used by members of a blockchain system [6].

An elaboration developed by Christopher and Peck [7] proposes separating SC risks into three main areas: organizational risk sources that are internal to the enterprise, network risk sources that are external to the enterprise but within the supply network, and environmental risk sources that are external to the network. This

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 689–698, 2021.
https://doi.org/10.1007/978-3-662-62138-7_69

classification clarifies the relevant dimensions of potential disturbances in a SC environment and thus provides the basis for a comprehensive risk analysis. The three main areas can be divided into five different categories of risks in the SCs (see Fig. 1).

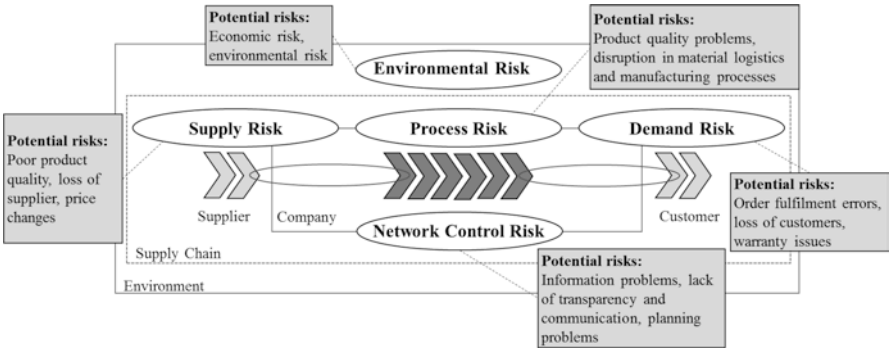


Fig. 1. Areas of supply chain risks (own illustration based on [3, 7, 8])

Based on the potential of the BCT, research approaches use the technology as an innovative solution for challenges in the area of Supply Chain Risk Management (SCRM). Due to their structural design, traditional SCRM tends to take actions reactively after damage has occurred. According to Min [9], the BCT provides the possibility of designing actions proactively and preventively as well as managing Network Control Risks by improved information flow in the SC. Additionally, Layaq et al. [10] reveal that the use of the blockchain and smart contracts can lead to improvements in material and information flows in the area of SC by increasing visibility (access to or sharing information through participants of the SC network [11]), transparency (disclosure of information to all stakeholders, including customers [12]) and transaction automation. The blockchain technology is already used by companies, e.g. to track products or raw materials along SCs in different industries [13, 14]. The success of selected SCRM mechanisms can depend on their alignment to existing or future SC structures. Types of SCs distinguish between functional and innovative products [15]. Functional products are basic goods with low variation possibilities and a predictable demand. Therefore, a SC structure should focus primarily on efficiency in order to operate with the lowest possible costs. In contrast, the main characteristics of innovative products are that they are continuously enhanced as well as highly variable, leading to an unpredictable demand. In that case, a market-responsive approach in order to react quickly to changing customer demands is an appropriate choice. Lee [16] added the supply aspect and classified SC strategies. Based on Fisher [15] and Lee [16], Perez [17] developed a roadmap which differentiates between efficient and responsive SCs and substrategies in each category:

- Oriented to Efficiency: Continuous-flow SCs, Efficient SCs, Fast SCs
- Oriented to Responsiveness: Custom-configured SCs, Agile SCs, Flexible SCs

This paper focuses on solving network control risks by applying blockchain smart contracts and improving transparency in the SC. First, a new extended smart contract-based approach is proposed, which represents a further development of conventional token-based approaches. Subsequently, the two solutions are compared and evaluated with respect to different SC strategies.

2 Related Works and Rationale of the Paper

In the context of this paper, the term 'tokenizing of assets' relates to the representation of physical assets on the blockchain [18]. On-blockchain tokens can represent all kind of assets such as currencies, stocks, properties and coupons [6]. Such tokens, which embody a 'claim' on the physical asset, can also be designated as asset-backed tokens [19].

First, tokenizing assets in SC use cases was used simply to prove the ownership and provenance of physical objects [20]. This approach was extended to solve transparency problems in SCs [21]. This development was supported by the definition and adaptation of token standards, especially on the Ethereum platform [22, 23]. Fundamentally, one can distinguish between two types of tokens, fungible token standards and non-fungible token standards [22]. *Fungible tokens*: different units of a fungible token are interchangeable and have no unique properties. *Non-fungible tokens*: each unit of a non-fungible token is unique from another, allowing the tracking of their ownership.

Token standards are the minimum specifications of required functions to allow an implementation. However, it is also possible to add further functions that are not part of the standard in order to specify tokens for their respective application [22]. Westerkamp et al. [23] used these possibilities to adapt the non-fungible token standard ERC-721 [22, 24] specifically to the needs of manufacturing SCs. Thereby, the ERC-721 standard is expanded by so-called 'token recipes' to map complex manufacturing processes. Such an approach enables the traceability of manufactured goods, including their components [23]. A closer look at the structure of token standards shows that ownership changes of tokens and the reflection of a user's balance are organized in transfer functions determined in a token contract [22]. Therefore, a transaction to change a token's ownership does not send the token itself to a new address; it is an interaction with the token contract assigning a new owner address to the token. Based on this fundamental functionality of tokens, a new framework was designed, which expands the approach of non-fungible tokens with the possibility of assigning clear authorities when deploying the smart contract [25]. Based on this framework an extended smart contract-based solution is being developed, specifically designed for the needs of complex SCs. This solution potentially allows mapping SCs holistically in a smart contract, including dynamic changes in the composition of products or the structure of the SCs. The next chapter presents the holistic smart contract structure of this new extended approach.

3 Extended Token-Based Solution

This chapter proposes an extended token-based solution aiming to create a smart contract with a static address on the network, which content remains dynamic after its deployment to enable the dynamic mapping of changes in the SC. At the end of this chapter, the holistic smart contract structure of the approach is presented.

3.1 Determination and Assignment of Authorities

As a first step, the identification and mapping of all stakeholders in the SC takes place. For this purpose the SC participants will be divided into SC partners (suppliers, producers and service providers), customers and certifiers [23]. The certifiers represent a special peculiarity providing non-physical assets such as certificates, licenses, and patents necessary to be authorized to produce or assemble components or manufacture the final product. The role of the certifier can be taken over by the manufacturer themselves, or by a third party. According to Dolgui et al. [26] SCs are subject to dynamic influences. Therefore, when modeling SCs in smart contracts, algorithms that allow continuous dynamic changes of the smart contract's content must be involved before deploying the contract on the network. Additionally, authorities in the SC entitled to create or modify virtual identities in the smart contract must be assigned to their respective functions.

3.2 Creation of Virtual Identities

As a second step, all processes in the SC must be logically linked with each other and reflected in the smart contract. For this to be possible, all requirements of the physical SC must be transferred to the virtual SC. In the physical world, the relationships between processes are obvious. For example, it is not possible to assemble a product without having the required parts on hand. These relationships must be mapped in the smart contract using two different types of assets. In this context, *Assets without dependencies* can be created without depending on previous actions, while *Assets with dependencies* can only be created when previous actions have been successfully completed.

In addition, in the physical world, only certain entities are able to create certain parts. This means, that only the entity 'owning' the process of creating a certain part in the physical world is able to have access to the function of the smart contract to create the virtual identity of that part. To guarantee this, the smart contract functions must be linked with the respective authority before deploying them on the blockchain network.

Once all relations of the product and its components have been defined, they must be transferred to the logical structure of the BCT. Grossman [27] describes the BCT as a "public database of timestamps". In traditional models, each system kept its own notion of time. In the BCT, however, all nodes of the network agree on a common time by implementing a consensus algorithm. As a result, the immutability of the network guarantees that a certain event on the network happened at a certain time. The

proposed smart contract-based solution aims to record all interactions with the smart contract on the blockchain in order to achieve transparency and the auditability of all assets.

A unique identification number (ID) refers to the virtual identity of an asset. The uniqueness of these IDs can be achieved by hashing the information they refer to. Therefore, Hash IDs are a logical result of their input data and provide information about the origin, composition, and time of creation. Hash IDs can also be referred to as the smart contract version of primary keys. Within the smart contract, these Hash IDs refer to the virtual identity of their physical counterparts. In order to map SC processes, the virtual identities must have the same ownership and conversion characteristics as their physical counterparts. They must be able to change their ownership when the physical product is transferred in the SC or sold. Furthermore, the virtual identities must be able to be summarized when combining several components to a new product. Since virtual identities can represent assets with and without dependencies, their creation on the blockchain must be linked to logical dependencies. To map such a logic, it requires the modelling, planning, and definition of further various smart contract functions. Therefore, it is necessary to carefully design all required conditions to create the virtual identity of an asset and the associated Hash ID. Westerkamp et al. [23] refer to this as the 'recipes' of virtual identities. However, the creation of virtual identities can also be affected by dynamic changes and these functions must be embedded into dynamic changing algorithms.

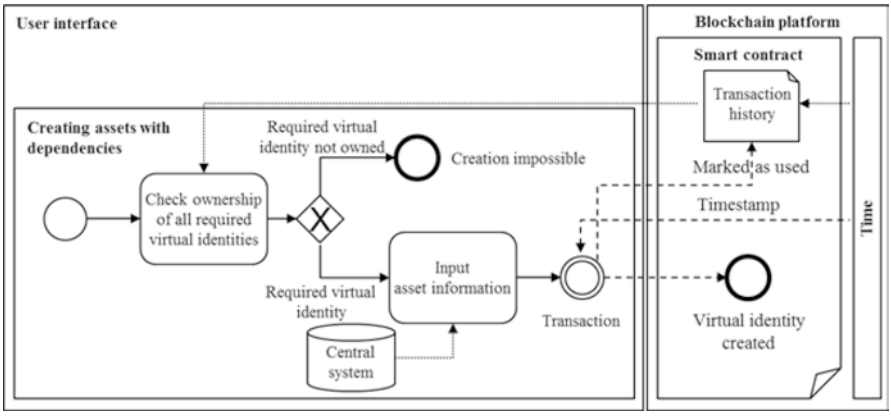


Fig. 2. Model to create virtual identities for assets with dependencies

As shown in Fig. 2, when the possession of the smart contract is clearly assigned and the ownership of all required virtual identities is clarified, information for creating a new virtual identity can be entered into the user interface of the smart contract. The information stored behind each Hash ID at the time of creation can come from a central IT system (e.g. ERP-System). This creates an interface between existing business software and the blockchain (see Fig. 2). The address of the creator is then authorized to send this virtual identity to a new owner.

3.3 Holistic Smart Contract Structure

The assigned authorities and the functions to create virtual identities are embedded into algorithms allowing the mapping of dynamic changes in the SC while the smart contract's address remains static. For these complex algorithms, existing approaches to update and continuously generate smart contracts can be used [26, 28].

As illustrated in Fig. 3, it is determined to what extent the authorities can have an impact on changes in the virtual identity logic. Such a structure allows a wide range of possibilities in terms of the allocation of authorities, which is determined in the dynamic updating algorithm related to the list of authorities.

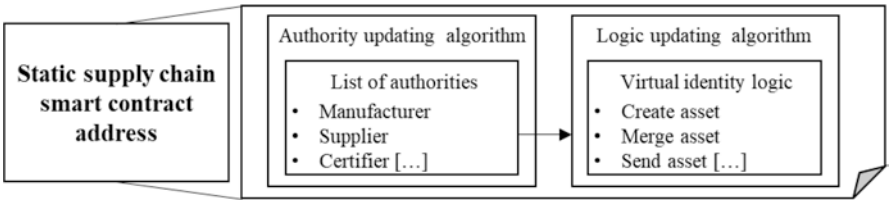


Fig. 3. Holistic smart contract structure

Decentralized and open SCs with an even distribution of power can be virtually represented, as well as regulated SC structures with only one central authority. It is important to state that such smart contract constructs are also affected by the immutability of the BCT. Consequently, any eventuality that could affect the SC or product structure must be considered in the program code before its deployment. In this context it is not necessary to describe every eventuality in detail, but to categorize certain types of eventualities in advance. If, for example, the possibility of adding new suppliers is not taken into account in the code, they cannot be added. Instead, a new smart contract with a new address must be deployed. Furthermore, the proposed structure allows involving privacy approaches to encrypt the content of the smart contract [29]. In this way, the interactions with the contract would remain visible to the public to provide SC transparency. The content of these transactions would, however, be encrypted and only be visible for a defined group of authorities.

4 Comparison of Solutions and Conclusion

This chapter compares the characteristics of blockchain-based SC control solutions. On the one hand are the characteristics of conventional ERC-721-token-based solutions for manufacturing SCs (see Chap. 2) and on the other hand the characteristics of the extended token-based solution proposed in this paper. The comparison is shown in Table 1. Subsequently, the two solutions are evaluated with respect to different SC strategies.

Table 1. Comparison of Blockchain-based supply chain control solutions

	Conventional token-based solution	Extended token-based solution
Authority	No central authority necessary, changes in the Supply Chain (SC) can be decided decentrally, but must be implemented downstream.	A defined authority concept manages and coordinates the changes in the smart contract. Different authority concepts can be applied (some authorities may receive additional power over the SCs).
Flexibility	Changes in the SC (e.g. parts, suppliers) require the distribution of new tokens downstream. Downstream SC partners can only react to changes. This leads to high change effort and uncoordinated actions in complex SCs.	The authorities enable product and structure changes by continuously updating the smart contract. This facilitates the change effort in complex SCs.
Effort	Fast technical implementation especially in linear and simple SCs, since the implementation does not require coordination of the SC partners and the 'token recipes'.	More elaborate planning necessary since all change contingencies must be considered in advance. The complexity of the technical implementation increases with the degree of individualization of the solution.
Accessibility	Standardized tokens (e.g. ERC-721) freely accessible to the public.	Non-standardized solution and not necessarily open to the public; private, semi-private or public smart contract on a public blockchain possible.
Storage	The level of detail of the information in the BC can be considered when deploying the token contract. Unused or outdated token lead to data waste of worthless tokens in the system.	The level of detail of the information can be considered when deploying the smart contract. 'Token recipes' can be dynamically adjusted to avoid data waste.
Privacy	Privacy is not given in public networks.	Parts of the smart contract can be encrypted for the public (only accessible for selected partners).

For the properties of the SC strategies the roadmap by Perez [17] is used. It distinguishes efficient and responsive SCs with three substrategies each.

Oriented to efficiency

Continuous Flow SCs focus on building synergies supported by collaborative relationships. They are characterized by long product life cycles with a low number of demand variations. The inventory consists of finished products having only one supplier for each key component. Efficiency oriented SCs with a continuous flow do not consider customization. These properties indicate a very low need for dynamic changes when adopting SC control solutions. Therefore, the conventional token-based solution is sufficient.

Efficient SCs focus on low costs, which is making them suitable for functional and commoditized products. The demand variation is medium to high with long product life cycles. Demand buffering is achieved with an inventory of finished products. Because such SCs have no customization requirements and the network is kept stable and simple (best-cost supplier on each occasion), dynamic changes are somewhat insignificant. Therefore, for efficient SCs the conventional token-based solution is recommended.

Fast SCs focus on competitive cost and the continuous renewal of the portfolio. Fashion-oriented SCs with functional and short life cycle products are typical representatives. They rely on supplier pools to cover their sourcing needs. With some exceptions, fast SCs do not consider customization. Even though they have few requirements in terms of dynamic structural and administrative adjustments, a fast and easy implementation of the conventional token-based solution is of greater importance. If the customization of the product and the supplier pool is subject to dynamic changes, the extended solution can be considered in certain cases.

Oriented to Responsiveness

Custom-configured SCs focus on product configurability allowing customers to choose product features from a set of limited options. Therefore, custom-configured SCs are suitable for products with a modular design and several configurations. Depending on the stability and continuity of the modular product composition, such SCs require an enhanced consideration of dynamic impacts on the structure of the product composition. In certain cases with simple, stable, and standardized configuration options, conventional standardized tokens could potentially be considered as a cost-effective solution.

Agile SCs focus on responsiveness to unpredictable changes in demand. The products are highly customizable and therefore usually designed only for small batches. The customization is relevant for manufacturing processes as well as for downstream processes. These characteristics of agile SCs place high demands on SC control solutions regarding dynamic structural and compositional product changes, making the extended smart contract-based solution the most suitable. When adopting this solution, it is particularly important to predict eventualities in advance.

Flexible SCs focus on unique solutions. The customization takes place during the design process. This requires adaptable processes making their reconfigurability a key factor for ensuring flexibility. Regarding SC control solutions, flexible SCs have high requirements in terms of the consideration of dynamic changes. The main challenge is to identify and cover all eventualities in advance in order to maintain flexibility. Therefore, SCs where the possible eventualities indicate commonalities, the extended smart contract-based solution can cover these with a substantial planning effort.

The summary of the recommended smart contract-based solution for the SC strategies is shown in Table 2.

Table 2. Deployment of blockchain-based smart contract solutions for supply chain strategies

	Supply chain strategies	Extended token-based solution	Conventional token-based solution
Responsive	Flexible	○	-
	Agile	+	-
	Custom-configured	+	○
Efficient	Fast	○	+
	Efficient	-	+
	Continuous-flow	-	+

+ Recommended ○ In certain cases - Not recommended

5 Discussion

An analysis of the two solutions in terms of suitability for SC types shows that the conventional non-fungible token standard is mainly suitable for simpler SCs following an efficient-oriented approach with a low number of changes over time. This facilitates the implementation and management of downstream changes in the SC. In responsive-oriented SCs, the flexibility associated with configuration and customization of the product results in the need for dynamic changes of the SC structure and product composition. Therefore, responsive SCs are preferably suitable in combination with the proposed extended smart contract-based solution. Especially in complex SCs, this solution allows for continuous administration of the SC and the adoption of compositional changes. To what extent the changes of such complex SCs can be foreseen when designing the smart contract, should be further investigated. Further research regarding the practical feasibility of this new approach is currently being conducted. Moreover, first implementation experiments in a research environment show that it is necessary to investigate the scalability and practical applicability in real industrial scenarios.

References

1. Gartner: Supply Chain Risk Management. www.gartner.com/en/supply-chain/insights/supply-chain-risk-management. Accessed 08 May 2020
2. Palm, D., Kales, P.: Efficiency-oriented risk prioritisation method for supply chains in series production. Int. Conf. Comput. Manuf. **2016**, 417–422
3. Kersten, W., Schröder, M., Indorf, M.: Potenziale der Digitalisierung für das Supply Chain Risikomanagement: Eine empirische Analyse. In: Seiter, M., Grünert, L., Berlin, S. (eds.) Betriebswirtschaftliche Aspekte von Industrie 4.0, pp. 47–74. Springer, Wiesbaden (2017)
4. Bosch, R., Penthin, S.: Blockchain als Treiber im modernen Supply Chain Management 4.0. BearingPoint (2018)
5. Iansiti, M., Lakhani, K.R.: The truth about blockchain. Harv. Bus. Rev. **95**(1), 118–127 (2017)
6. Buterin, V.: Ethereum White Paper: A Next Generation Smart Contract & Decentralized Application Platform (2013). https://blockchainlab.com/pdf/Ethereum_white_paper-a_next_generation_smart_contract_and_decentralized_application_platform-vitalik-buterin.pdf. Accessed 20. Oct. 19
7. Christopher, M., Peck, H.: Building the resilient supply chain. Int. J. Logistics Manage. **15**(2), 1–14 (2004)
8. Kales, P.: Development of an efficiency-oriented risk prioritisation method for supply chains in series production. Thesis, Reutlingen University (2015)
9. Min, H.: Blockchain technology for enhancing supply chain resilience. Bus. Horiz. **62**(1), 35–45 (2019)
10. Wasim Layaq, M., Goudz, A., Noche, B., Atif, M.: Blockchain technology as a risk mitigation tool in supply chain. Int. J. Trans. Eng. Technol. **5**(3), 50–59 (2019)
11. Barratt, M., Oke, A.: Antecedents of supply chain visibility in retail supply chains. J. Oper. Manage. **25**(6), 1217–1233 (2007)

12. Doorey, D.: The Transparent supply chain: from resistance to implementation at Nike and Levi-Strauss. J. Bus. Ethics **103**, 587–603 (2011)
13. Wazid, M., Das, A.K., Khan, M.K., Al-Ghaiheb, A.A.-D., Kumar, N., Vasilakos, A.V.: Secure Authentication Scheme for Medicine Anti-Counterfeiting System in IoT Environment. IEEE Internet Things J. **4**(5), 1634–1646 (2017)
14. Arena, A., Bianchini, A., Perazzo, P., Vallati, C., Dini, G.: BRUSCHETTA: An IoT Blockchain-based framework for certifying extra virgin olive oil supply chain. IEEE (SMARTCOMP) **7**, 173–179 (2019)
15. Fisher, M.: what is the right supply chain for your product? Harv. Bus. Rev. **75** (1997)
16. Lee, H.: Aligning supply chain strategies with product uncertainties. Calif. Manage. Rev. **44**(3), 105–119 (2002)
17. Perez, H.D.: Supply chain strategies: Which one hits the mark? CSCMP's Supply Chain (Quarterly) (2013)
18. Primavera de F., Benedikt, S., Constance, C., Carla, R. et al.: Regulatory Framework for Token Sales (2019)
19. Hahn, C., Wons, A.: Initial Coin Offering (ICO). Unternehmensfinanzierung auf Basis der Blockchain-Technologie. Springer, Wiesbaden (2018)
20. Van de Velde, J., Scott, A., Sartorius, K., Dalton, I., Shepherd, B., Allchin, C., Dougherty, M., Ryan, P., Rennick, E.: Blockchain in capital markets - the prize and the journey. Oliver Wyman & Euroclear (2016)
21. Chang, S.E., Chen, Y.-C., Lu, M.-F.: Supply chain re-engineering using blockchain technology. Techno. Forecast. Soc. Change **144**, 1–11 (2019)
22. Antonopoulos, A.M., Wood, G.A.: Mastering Ethereum - Building smart contracts and DApps. O'Reilly Media, Tokyo (2019)
23. Westerkamp, M., Victor, F., Küpper, A.: Tracing manufacturing processes using blockchain-based token compositions. Digital Comm. Net. **6**(2) (2019)
24. Entriken, W., Shirley, D., Evans, J., Sachs, N.: EIP 721: ERC-721 (2018). https://eips.ethereum.org/EIPS/eip-721. Accessed 31 March 2020
25. Dietrich, F., Palm, D., Louw, L.: Smart contract based framework to increase transparency of manufacturing networks. CIRP Design Conference 30 (2020)
26. Dolgui, A., Ivanov, D., Potryasaev, S., Sokolov, B., Ivanova, M., Werner, F.: Blockchain-oriented dynamic modelling of smart contract design and execution in the supply chain. Int. J. Prod. Res. **58**(1), 1–16 (2019)
27. Grossman, N.: The Blockchain as verified public timestamps (2015). https://www.nick-grossman.is/2015/the-blockchain-as-time/. Accessed 21 Oct 2019
28. Frantz, C., Nowostawski, M.: From institutions to code: towards automated generation of smart contracts. Int. Workshops Found. Appl. Self-Syst. **1**, 210–215 (2016)
29. Yuan, R., Xia, Y.-B., Chen, H.-B., Zang, B.-Y., Xie, J.: ShadowEth: Private Smart Contract on Public Blockchain. J. Comp. Sci. Technol. **33**(3), 542–556 (2018). https://doi.org/10.1007/s11390-018-1839-y

Determination of a Dedicated, Cost-Effective Agility in Manufacturing Networks

J. Ays(✉), A. Gützlaff, K. Thomas, F. Berbecker, and G. Schuh

Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
j.ays@wzl.rwth-aachen.de

Abstract. Global manufacturing companies are operating their manufacturing networks in an increasingly uncertain and complex environment. Consequently, they have to adapt their networks much faster to changing conditions. Thus, agile manufacturing networks that are able to react and adapt quickly at reasonable cost to those conditions are required. One of the main challenges in this context is the balancing of cost- and investment-intensive measures to increase flexibility and transformability on the one hand and the impediment of the overall changeability/agility by the associated increased capital tie-up on the other. Hence, this paper presents an approach to quantify and evaluate the benefits and costs of changeability measures. By systematically discussing a reasonable trade-off between costs and benefits, a dedicated, cost-effective agility for a manufacturing network is determined, enabling companies to improve their network design in terms of cost-effective changeability and thus ensure their competitiveness in the long term.

Keywords: Agile manufacturing networks · Changeability · Dedicated agility

1 Introduction

Global manufacturing companies today operate in an ever more uncertain environment [1]. A multitude of internal and external factors has a strong impact on manufacturing networks. Internally, the design of networks is mainly influenced by the iterative cycle thinking of agile product development and short-term changing specifications in the form of rapid change requests [2]. In addition, manufacturing networks are externally influenced by today's highly volatile, uncertain, complex and ambiguous (VUCA) environmental conditions, which require disruptive changes in the network at equally short notice [1].

This results in an increasingly difficult long-term prediction of external and internal conditions for companies and a growing uncertainty that must be taken into account when designing manufacturing networks [3]. In order to counter these uncertainties, in practice and in most theories the probabilities of occurrence for scenarios are estimated with the help of subjective opinions or statistical calculations based

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 699–706, 2021.
https://doi.org/10.1007/978-3-662-62138-7_70

on existing data. Afterwards, the manufacturing network is designed as optimally as possible by means of e.g. scenario analysis [4, 5, 6]. However, the challenge here is to assume the possible prediction of such probabilities and thus to assume the pure existence of risk situations. In practice, it has been shown that due to the high complexity of networks, among other factors, serious misjudgments are regularly made and networks must change drastically [7]. It can therefore be deduced that uncertainty scenarios without predictable probabilities of occurrence (e.g. so-called black swan events) do exist [8]. The last few years have shown a wide variety of disruptive examples in the areas of political stability, currency development or epidemics e.g. the Brexit or the COVID-19 crises, which imposed manufacturing companies with disruptive changing environmental settings on which they had to react on.

In order to be able to cope with these uncertainties, today's companies must actively face the existing dynamic environment and consciously shape their own changeability for long-term resilience. To achieve this, agile network structures are needed that are highly responsive, cost-efficient and ideally proactive in dealing with change [9]. One of the biggest challenges in designing these agile manufacturing networks is the estimation of costs and benefits of agility, since without probabilities of occurrence, statistical methods are not applicable and benefits are difficult to quantify [10].

Therefore, the aim of this paper is to determine systematically a dedicated, cost-efficient agility for global manufacturing networks. The question is to be answered, which dedicated level of agility is reasonable for a global manufacturing network and which agility costs seem to be appropriate. For this purpose, first the term agility in the context of manufacturing networks will be characterized in more detail. This is followed by a methodical consideration of the costs and benefits of measures to increase agility. A meaningful quantification of the agility benefits ensures the comparability of the measures. In the final step, a company-specific trade-off between costs and benefits is performed, which also considers the tied-up fixed costs associated with the measures to ensure agility in all terms.

2 Prior Research

In the following, a short overview of relevant existing literature on the design of changeability and agility in manufacturing networks under consideration of associated costs and benefits serves to classify the paper.

Basically, two approaches can be distinguished in the literature for dealing with uncertainty in the design of manufacturing networks. On the one hand, the best possible design of a network for a defined probable operating point is sought in order to keep the uncertainties and the associated change as small as possible. For this purpose, complex mathematical models are designed, which try to determine the best possible design e.g. by means of multi-criteria optimization [4], and a continuous improvement of predictions is targeted, e.g. by using neural networks [6]. On the other hand, the intentional preparation for change is focused in the topic area of changeability. This changeability is much discussed in the literature, as an overall term for the various manifestations such as flexibility, transformability or agility [11].

Flexibility refers to the ability to change at reasonable cost within a predefined flexibility corridor (e.g. due to not fully utilized capacities) to react on fluctuating demand. In case of a minor change request, which forces the company to adjust the volume beyond the flexibility corridor, transformability is needed. Transformability describes a possible shift of the flexibility corridor within an anticipated transformability corridor (e.g. an increased universality of machines can be named here, which allows a fundamental change of the machine within a certain framework) [12]. In a transformable system, the reachable states are not limited a priori [13]. However, if a major change request hits the network, the adaption within this corridor may not be adequate anymore and a disruptive change may be necessary beyond any former expectations.

In literature, the term agility is understood as a further increase of transformability and is associated in particular with manufacturing networks, as well as with strategic and proactive action [11, 14]. Nevertheless, this term is still the least concrete or most inconsistently used one. For this reason, a proposal for a definition of agility in the context of manufacturing networks is made in the following chapter.

Some suggestions for the design of changeability in networks exist in literature. Monauni proposes a network-wide capacity pool, which should be built up redundantly or complementary, depending on the type of volatility [15]. Mikusz et al. examine different concepts of changeability in the network and present six hypotheses for the design of changeability [16]. Ramsauer et al. suggest concrete measures to increase agility, but only partly refer to manufacturing itself [14]. Fisel et al. examine the frequency and amplitude of external factors to derive suitable countermeasures [12]. Lyu et al. present an approach to network-wide capacity configuration and allocation [17]. Sudhoff examines the mobility of manufacturing systems between two locations as an enabler of change and considers the monetary benefits [18]. Moser et al. develop an approach for calculating the potential of changeability measures based on given probability scenarios [19].

In the mentioned approaches, there is often a lack of practical concretization or a focus on a certain sub-area of changeability. Furthermore, a systematic, monetary evaluation of proposed measures of changeability is hardly mentioned or not elaborated in detail. However, there are some economic considerations of the changeability of a single location [20]. Moser et al. present the most applicable approach for the cost-benefit assessment of the ability to change, but base all considerations on existing probabilities and thus ignore unforeseeable events [19].

3 Concept of a Dedicated, Cost-Effective Agility

The presented approach allows to evaluate all types of changeability measures (flexibility, transformability or agility measures) in terms of cost and especially monetary benefits and suggests a reasonable trade-off between measures to enable an agile manufacturing network. Thereby, special focus lays on the consideration of unexpected events without a given probability. As a basis for this, the term agility is first defined in the context of manufacturing networks and put in relation to the terms flexibility and transformability. This is followed by an evaluation of the benefits and costs of the various measures, and the trade-off between different measures.

3.1 Characteristics of Agile Manufacturing Networks

Agility enables to deal with changeability needs outside transformability corridors by quickly adapting to any kind of situation: for example enter an entire new production field (product- or location-wise), and close or significantly scale down the former production field. New flexibility and transformability corridors are developed within the alternative field and the volume in the initial field is shut down (see Fig. 1). Agile manufacturing networks combine and extend all the mentioned characteristics of flexibility and transformability. They can react within the flexibility and transformability corridors as well as move these corridors by quickly entering new/closing less needed fields due to their agility.

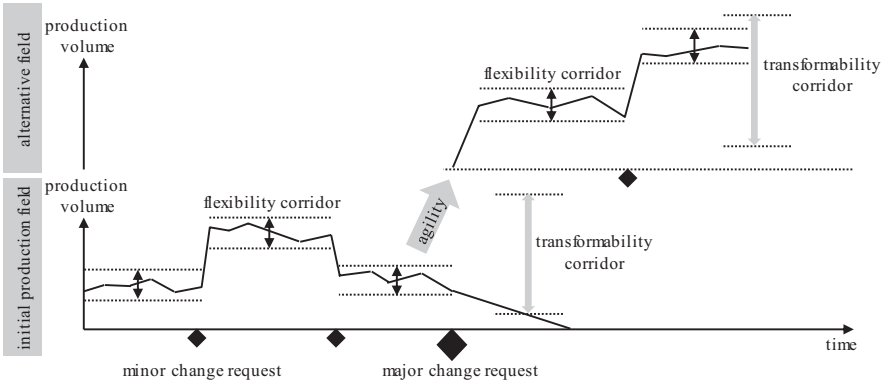


Fig. 1. Differentiation between flexibility, transformability and agility

To increase these three types of changeability, companies can use different measures. For example, various possibilities exist when an investment for a new machine is planned. For a flexibility increase, a machine with many different functions would be beneficial. A transformable machine could be expandable due to its modular structure. However, agility could be reached by renting the machine with a short contract period instead of buying it or buying the least expensive machine to reduce sticky costs in the case of change.

3.2 Benefit and Cost Valuation of Agility Measures

When a decision needs to be made between competing agility measures, a cost and benefit valuation needs to be done for every individual measure. The total costs of a projected measure ($C_{measure}$) can generally be calculated by addition of the operating expenditures (C_{opex}) over the valuation period (t_{VP}) and the capital expenditures (C_{capex}) (Formula 1). The valuation period should have the same duration for all competing agility measures. Regarding the expected benefit of a measure, a differentiation between flexibility and transformability/agility perspective is made due to the diverging underlying assumptions. The flexibility assessment is based on the classical

estimated probabilities of scenarios to build up the most probable flexibility corridor in the first place, while the assessment of transformability or agility needs to be evaluated differently to counter unknown and unexpected events and changes. However, in some cases of transformability the database for the scenario probabilities might be sufficient enough to also base the benefit evaluation on estimated probabilities.

For the expected flexibility benefit ($E(B_{flex})$) (formula 2), different capacity scenarios ($s = 1..n$) must be considered. The profit of the scenario ($E(P_s)$) is the amount that could be additionally generated with the measure in case of the scenario (p_s). Multiplying the result with an overall forecast reliability (FR_{ps}) leads to the expected flexibility benefit of the measure. This is done to differentiate between different classes of probability estimations (e.g. 1.00 for a high reliability down to 0,25 for a low reliability of the probability estimation) and counter the rising uncertainty (and not the rising risk) (compare Table 1). Hence, the forecast reliability reduces the expected benefit in cases of a weak information basis. This leads in many cases to worse business cases for flexibility measures, but should be considered to increase agility, due to less sticky, tied-up capital costs, instead of flexibility.

Table 1. Forecast reliability of probabilities or factors

Forecast is based on extensive data information and experience	1.00
Forecast is based on moderate data information or experience	0.75
Forecast is based on some references from the past	0.50
Forecast is based on the gut feeling of the decision-makers	0.25

Before the transformability/agility benefit of measures can be calculated, a substitute element for the former used probability must be found. The authors propose therefore the degree of volatility (DOV) (Formula 3). Because reliable probabilities cannot be estimated in an uncertain environment, the DOV describes the general volatility in the decision-making environment. First, the most relevant internal and external influence factors ($i = 1..m$) must be nominated. Examples for influence factors could be the economic situation in a specific country (external factor) or the salary structure within a company (internal factor). The development of each influence factor should already be considered in the overall network design. Thus, the question is how likely the assessed factors will deviate from the expected path and how strong a possible impact (IM_i) in case of a deviation would be. The first estimation is done based on the introduced forecast reliability in Table 1 (FR_{ps}), while the second could basically be determined with: high (1.00), moderate (0.75), low (0.50). By this, high-impact factors with a weak information base imply a high uncertainty resp. volatility of the manufacturing environment, in which the manufacturing networks could benefit more from transformability and agility measures. The last part of the DOV is the time factor (TF) which is based on the duration of the valuation period (t_{VP}). A long period leads to a high time factor ($TF = 1.0$), while a short period leads to a low time factor ($TF = 0.75$) due to a more manageable data basis and probably better forecasts.

By using the DOV, the expected benefit from the transformability/agility perspective ($E(B_{tran/agil})$) can be calculated (Formula 4). In any case of a change request on the manufacturing network there will be costs for the company. The main goal of measures within the transformability/agility perspective is the reduction of those

future costs. The main four examples of those costs are given in the method. First, the sticky costs which are not needed anymore after a change request but cannot be cut directly (C_{sticky}) [10]. Second, contractual penalties which the company needs to pay due to delay in delivery ($C_{penalty}$). Third, the company's lost profit due to the necessary change and the waste of capacities ($C_{lostprofit}$). Last, costs of a new ramp-up phase, in case of changing the production (field) (C_{rampup}).

$$C_{measure} = C_{capex} + C_{opex} * t_{VP} \tag{1}$$

$$E\left(B_{flex}\right) = \sum_{s=1}^{n} \left(E(P_s) * p_s * FR_{p_s}\right) \tag{2}$$

$$DOV = \frac{\sum_{i=1}^{m} ((1 - FR_i) * IM_i)}{m} * TF \tag{3}$$

$$E\left(B_{tran/agil}\right) = \left(\Delta C_{sticky} + \Delta C_{penalty} + \Delta C_{lostprofit} + \Delta C_{rampup}\right) * DOV \tag{4}$$

$$E(P_{measure}) = E\left(B_{flex}\right) + E\left(B_{tran/agil}\right) - C_{measure} \tag{5}$$

C	costs [€]	t_{VP}	valuation period [yrs]
E(B/P)	expected benefit/profit [€]	FR	forecast reliability [%]
s	scenario 1…n	p_s	probability of scenario s [%]
DOV	degree of volatility [%]	TF	time factor
i	influence factor 1…m	IM_i	impact of factor [0.50;0.75;1.00]

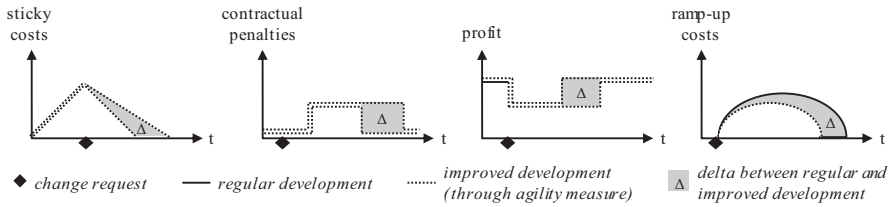


Fig. 2. Expected benefit from transformability/agility perspective

When a company can lower any of those listed costs, the delta is equivalent to the generated benefit (see Fig. 2). The expected benefit in terms of transformability/agility is the potential profit/cost reduction in case of a change request multiplied with the DOV. Finally, the expected overall benefit of a measure ($E(P_{measure})$) can be calculated by balancing all calculated benefits and costs (Formula 5).

3.3 Determination of Applicable Measures

After all possible measures are valuated, a trade-off between the various measures for a production task is necessary. A useful measure must fulfil the following two criteria.

First, the overall expected profit (E(P)) must be positive which means that the costs during the valuation period must be lower than the calculated expected benefit from flexibility and transformability/agility perspective.

Second, the amount of new sticky fixed costs ($C_{new\ sticky\ fixed}$) should be lower than the expected profit. The fixed cost quotient (FCQ) be the indicator to determine the balance between expected profit and new sticky fixed costs (Formula 6). New sticky fixed costs are all capital and operational expenditures within the valuation period, which cannot be cut back easily in the future, for example capital investments or binding contracts. If the FCQ is lower than one for a measure, the measure causes too much sticky fixed costs which increases the risk of misinvestments due unconsidered effects and reduces the effect of total agility increasement.

$$FCQ = \frac{E(P)}{C_{new\ sticky\ fixed}} \tag{6}$$

In the final step, all remaining measures can be ranked by their expected profit and the fixed cost quotient. The management must make a final decision based on the classifications of the measures and their overall strategy.

4 Conclusion

In today's fast changing environment, companies must be able to determine a dedicated and cost-effective agility in their manufacturing network. The described concept will support them in this process. After the definition of agility in manufacturing networks, the main method of cost and benefit valuation of agility measures was presented and a trade-off between costs and benefit was made. The concept offers a practical approach, which enables companies to determine the best possible agility measure for their production task. An evaluation of the approach in practice is still pending, but is currently being conducted.

Although most of the mentioned factors are easy to understand, the specific determination of each input value may be challenging for the users of the concept. In further research, the weight of each factor must be focused to better understand the influences of the input data. Moreover, the relation between the degree of volatility and the effect of agility measures is an important future research field, which helps to develop the right, dedicated level of agility in manufacturing networks.

Acknowledgements. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

1. Millar, C.C.J.M., Groth, O., Mahon, J.F.: Management innovation in a VUCA world: challenges and recommendations. Calif. Manage. Rev. **61**(1), 5–14 (2018)
2. Schuh, G., Gartzen, T., Soucy-Bouchard, S., Basse, F.: Enabling agility in product development through an adaptive engineering change management. Procedia CIRP **63**, 342–347 (2017)
3. Krebs, P., Reinhart, G.: Evaluation of interconnected production sites taking into account multidimensional uncertainties. Prod. Eng. Res. Develop. **6**, 587–601 (2012)
4. Lanza, G., Moser, R.: Multi-objective optimization of global manufacturing networks taking into account multi-dimensional uncertainty. CIRP Ann. – Manuf. Technol. **63**(1), 397–400 (2014)
5. Olivares-Aguila, J., ElMaraghy, W.: System dynamics modelling for supply chain disruptions. Int. J. Prod. Res. **58**, 1–19 (2020)
6. Bandara, K., Shi, P., Bergmeir, C., Hewamalage, H., Tran, Q., Seaman, B.: Sales demand forecast in e-commerce using a long short-term memory neural network methodology. In: International Conference on Neural Information Processing, Springer, pp. 462–474 (2019)
7. Sakai, Y.: On the economics of risk and uncertainty: a historical perspective. Discussion Papers CRR Discussion Paper Series A: General 28, Shiga University (2018)
8. Bevilacqua, M., Ciarapica, F.E., Marcucci, G.: A modular analysis for the supply chain resilience triangle. IFAC-PapersOnLine **51**(11), 1528–1535 (2018)
9. Schuh, G., Prote, J.-P., Franken, B., Ays, J., Cremer, S.: Dedicated agility: a new approach for designing production networks. In: 2018 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM), Bangkok, pp. 1–5 (2018)
10. Schuh, G., Prote, J.-P., Gützlaff, A., Ays, J., Donner, A.: Fixed cost management as an enabler for agile manufacturing networks. Procedia Manuf. **39**, 625–634 (2019)
11. Wiendahl, H.-P., Reichardt, J., Nyhuis, P.: Handbook Factory Planning and Design. Springer, Heidelberg (2015)
12. Fisel, J., Duffie, N., Moser, E., Lanza, G.: Transformability – a frequency perspective. Procedia CIRP **79**, 186–191 (2019)
13. Zaeh, M.F., Moeller, N., Vogl, W.: Symbiosis of Changeable and Virtual Production – The Emperor's New Clothes or Key Factor for Future Success. In: International Conference on Changeable, Agile, Reconfigurable and Virtual Production (2005)
14. Ramsauer, C., Kayser, D., Schmitz, C.: Erfolgsfaktor Agilität. Chancen für Unternehmen in einem volatilen Marktumfeld. Wiley, Welnheim (2017)
15. Monauni, M.: Agility enablers in production networks: pooling and allying of manufacturing resources. Procedia CIRP **17**, 657–662 (2014)
16. Mikusz, M., Heber, D., Katzfuß, C., Monauni, M., Tauterat, T.: Changeable manufacturing on the network level. Procedia CIRP **41**, 27–32 (2016)
17. Lyu, G., Cheung, W.C., Chou, M.C., Teo, C.P., Zheng, Z., Zhong, Y.: Capacity allocation in flexible production networks: theory and applications. Manage. Sci. **65**(11), 5091–5109 (2019)
18. Sudhoff, W.: Methodik zur Bewertung standortübergreifender Mobilität in der Produktion. Utz, München (2007)
19. Moser, E., Huss, A.K., Liebrecht, C., Lanza, G.: A portfolio theory approach to identify risk-efficient enablers of change in global production networks. Procedia CIRP **63**, 768–773 (2017)
20. Klemke, T.: Planung der systematischen Wandlungsfähigkeit von Fabriken. Dissertation, Leibniz Universität Hannover (2013)

Requirements for an Event-Based Visualization of Product Complexity

J. Koch, C. Dölle, M. Riesener, and G. Schuh(✉)

Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
jan.koch@wzl.rwth-aachen.de

Abstract. Analysis of data is currently being discussed in research and practice across a wide range of areas, but mainly with focus on production. Besides production, this is also relevant in areas such as complexity management. A major challenge in complexity management is the trade-off between standardization and customer-specific solutions. This is strengthened by increasing cost pressure, shortened product life cycles and a rising speed of innovation. Therefore, companies need transparency with regard to product-induced complexity in order to respond with an appropriate product variety and complexity management. In this paper a methodology is presented, which defines design guidelines for a targeted and event-based visualization of stakeholder-specific information requirements. To achieve this, a requirement morphology is developed to describe general as well as complexity-specific visualization requirements. On that basis, a dashboard structure is derived and a stakeholder-specific assignment of visualization types is considered.

Keywords: Complexity management · Data analytics · Data visualization

1 Introduction

Manufacturing companies across all industries are faced with the challenge of producing and offering economically and technically marketable products in the area of conflict between standardization and implementation of customer-specific requirements [1, 2]. The transformation from a supplier's to a buyer's market and the differentiation of customer needs have led to an increase in number of variants [3]. This development is intensified by increasing cost pressure, reduced development times, shorter product life cycles and increasing speed of technology as well as innovation. Many companies are therefore pursuing a differentiation strategy with the aim of meeting customer requirements more effectively than competitors are and thus countering increasing price pressure [4]. As a result, companies have to reduce costs along the value chain by implementing customized solutions to address increased costs. Furthermore, due to growing product programs the increase of product-induced complexity must be countered by suitable complexity management. This has to be done for both, product

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 707–716, 2021.
https://doi.org/10.1007/978-3-662-62138-7_71

program breadth through diversification as well as product program depth through variation.

However, in most cases, product-induced complexity cannot be identified and automatically analyzed by companies at an early stage. Therefore, visibility and transparency of product complexity is necessary in a first step. In a recent study on "Global complexity management", conducted by the Laboratory of Machine Tools and Production Engineering WZL of RWTH Aachen University [5], 68% of interviewed executives stated that they did not have comprehensive transparency of the entire product program and its complexity. In order to create transparency and make decisions based on data, it is necessary to identify relevant information requirements and to visualize this information in a context-based manner [6]. However, decision making in the majority of companies is not based on data-based results [7]. In this respect, another 71% of the participants in the above-mentioned study stated that data analytics methods for optimizing product complexity are currently not sufficiently applied [5]. Even employees with extensive experience typically cannot include more than three influencing factors in a systematic decision making process [7]. Therefore, this paper describes a methodology to visualize stakeholder-specific information requirements in order to create transparency on relevant events related to the management of product complexity.

2 Relevant Terminology

In the following, product complexity as well as relevant terms for the visualization of product-related information are defined.

Product complexity can be described by means of two dimensions, complexity of the individual product and complexity of variations. Features of individual product complexity include the number and variety of elements as well as variety in processes and resources needed in manufacturing. Variation complexity describes the variety of offered product variants within the product program and the resulting complexity in order processing [8]. Product complexity is defined within this paper according to Schuh and Riesener [3] as: Combination of external, i.e. market-induced, and internal, i.e. product-induced variety, which must be designed, developed and controlled in a coordinated manner [3].

Visualization according to Diehl [9] is defined as a process of transforming information into a visual form that allows the user to visually discover hidden aspects in data that are essential for exploration and analysis [9]. In literature, a differentiation is made between three areas of visualization: Scientific visualization, information visualization and visual analytics. The focus of this paper is on information visualization. Information visualization deals with the computer-generated, interactive and visual representation of discrete and mostly abstract data without physical or spatial reference [10]. The spectrum of data to be visualized is therefore wider than in scientific visualization and the specific selection of suitable visualization types is of central importance [10]. Moreover, the differentiation between user groups is an important aspect.

In the industrial context, visualization takes place mainly in the form of static reports or as data-based and customizable visualization in dashboards. The latter will be the focus of this paper. A dashboard is expected "to collect, summarize and present information from multiple sources" [11]. It shows the most important information monitored at a glance, contains a stakeholder-oriented visualization with simple usage as well as possibilities for customization. Therefore, visualization of product complexity-relevant events in dashboards enables transparency and controllability.

3 Related Work

There are numerous approaches in complexity management to master the variety demanded by the market, but there is no concept for systematic and data-based analysis as well as visualization of product complexity.

Approaches, such as Budde [12], Brosch [13] and Orfit et al. [14] consistently illustrate that for a holistic analysis of product complexity, the entire product program as well as various disciplines within the company must be considered. Moreover, the approaches include different information requirements and a key figure-based assessment. However, company data for a successful application as well as specific visualization requirements are not sufficiently addressed [12–14].

In addition, approaches focusing on a data-based management of product complexity, e.g. Schmidt et al. [15], Kreimeyer et al. [16] and Junk et al. [17], were analyzed. Data-based approaches address very specific information requirements and do not sufficiently consider a comprehensive assessment of product complexity. The consideration of visualization requirements is underrepresented and shows a need for research [15–17].

In addition, approaches for visualizing product complexity, such as Windheim et al. [18], Wu et al. [19] as well as Gebhardt and Krause [20], were analyzed. The definition of stakeholder-specific information requirements is considered to a greater extent. Nevertheless, these approaches lack a holistic view and data-based visualization [18–20].

In summary, it can be stated that no approach considers a comprehensive assessment of information requirements as well as a corresponding visualization for a holistic and stakeholder-specific analysis of product complexity. The described theoretical deficit represents the research needs of the present paper.

4 Methodology

As stated in the introduction, there is a lack of data-based support and visualization in decision-making for managing product complexity. Various approaches address the need to assess product complexity, but do not sufficiently integrate and visualize stakeholder-specific information requirements. Therefore, a methodology for an event-based visualization of product complexity is presented below.

Within the overall research project, stakeholder-specific information requirements are systematically identified for the management of product complexity. These information requirements are evaluated using key figures in a next step, which are synchronized with the data available in corresponding information systems. Therefore, real-time access to product complexity-relevant data can be achieved through a data model. In a final step, however, the information requirements must be made accessible to various stakeholders, which is to be achieved through a data-based visualization in dashboards. For this purpose, the relevant steps are described in this paper. The first step is the definition of design requirements for product complexity dashboards. This is detailed by corresponding dashboard elements and the assignment of visualization types.

4.1 Dashboard Design Requirements

The visualization of complexity-relevant questions should consider different design requirements depending on the stakeholders (see Fig. 1).

First, according to the hierarchy level of the stakeholder, a distinction must be made between strategic, tactical or operational presentation. Three superordinate dashboard types are derived for this paper based on these three characteristics. The second differentiation criterion is the purpose of analysis which is differentiated in terms of a financial focus and a domain focus according to the three typical areas of complexity management: Market, product and process. Furthermore, the scope of the analysis presented will be differentiated in the context of a company-wide, site-specific, domain-specific or individual scope. With regard to the visual presentation of complexity-relevant information, a further distinction can be made between graphical or tabular presentation as well as the flexibility of visualizations. For the data considered, quantitative or qualitative analyses are differentiated as well as the update frequency, which can range from monthly to real-time. In the following, the defined criteria as well as assigned characteristics are presented in form of a morphology, which is derived from literature and validated within industrial working group meetings [21–24].

Criteria	Characteristic				
Audience	Strategic	Tactical	Operational		
Purpose	Planning	Monitoring	Analysis	Implementation	Communication
Domain	Finance	Sales	Engineering	Production	
Scope	Company-wide	Site-specific	Domain-specific	Individual	
Flexibility	Cross-highlighting	Slicers and filters	Drill down and up	Static	
Display	Graphical	Tabular	Integration of both		
Data	Qualitative	Quantitative			
Update	Monthly	Weekly	Daily	Hourly	Real time

Fig. 1. Morphology of dashboard design requirements

For a successful management of product complexity different information requirements and different perspectives have to be considered. Therefore, three dashboard types are described below: Strategic, tactical and operational dashboards (see Fig. 2).

Strategic dashboards aim at a quick overview with simple display mechanisms. The goal of executives is to get a holistic overview on the enterprise by consolidated data from various divisions and business segments. Therefore, high-level measures are derived for long-term decisions and profitability, growth and risks are mainly in focus. Hence, strategic dashboards are not designed for interaction and do not require real time data. In the context of complexity management analyses should consider financial issues as well as high-level insights into the three domains: Market, product and process. This implies that, from a market perspective the entire product program, from a product perspective all technologies and products, and from a process perspective manufacturing and assembly at all sites is analyzed. Thus, strategic dashboards enable to assess developments of the corporate environment by considering external information as well as the concrete company development in terms of market share, sales development, innovative power, etc. Number and variety of elements considered as well as their variability are thus monitored in a highly aggregated manner and analyzed on the basis of long-term developments.

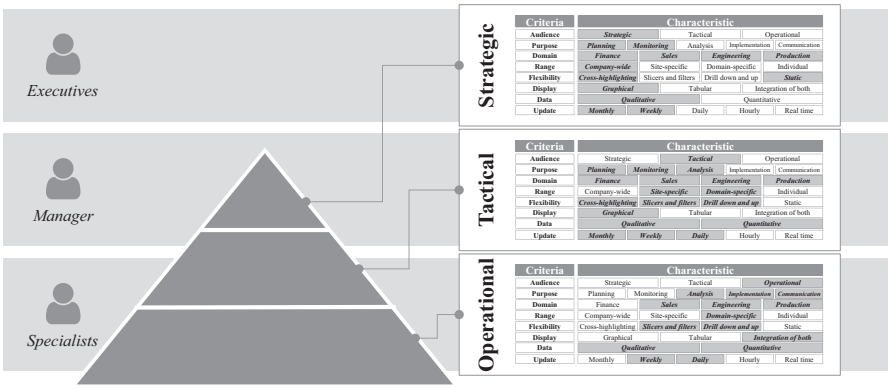


Fig. 2. Dashboard types for a comprehensive management of product complexity

Thus, tactical dashboards are applied besides planning and monitoring for analytical purposes with a divisional focus. Therefore, key figures are specified to the divisional area including more complex data, a greater context including a more extensive history, rich comparisons and a more sophisticated display. In particular, it should enable greater interaction with the data such as drill-down functions or specific slicer and filter functions compared to strategic dashboards. This includes more detailed analyses, e.g. for the market perspective with regard to customer structure and product program. Analyses visualized in a tactical dashboard for the market perspective should contribute to improving competitiveness, optimizing the product range or increasing customer satisfaction.

In contrast, operational dashboards provide more specific information at a deeper level of detail for monitoring operations. This implies that analyses are detailed down to product variant level for the market perspective or to component level for the product perspective. Therefore, the focus here is less on current trends like as-is vs. to-be comparisons or profitability analyses, but rather on searching and finding relevant information as well as visualizing the status quo. As a result, operational activities of complexity management are ensured, such as increasing the reuse of existing components in new product development projects.

In summary, this step suggests to distinguish between three different dashboard types for the management of product-induced complexity in order to address information requirements sufficiently based on the criteria described above.

4.2 Dashboard Structure and Elements

In the following, the structure as well as relevant elements are derived for the design of dashboards for product complexity. The overall dashboard structure consists of four areas. First, the integration of a role concept and thus the associated adaptation to specific stakeholder needs. The filtering with regard to time and relevant analysis objects for managing product complexity as well as the layout for visualizing the different analyses (see Fig. 3).

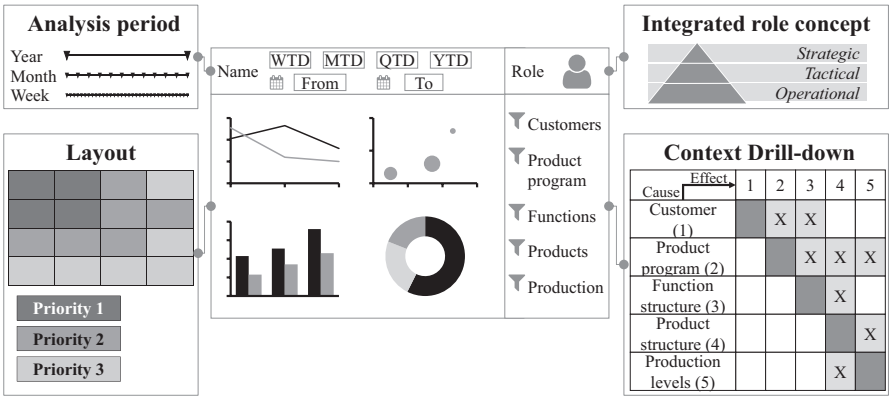


Fig. 3. Dashboard structure and elements

The analyzed time span can be differentiated between Year to Date to Week to Date or based on an individual analysis period. Moreover, the integrated role concept enables a consideration of stakeholder-specific information after login and to store individualized settings. A central element for the analysis of product complexity are drill-down functions for different complexity relevant objects to be analyzed. A distinction for the analysis of product-induced complexity is made between customer structure, product program, functional structure, product structure and different production levels. The customer structure enables filtering based on different market

segmentation criteria, the product program between levels of the product family through product lines and product groups to product variants, the function structure from overall functions to sub-functions, the product structure from modules, assembly groups to components and the production levels from site to areas, lines and stations. Similar to the functionality of a multidimensional data model, the dashboard should realize filtering between different levels of each analysis object as well as across the different objects. These filter functions are of particular interest for more detailed analyses and thus for tactical and operational dashboards. Relevant objects for drill down from market, product or process perspective were determined by a cause-effect analysis. This analysis provides information on which perspectives are affected by which objects and therefore need to be considered for drill down analyses in the dashboard (see Fig. 3). The market perspective includes customer structure and product program. The product perspective is affected by all dimensions. Therefore, all filter options with regard to the mentioned analyses objects must be provided. For the process perspective, in addition to the production levels, the product program and the product structure must also be taken into account. The product program shows the greatest impact across all perspectives. For the layout as main area of the dashboard, number and structure of analyses have to be considered. The layout can be divided into three zones with different priority levels based on an eye-tracking study [11]. In most cases the view falls on the upper left area, therefore the most important analyses should be presented here. Along the diagonal line from top left to bottom right the priority decreases. Furthermore, for the viewer a presentation of up to maximum six to seven analyses is currently sufficient. Therefore, it is deduced that for the strategic dashboard four graphs should be considered for the corresponding three perspectives as well as financial information. In the top left of the display, the financial information should be presented as central criteria. For tactical as well as operational dashboards, the presentation of six analyses will be focused, whereas for operational dashboards not everything might be visible at a glance and have to be added by further tabs.

Thus, dashboard structure and its relevant elements for the previously defined strategic, tactical or operational dashboards were defined.

4.3 Visualization Requirements

The final step of the presented methodology focuses on a procedure for assigning specific visualization types to identified information requirements with regard to the management of product complexity. The following section therefore describes how to assign visualization types to these attributes (see Fig. 4).

The assignment of visualization tasks to complexity attributes is based on Abela [25]. The following four visualization tasks are considered, depending on what your data should demonstrate: *make comparisons, analyze the compositions, show relationships* or *display distributions* of a data set. For the visualization of information requirements in complexity management, these tasks are assigned to the complexity-relevant attributes *variety, diversity, relationships* and *variability*. Visualization types that represent a static comparison between different values are suitable for visualizing a large number of elements. Therefore, comparison-oriented

visualization types should be used to visualize the variety of elements. The diversity of elements should be analyzed by identifying differences or similarities. This is realized by visualizations with focus on displaying and analyzing compositions. The relationship between elements can be visualized by scatter or bubble charts. The variability of values over time can be analyzed by comparison or composition of elements over time. The most frequently used visualization types for this complexity attribute are line, area or column charts.

Therefore, the last method step provides an assignment of visualization types to complexity relevant information requirements according to the overall complexity attributes.

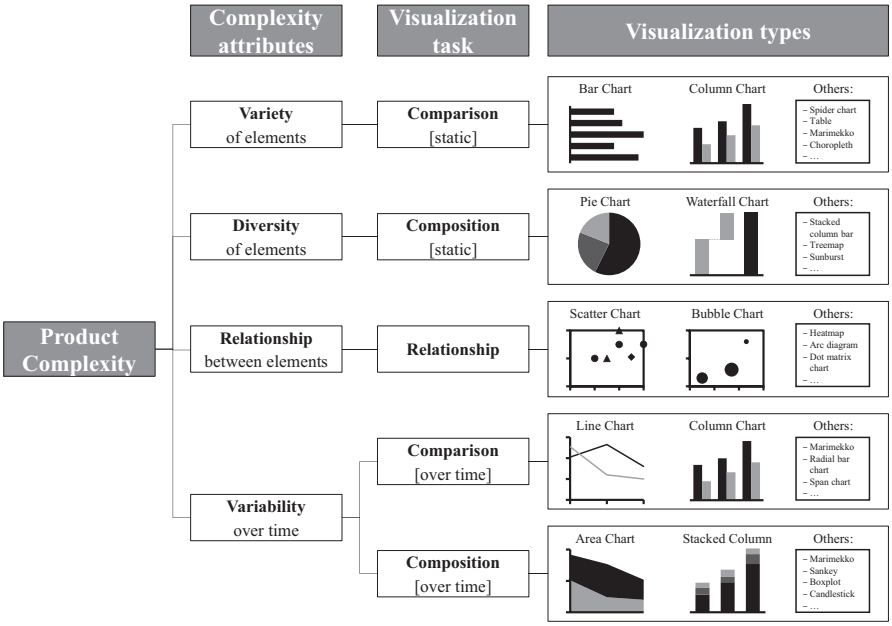


Fig. 4. Assignment of visualization types to complexity attributes

5 Conclusion and Further Research

Growing complexity of technical systems and shortening product lifecycles in a globalized competition lead to a demand for data-based transparency and visualization of stakeholder-specific information requirements. This paper presents a methodology for the design of complexity dashboards with the aim of creating visibility and transparency of product complexity-relevant events. Therefore, three steps are introduced to illustrate the approach. In a first step, visualization requirements were derived and classified with regard to different characteristics for the considered stakeholders. In a second step, a generic structure for visualizing information requirements in a dashboard was derived and detailed for the different perspectives for managing product

complexity. The third step concludes with an assignment of visualization types with regard to the different analysis requirements.

Further research has to focus on the specification of each step and the validation of the proposed methodology within several industry case studies. The validation and further investigation of the presented methodology is currently part of the research activities of the authors.

Acknowledgements. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

1. Riesener, M., et al.: Information requirements for a data-based analysis of product and service complexity. Procedia CIRP **83**, 279–284 (2019)
2. Kwasniok, T., Kilimann, S.: Mastering Product Complexity, Düsseldorf (2012)
3. Schuh, G., Riesener, M.: Produktkomplexität managen. Strategien - Methoden - Tools, 3rd edn. Hanser, München (2017)
4. ElMaraghy, H., et al.: Product variety management. CIRP Ann. **62**(2), 629–652 (2013)
5. Schuh, G. et al.: Globales Komplexitätsmanagement. Zusammenfassung der Ergebnisse des Konsortial-Benchmarkings (2017)
6. Schuh, G. et al.: Industrie 4.0 Maturity index. Managing the digital transformation of companies. Update 2020. acatech - National Academy of Science and Engineering (2020)
7. Blase, P. et al.: Data-driven: Big decisions in the intelligence age (2016)
8. Closs, D.J., et al.: The differential impact of product complexity, inventory level, and configuration capacity on unit and order fill rate performance. J. Oper. Manage. **28**(1), 47–57 (2010)
9. Diehl, S.: Software Visualization. Visualizing the Structure, Behaviour, and Evolution of Software. Springer-Verlag, Berlin (2007)
10. Telea, A.: Data visualization. Principles and practice. CRC Press, Boca Raton (2015)
11. Yigitbasioglu, O.M., Velcu, O.: A review of dashboards in performance management. Implications for design and research. Int. J. Accounts Inf. Syst. **13**(1), 41–59 (2012)
12. Budde, L.: Integriertes Komplexitätsmanagement in produzierenden Unternehmen. Ein Modell zur Bewertung von Komplexität. Dissertation (2016)
13. Brosch, M.: Methode zur Reduzierung der produktvarianteninduzierten Komplexität (2015)
14. Orfi, N. et al.: Harnessing Product complexity. Step 1 - establishing product complexity dimensions and indicators. Eng. Econ. **56**(1), 59–79 (2011)
15. Schmidt, M. et al.: Graph-based similarity analysis of BOM data to identify unnecessary inner product variance. In: Proceedings of the 21st International Conference on Engineering Design. vol. 1: Resource Sensitive Design, Design Research Applications and Case Studies, pp. 489–498. The Design Society, Glasgow (2017)
16. Kreimeyer, M. et al.: An integrated product information model for variant design in commercial vehicle development. In: International Design Conference - Design 2016, pp. 707–716, Dubrovnik, Croatia (2016)
17. Junk, C. et al.: Model-Based Variant Management with v.control. In: Stjepandić, J., Borsato, M., Wognum, N., Curran, R., Verhagen, W.J.C. (eds.) Transdisciplinary lifecycle analysis of systems. Proceedings of the 22nd ISPE Inc. International Conference on Concurrent

Engineering, July 20–23, 2015. Advances in transdisciplinary engineering, vol. 2, pp. 194–202. IOS Press, Amsterdam (2015)
18. Windheim, M. et al.: Towards a decision-making framework for multi-criteria product modularization in cooperative environments. 28th CIRP Design Conf. **70** 380–385 (2018)
19. Wu, Q., et al.: Multi-dimensional information visualization analysis of business circles and products based on density clustering. IFAC-PapersOnLine **51**(17), 811–816 (2018)
20. Gebhardt, N., Krause, D.: A Method for designing visualisations as product development tools. In: DS 84: Proceedings of the DESIGN 2016 14th International Design Conference, pp. 611–620 (2016)
21. Rainer, R.K., et al.: Introduction to information systems. Supporting and transforming business. Wiley, Hoboken (2014)
22. Few, S.: Information dashboard design. The effective visual communication of data. O'Reilly, Beijing (2006)
23. Gluchowski, P. et al.: Management Support Systeme und Business Intelligence. Computergestützte Informationssysteme für Fach- und Führungskräfte, 2nd edn. Springer, Berlin (2008)
24. Kerzner, H.: Project Management Metrics, KPIs, and Dashboards. A Guide to Measuring and Monitoring Project Performance, 1st edn. Wiley, Hoboken (2011)
25. Abela, A.: Advanced Presentations by Design. Creating Communication that Drives Action, 2nd edn. Wiley, New York (2013)

Design of Tailored Subscription Business Models – A Guide for Machinery and Equipment Manufacturers

Y. Liu(✉), A. Gützlaff, S. Cremer, T. Grbev, and G. Schuh

Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Campus-Boulevard 30, 52074 Aachen, Germany
y.liu@wzl.rwth-aachen.de

Abstract. Machinery and equipment manufacturers especially from high wage countries are increasingly under pressure to extend their competitive differentiation. Studies show that innovation in business models can create stronger differentiation than innovation in products and processes. Especially in the software industry, subscription-based business models have recently formed new champions who dominate the markets. Unfortunately, it is so far unclear what "subscription" means in the machinery and equipment industry, what offers it can contain and how the offers can be delivered. This work supports machinery and equipment manufacturers by answering these three fundamental questions. Firstly, an adequate definition of subscription business models in the machinery and equipment industry is introduced. Secondly, a holistic understanding of potential offers within a subscription business model is provided. Thirdly, supporting success factors and requirements for the creation of the offers are identified.

Keywords: Machinery and equipment industry · Subscription · Business model

1 Introduction

The German machinery and equipment (hereinafter M&E) industry generates appr. EUR 233 bn in revenue each year. It ranks third place in the global competition after China and the USA [1]. As competitors from emerging countries are catching up on innovative capability but remaining at relatively low cost structures [2], the competition is becoming increasingly fierce. Especially companies from high wage countries need to find ways to differentiate from competitors. Traditional means of product and process innovations seem no longer sufficient. Instead, innovations in business models are gaining more importance as they can fulfill the customers' needs better [3].

Among various new business models, the so-called subscription-based business models (hereinafter subscription models) have gained particularly broad attention in the recent years. Especially in the software industry, new champions were formed

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 717–727, 2020.
https://doi.org/10.1007/978-3-662-62138-7_72

who now dominate the markets. Subscription companies grow their revenue in average more than nine times faster than the S&P 500[1] [4]. It is because of this success that manufacturers from the M&E industry are also developing and testing subscription models [5]. The potential benefits for both customers and manufacturers are manifold and partly comparable to the software industry. On the customers' side, these benefits include, for example, higher financial flexibility and customer centricity. On the manufacturers' side, benefits include, for example, better financial planning basis and a long-term customer relationship [4]. At the same time, manufacturers may also be able to monetize Industrie 4.0 potentials more effectively than so far [6].

By looking at existing successful subscription models, it also becomes obvious that a variety of customer-centered values need to be generated [4, 7]. This implies that M&E manufacturers can only be successful if their customers, i.e. industrial companies that are using the machines and equipment, to be successful in their own markets. One of the major keys for any industrial company is to constantly increase the productivity [8]. Yet despite all digitalization efforts in the recent years, the German manufacturing industry has experienced a total productivity increase of only 1.3% between the years 2011 and 2018 [9]. This phenomenon is also referred to as the "productivity paradox" [8]. By helping their customers to increase their productivity, manufacturers of the M&E industry can strengthen their own competitive position.

The introduction of subscription models is a new frontier in the M&E industry. Since it is significantly different from the software industry, a direct imitation of existing subscription models is not possible. The complexity of customers' needs and potential solutions to meet them poses a great challenge to M&E manufacturers. Therefore, the authors of this work aim to answer three fundamental questions in order to support M&E manufacturers in tailoring their own, unique subscription model:

1. What does "subscription" mean in the M&E industry?
2. Which offers can be included in subscription models?
3. Which success factors and requirements need to be considered?

2 Research Approach

In order to answer these questions, an extensive literature research was conducted. It was found that subscription models have not yet gained much attention in the production research due to its novelty in the manufacturing industry. However, service-related business models, product service systems and smart services partly show similarities to subscription models. The literature research was therefore extended to these academic fields in order to identify relevant existing approaches.

[1] Standard & Poor's 500: Stock market index comprising 500 large companies listed on stock exchanges in the United States of America.

It was found that the research questions can only be partly answered with existing approaches. They provide a basic understanding for subscription businesses in general, but no definition for its application specifically in the M&E industry. Only few works provide a selection of potential offers, but no extensive overview and structure are available. Success factors and requirements are only singularly mentioned in some literature. However, no holistic overview or model was found.

Based on these findings, a case study of 107 existing subscription models was conducted. While the literature review was mostly focused on the M&E industry, the case studies included subscription models both from within and outside of the M&E industry. Especially successful cases were found in the software, service and entertainment industries where subscription models are widespread by now. The case studies were conducted by means of scientific publications, company information, third party reports and personal interviews.

In the first step, an adequate definition of "subscription models in the M&E industry" is developed. The elementary subscription characteristics is provided by existing literature, e.g. by Gassmann et al. [10]. However, according to Osterwalder and Pigneur, business model innovation is about creating customer value [11]. Based on the same idea, this work expands the elementary definition by outlining core value propositions for subscription customers in the M&E indsutry. This is done by extracting the most constitutive value propositions from case studies of existing, successful subscription models. However, since the vast majority of the analyzed cases are located outside of the M&E industry, their transferability to the M&E industry is also analyzed.

In the second step, potential offers are identified which can create the value propositions. Three sources are used: traditional offers within the M&E industry based on literature review, existing subscription offers within the M&E industry based on case studies, and existing subscription offers outside of the M&E industry based on case studies. As a continuous increase of productivity is one of the major challenges for the customers of the M&E industry, i.e. manufacturing companies, the identified offers are especially evaluated on their potential to increase the machines' and equipment's performance and utilization.

In the third step, success factors which can increase the fulfillment level of the value propositions from step 1 and therefore increase the customer value are identified. While success factors from existing subscription models within the M&E industry can be directly considered, those from subscription models outside of the M&E industry need to be evaluated on their transferability.

Based on the aggregated results of value propositions, offers and success factors, requirements for the M&E manufacturer are identified in the fourth step. These requirements would enable the transformation and the operation of subscription models. They include both the technical and the organizational dimension.

The research approach is visualized in Fig. 1.

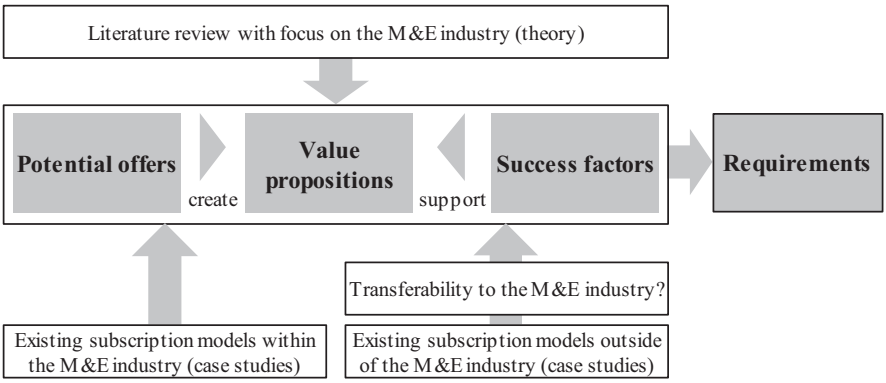


Fig. 1. Research approch for the presented work.

3 Definition of "subscription" in the Machinery and Equipment Industry

Traditionally, subscription models describe recurrent payments for recurring values. They were first applied in the book and newspaper industry in the 17th century [10]. Today's successful subscription models especially emphasize the importance of *customer centricity* and suggest a wider definition in order to capture the nature of their success [4]. Consequently, an appropriate understanding of modern subscription models needs to be developed. Based on the paradigm *customer centricity*, the authors of this work identified five core value propositions for the M&E industry. They are briefly introduced in the following.

Customized Offers Providers can greatly increase the customers' satisfaction by focusing on their problems instead of high quality products alone. Since all customers are different, individual solutions are required [12]. One way to do so is to provide customized physical products, digital products and services. Furthermore, a customized mix of standardized offers may also generate customized solutions but still maintaining an economic cost structure.

Continuous Productivity Increase Increasing the customer' productivity only once is not enough. The case study shows that subscription customers also need to be provided with new offers on a regular basis in order to achieve long-term satisfaction. Likewise, M&E manufacturers not only need to continuously improve their physical products, digital products and services, but also frequently generate new offers [6]. By doing so, the customers' productivity can be continuously improved.

Flexibility Subscription customers seem to value the freedom in products and services over the actual price [13]. In consequence, M&E manufacturers need to allow their customers to change between the products and services flexibly. Although this might be difficult to apply to the machines, it might be possible to apply to machine components, given a modular design. Customers also enjoy a significant financial

flexibility since capital expenditures (CAPEX) for the purchasing of machines are turned into operational expenditures (OPEX).

Risk Sharing In order to establish a trustful long-term relationship with the customers, successful subscription providers take upon more risks than in traditional business models [14]. At the same time, risk sharing creates an intrinsic motivation for providers to improve their performance continuously. The risks and potential benefits for the providers need to be reflected in a smart pricing strategy. Also, providers must not offer subscription model to any customer, but select the promising ones with care [15].

Convenience Company processes are often distinguished between core, support and management processes [16]. M&E manufacturers can create additional value by helping customers to focus on their core processes [17]. This can be done by either simplifying the customers' non-core processes or by carrying out the processes for them. However, M&E providers need to find ways to deliver this core value proposition by digital means in order to maintain an economic cost structure.

Based on these core value propositions, a more adequate definition for subscription models in the M&E industry was developed (see Fig. 2). It is worth mentioning that more value propositions and facets of value propositions were identified than described above. However, their transferability to the M&E industry was found hardly possible.

Subscription business models in the machinery and equipment industry describe a long-term relationship between suppliers and customers who receive recurring values for recurring payments. Based on consequent customer centricity, all generated values are designed towards the customers' problems and needs, especially to a continuous increase of their productivity. The values are transported not only by physical products, but also by all digital products and services that are required and desired by the customers.

Fig. 2. Definition of subscription business models in the machinery and equipment industry.

4 Subscription Offers for Customers

According to this definition, subscription providers need to develop customer offers with which the customers' problems and needs are met. In order to do so, M&E manufacturers can choose from a wide variety of potential offers to provide. These include both conventional offers, e.g. maintenance support, and modern offers by means of digitalization, e.g. analytics platforms. The challenge for the manufacturers is a missing overview and structure for the variety of potential offers. Fig. 3 shows the structure proposed by this work. The most significant offers that were identified in the case study are described for each category of the structure. They include traditional and subscription offers within as well as subscription offers outside of the M&E industry which were found transferable.

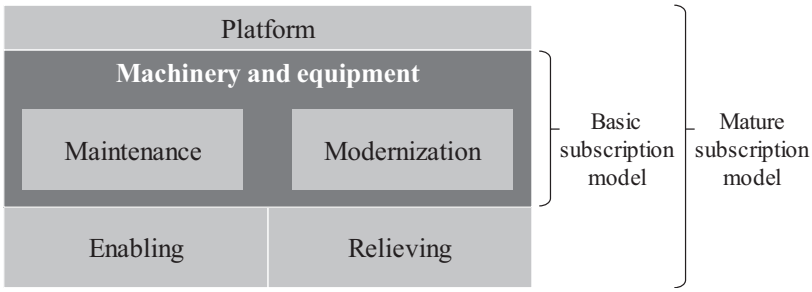


Fig. 3. Structure for subscription-based offers in the machinery and equipment industry.

Within this structure, all potential subscription offers are divided into six categories. Even though the paradigm *customer centricity* suggests that the productivity of machineries and equipment are more important to the customers than the possession of these, the physical products are still the foundation of any subscription model. The majority of all offers in the other categories are designed to increase their productivity. Therefore, excellent machine qualities towards a high productivity are still an important source of competitive advantage and thus indispensable. *Maintenance* is also considered indispensable not only for subscription models but for any business model. It achieves a constant level of productivity which is inherent to the machine and equipment itself, but does not contribute to a continuous productivity increase. In contrast, *modernization* of physical components, digital solutions and services allows a gradual productivity increase over time. Therefore, the combination of these three categories is required in order to form a minimum set of subscription offers in order to achieve the five core value propositions.

The level of fulfillment for the core value propositions can be enhanced by adding further subscription offers. With *enabling* offers, customers receive the required support, tools and knowledge in order to increase the overall performance by themselves. With *relieving* offers, subscription providers disburden their customers of non-core processes so the customers can focus on their core processes. In addition to these five categories, *platforms* are already commonly provided by many M&E manufacturers. It seems obvious to include them into subscription models as well. They can both include additional offers and support offers of the other categories, and are often designed modularly in order to reduce the customization effort. Platforms are essentially important to the economic success of providers. On the one hand, providers can reduce their operational costs by enabling customer self-services. On the other hand, platforms can significantly increase rate of upselling (increase the utilization of a product category), cross-selling (sale of additional product categories) and renewal of subscription contracts [7]. Together, all six categories form a “mature subscription model” which provides more opportunities to increase the customers’ performance and productivity. This way, providers can establish a strong customer relationship that is difficult for competitors to substitute.

An overview of some of the most significant offers that were identified in the case study are shown in Fig. 4. With the proposed structure, companies may identify specific offers with which their customers' problems and needs may be met. Providers may also identify offers which are desired but cannot be delivered today due to missing capabilities. Therefore, the structure may also serve as a basis for the company's future development of capabilities. However, it is worth mentioning that this collection only represents a snapshot of technologically available offers. With progressing technologies, further subscription offers may be developed in the M&E industry.

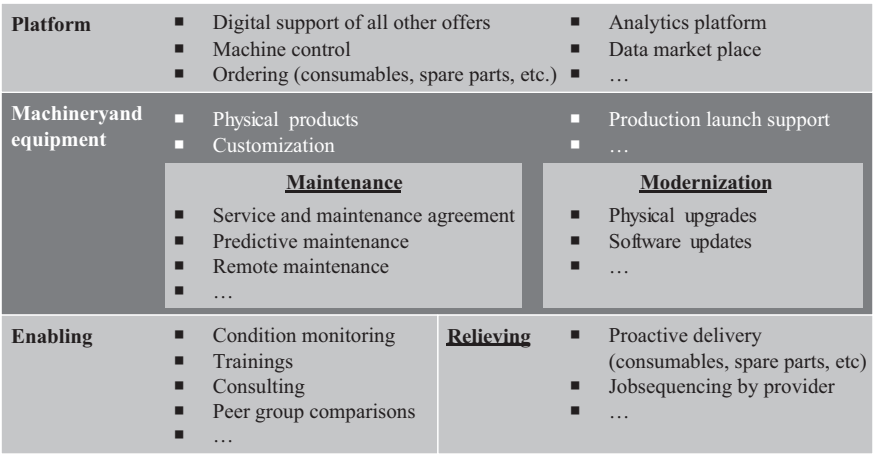


Fig. 4. Overview of most significant subscriptoin-based offers in the M&E industry.

5 Success Factors and Requirements

The value propositions and potential offers were identified from the customers' point of view. However, successful subscription models differ from others also from the providers' point of view. While success factors can increase the effect of the offers, necessary requirements would enable the offers in the first place.

5.1 Success Factors

Success factors can increase the level of fulfillment for the core value propositions. For subscription models in the M&E industry, seven success factors were identified.

Continuous Innovation As the core value proposition *continuous productivity increase* implies, providers need to generate innovations on a regular basis. Apart from competences that are required, providers need to establish organizational capabilities which support the internal development of innovations and incorporate external innovations systematically [18, 19].

Customer Integration Customers should be actively integrated into the providers' innovation processes. The goal is to identify both their explicit and implicit needs and to transfer them into product and service specifications. This way, a high level of fulfillment for the customers' needs can be achieved. However, not all customers are suitable for integration. They need to be carefully selected [7].

Customer Analysis Another way to identify implicit customer requirements are data-based analyses of the customers' interaction with the machines and of the machine performance. The derived insights can be used both to develop more targeted, customized solutions and to channel the development of general future machine generations.

Transparency M&E manufacturers establish a close partnership with their customers by focusing on solving their problems. In this partnership, a high level of trust is essential. This can be fostered by keeping all relevant information and decisions transparent, e.g. real-time machine data and reasons for maintenances. In subscription models which include outcome-based pricing mechanism, transparent billing structures are especially important.

Automated Processes By solving the customers' problems individually, the efficiency potentials of standardized offers and processes are partly eliminated. In order to keep a reasonable cost structure, a high level of automation for the provider's core and support processes by digital means is required. This may include for example processes in customer analysis and accounting, but also sales processes via the platform.

Value-Based Pricing The identification of a customer's willingness to pay is a major challenge for subscription providers. The case study shows that successful providers achieve higher profitability by following a value-based instead of cost-based approach in designing their pricing strategy. However, this change of mentality seems to be especially difficult for the M&E industry.

Customer Success Management Following the developed definition of subscription models, the provider's success highly depends on the customer's success. A constant evaluation of the customer's performance and its proactive support is required by the provider. This may even result in the formation of a specialized organization unit [7].

5.2 Requirements

Based on the previous results, four technical and four organizational requirements that subscription providers in the M&E industry need to fulfill could also be identified.

Modularity of Offers The continuous delivery of customized offers may cause excessive costs for providers and threaten the profitability of the overall business model. In order to maintain an economic cost structure, both physical and digital products as well as service offers need to be designed modularly. Nevertheless, the variance of the modules should be kept at a manageable range.

Real-Time Connectivity For some subscription offers, real-time capability is required, e.g. for condition monitoring, job sequencing by provider and interactions via the platform. Therefore, the machines and equipment need to be able to gather, process and transmit the relevant data in real-time.

Data Sharing Agreement Since the data can reveal the customers' performance and knowhow, some customers may be reluctant to share them. Therefore, successful subscription providers require a dedicated data sharing agreement with the customers. It forms the foundation for a trustful relationship.

Internet of Production The identification of improvement potentials may require customer analyses combining data from different domains. In order to do so, the concept of the Internet of Production, developed by the Laboratory for Machine Tools and Production engineering (WZL) of RWTH Aachen University, supports providers in developing an effective and efficient data infrastructure.

Digital Dynamic Capabilities The transformation to subscription models requires digital and data analytics competences which are sometimes scarce at M&E manufacturers. The acquisition or buildup of these competences may be a long-term task. At the same time, companies need to combine the digital capabilities with organizational dynamic capabilities in order to constantly generate new digital offers [20].

Specialized Sales Forces Traditional sales forces are experienced in selling one-off transaction contracts. Existing subscription providers have experienced great difficulties in training them for selling subscription offers. This may be due to the importance of value-based pricing instead of the traditional cost-based approach. Therefore, many subscription providers found it easier to hire subscription specialized sales agents. Furthermore, the evaluation system for the sales teams also needs to be adapted.

Adapted Organization Not only the sales forces, but the entire organization needs to be adapted to subscription business. The aforementioned definition of subscription models implies a process-oriented organizational structure which seems to be rare in the M&E industry. At the same time, the organization needs to be able to foster a customer-centered culture for leveraging the organizational dynamic capabilities.

Financial Resources Providers need to be able to afford the transition to subscription models financially. Subscription revenues are generated in regular small transactions. In comparison to traditional financial reports, revenues and profits will most likely experience a significant drop during the transition. At the same time, transition costs, e.g. new IT infrastructures and trainings, will further weigh down the financial results. In literature, this effect is often described in the so-called fish diagram [4, 6].

An aggregated overview of this work's results is shown in Fig. 5.

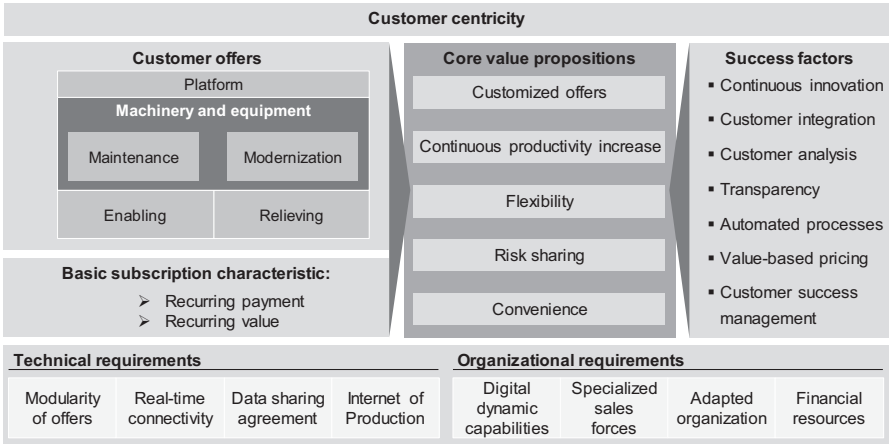


Fig. 5. Overview of aggregated results of the presented work.

6 Outlook

The presented work provides a guide for M&E manufacturers in designing their own tailored subscription models. It includes an adapted definition for subscription models in the M&E industry, an overview and structure for potential offers for customers, as well as success factors and requirements for manufacturers. The results were validated with four M&E manufacturers who have already introduced subscription models or are planning to do so. These manufacturers are well established companies for industrial electric tools, industrial printing machines, high-end medical devices and air compressor systems. The results were found generally valid and applicable. The structure enables M&E manufacturers to define a specific set of value propositions and corresponding offers which suit the company's strategy. By doing so, companies can design tailored subscription models. However, it was agreed that the results, especially the potential offers, depend on available technologies and may change over time. Furthermore, the significance of data analytics for both development and provision of the offers was found essentially important.

The results of this work reveal the necessity for further research in order to support M&E manufacturers in introducing subscription models successfully. While most of the aforementioned technical requirements can already be met with existing technologies, the organizational requirements still lack scientific insights. Researchers on the RWTH Aachen Campus are currently working on bridging some of these scientific gaps. At the same time, subscription models have already attracted companies from adjacent industries to enter the M&E industry. Component suppliers, e.g. for sensors, as well as service companies start to extend their businesses by introducing subscription offers in the categories maintenance, enabling, relieving and platform. Therefore, it is important for M&E manufacturers to take leadership on subscription models and defend their market positions against existing and upcoming competitors.

Acknowledgements. The authors thank the Deutsche Forschungsgemeinschaft (DFG, German research Foundation) for funding the research under Germany's Excellence Strategy – EXC-2023 Internet of Production – 390621612.

References

1. VDMA: Maschinenbau in Zahl und Bild 2019. Frankfurt a. M. (2019)
2. Dutta, S., Lanvin, B., Wunsch-Vincent, S.: The Global Innovation Index 2019. Cornell University; INSEAD; World Intellectual Property Organization, Geneva (2019)
3. Becker, W., Eierle, B., Fliaster, A., Ivens, B., Leischnig, A., Pflaum, A., Sucky, E. (eds.): Geschäftsmodelle in der digitalen Welt. Springer Fachmedien, Wiesbaden (2019)
4. Tzuo, T., Weisert, G.: Subscribed. Portfolio/Penguin, New York (2018)
5. Stich, V., Hicking, J.: Smartifizierung von Maschinenbauprodukten mittels einer zielorientierten Methode. HMD Praxis der Wirtschaftsinformatik **56**(3), 542–556 (2019)
6. Schuh, G., Frank, J., Jussen, P., Rix, C., Harland, T.: Monetizing Industry 4.0. In: Proceedings 2019 IEEE International Conference on Engineering, Technology and Innovation (ICE/ITMC), pp. 1–9, IEEE, Piscataway (2019)
7. Lah, T., Wood, J.B.: Technology-as-a-Service Playbook. Point B Inc, s.l. (2016)
8. VDMA: 4. Ordnungspolitisches Kolloquium des VDMA und der IMPULS-Stiftung. Frankfurt a. M. (2017)
9. Statista: Umsatz je Beschäftigten im Maschinenbau in Deutschland. https://de.statista.com/statistik/daten/studie/235375/umfrage/umsatz-je-beschaeftigten-im-deutschen-maschinenbau/. Accessed 8 Apr 2020
10. Gassmann, O., Frankenberger, K., Csik, M.: Geschäftsmodelle entwickeln. Hanser, München (2013)
11. Osterwalder, A., Pigneur, Y.: Business model generation. Flash Reproductions, Toronto (2010)
12. Chen, T., Fenyo, K., Yang, S., Zhang, J.: Thinking inside the subscription box: New research on e-commerce consumers. McKinsey (2018)
13. Zuora: The end of ownership: How generations across the U.S. really feel about the subscriptino economy. https://www.zuora.com. Accessed 18 Apr 2020
14. Rudolph, T., Bischof, S.F., Böttger, T.M., Weiler, N.: Disruption at the door: a taxonomy on subscription models in retailing. Market. Rev. St. Gallen **5**, 18–25 (2017)
15. Kindström, D., Kowalkowski, C.: Service-driven business model innovation. In: Foss, N. J., Saebi, T. (eds.): Business Model Innovation and the Organizational Dimension, pp. 191–216, Oxford Univ. Press, Oxford (2015)
16. Hammer, M.: What is Business Process Management? In: Vom Brocke, J., Rosemann, M. (eds.) Handbook on Business Process Management 1, pp. 3–16, Springer, Berlin (2015)
17. Wang, C.L., Ye, L.R., Zhang, Y., Nguyen, D.-D.: Subscription to fee-based online services: What makes consumer pay for online content? J. Electron. Commer. Res. **6**(4), 304–311 (2005)
18. Teece, D.J.: Business models and dynamic capabilities. Long Range Plan. **51**(1), 40–49 (2018)
19. Foss, N.J., Saebi, T. (eds.) Business Model Innovation and the Organizational Dimension, 1. ed. Oxford Univ. Press (2015)
20. Warner, K.S.R., Wäger, M.: Building dynamic capabilities for digital transformation: An ongoing process of strategic renewal. Long Range Plan. **52**(3), 326–349 (2019)

Author Index

B.-A. Behrens et al. (Eds.): WGP 2020, LNPE, pp. 729–732, 2021.
https://doi.org/10.1007/978-3-662-62138-7