

# Striving for Zero Defect Production: Intelligent Manufacturing Control Through Data Mining in Continuous Rolling Mill Processes

Benedikt Konrad, Daniel Lieber and Jochen Deuse

**Abstract** Steel production processes are renowned for being energy and material demanding. Moreover, due to organizational and technological restrictions in flow production processes, the intermediate product's internal quality features cannot be assessed within the process chain. This lack of knowledge causes waste of energy and material resources, unnecessary machine wear as well as reworking and rejection costs, when defective products are passed through the entire process chain without being labeled defective. The process control approach presented in this paper provides the opportunity of gaining transparency on quality properties of intermediate products. This aim is achieved by predicting intermediate product's quality by means of data mining techniques. This approach can be applied in a wide field of production environments, ranging from steel and rolling mills to automated assembly operations. Concerning this concept, the authors derive a methodology for representing different quality properties in a way that it can be applied in the process control. Beyond that, first results of statistical analyses on the quality-related significance of process parameters are disclosed.

**Keywords** Process data mining · Real-time quality prediction · Intelligent manufacturing process control

---

B. Konrad (✉) · D. Lieber · J. Deuse  
Chair of Industrial Engineering, TU Dortmund University, Dortmund, Germany  
e-mail: benedikt.konrad@tu-dortmund.de

D. Lieber  
e-mail: daniel.lieber@tu-dortmund.de

J. Deuse  
e-mail: jochen.deuse@tu-dortmund.de

# 1 Introduction

Resource and energy efficiency of interlinked manufacturing processes can be increased in many ways: process engineering solutions or design improvements, higher quality of raw materials as well as optimized operating parameters [1]. In this context, based on a rolling mill case study, real-time optimization and adaptation of predefined operating modes on the basis of continuous inline quality control by means of data mining techniques poses a promising approach to sustainably increase the efficiency of production processes [2, 3].

Continuous quality control across all stations of the process chain is intended to ensure processing conditions within certain tolerance limits in order to guarantee the final product's quality standards defined by the customer. Nevertheless, up to now, processing by rigid and inflexible program guidelines is still common practice in continuous flow production systems. In case of missing quality gates due to organizational and technological restrictions, processing steps are basically targeted to pre-fixed operation settings instead of orienting and adjusting on the basis of actual product quality levels or states of processing.

Based on the Toyota Production System, the development of a new production control approach is motivated by the Jidoka principle, which focuses on avoiding waste, such as rejection and rework, as early as possible within production processes [4]. Although these process-immanent quality checks are state of the art in the automotive industry, they cannot be transferred to high complex and interlinked production systems in the steel industry one-to-one [5].

The rigidly linked sequence of production steps as well as rough environmental conditions during processing and technological restrictions within the value chain obstruct assessing the physical quality of intermediate products [6]. For that reason, steel production is still characterized by final quality diagnostics measurements that are done at the finishing stands at the end of the process chain [7]. This is a major drawback as failures during production lead to high internal costs and waste of resources when only the final product's quality properties can be tested [6].

Therefore, this paper is focused on avoiding inferior product quality by real-time Inline Quality Prediction (IQP) tools based on data mining and artificial intelligence and its integration into a comprehensive Intelligent Manufacturing Process Control (IMPC) approach for industrial application.

The considered case study is provided by a leading German steel producer and is representative for the production control approach discussed in this paper. The process chain consists of five major processing steps starting with the heating process at the rotary hearth furnace where steel bars are heated to forming temperatures. In the following facilities the bar's profile is reduced to customer specifications before the entire bar is separated into rods of customer defined length (see Fig. 1). A more detailed description of the production steps as well as preliminary work on processing and storing the collected static, process and quality data can be found in [6, 8].

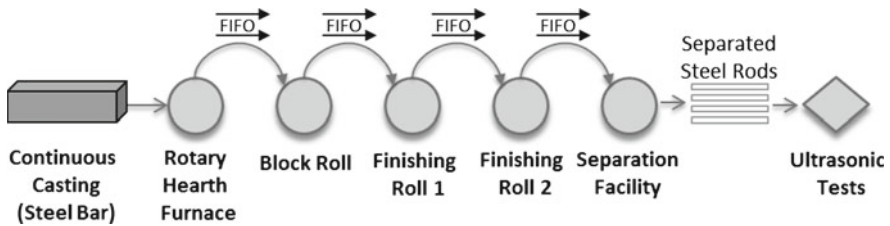


Fig. 1 Process chain of rolling mill case study [6]

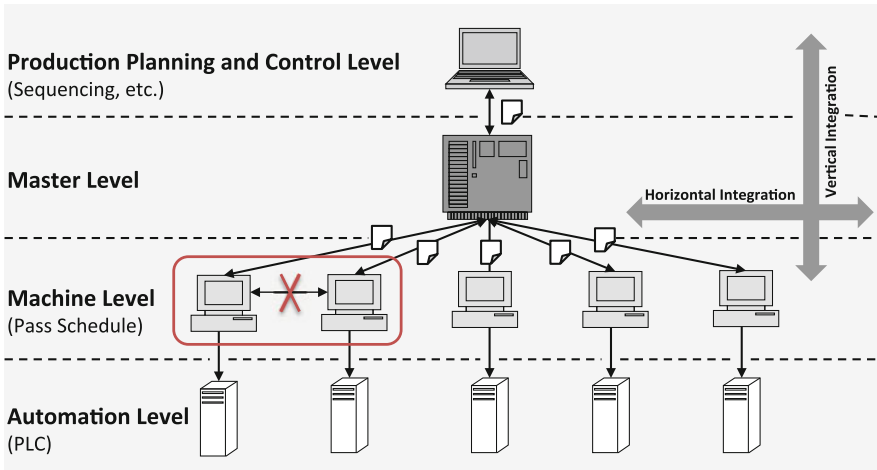
The remainder of this paper is arranged as follows: Sect. 2 summarizes the IT-infrastructure and schemes applied for production control in the case study. Section 3 introduces the Intelligent Manufacturing Process Control concept as an answer to the previously identified drawbacks. Sections 4 and 5 focus on deriving quality criteria and statistically analyzing the impact of various process parameters on quality properties on the basis of data collected at the first step of the rolling mill process chain, the rotary hearth furnace. Finally, in Sect. 6 a conclusion is given and the next steps in developing the process control concept are lined out.

## 2 Manufacturing Process Control Tools in Flow Production

### 2.1 Manufacturing Process Control IT-Infrastructure in a Rolling Mill Case Study

IT infrastructures—in the case study considered—are characterized by a hierarchical organization of all required process planning, control and execution entities. Figure 2 depicts information flows and communication interfaces of the process control tools within the IT-structure of the case study.

On the top level, the production planning and control level, customer order specifications such as material, order volume and final dimensions are transformed to unique internal production order specifications including a predetermined work flow sequence as well as a specific assignment of predefined rolling schedules dependent on input and output dimensions of the material. This basic production guideline is transmitted to the master level, in which a centralized master computer coordinates and traces all individual intermediate products of the corresponding production order within the value chain by their unique identification code. The master level spreads all relevant processing information to the machine level, so that the operator is informed of current and upcoming tasks. Machine-to-Machine communication is not provided. After confirming the incoming production telegram, the corresponding rolling program is loaded and automatically passed to the automation level, where each processing step of the predefined rolling schedule is executed, supervised and controlled by a Programmable Logic Controller (PLC).



**Fig. 2** Information flows across function levels

Nevertheless, the operator is able to intervene and adapt rolling parameters manually at any time according to unforeseen interruptions or failures during processing. After finishing the intermediate product's processing steps, the operator confirms the finished production telegram which is then automatically fed back to the master level and transferred with new input parameters to the next production step.

## ***2.2 Structural Weaknesses of Process Control Systems in Flow Production***

This type of manufacturing process control strategy can be described as a non-adaptive automated production control, which is based exclusively on predefined pass schedules in a one-way feed-forward approach. Despite of manual interventions by the operator due to obvious abnormalities from regular production, there is no additional supporting entity integrated that can monitor and verify the status of processing in real-time or even derives adequate counteractions in case of deviations to guarantee quality specifications of the final product. In fact, no feed-back loop which, for instance, compares target and actual processing parameters regarding the intermediate product's quality properties is provided.

To react on quality deviations of intermediate products in real-time, the automated process control has to be enhanced with intelligent process control modules based on artificial intelligence which allow real-time optimization and adaptation of pass schedule parameters to face challenges such as resource and energy efficient production processes [2, 9]. Therefore, an Intelligent Manufacturing Process Control applicable in hot-rolling production processes in the steel industry has to be developed.

### 3 Intelligent Manufacturing Process Control in Flow Production Systems

As lined out in the previous chapters, a new control scheme has to be developed. With the intention of implementing the Jidoka principle in steel industry, this approach has to incorporate features such as process data acquisition and interpretation besides the fundamental aspects of order control and tracking. The general idea of continuously monitoring and adjusting process parameters is known as Advanced Process Control in the process industry [10, 11]. The Intelligent Manufacturing Process Control (IMPC) introduced here transfers this basic idea from process industry to other industry branches by introducing data mining techniques to predict the quality of an intermediate product and adjust further processing steps according to the prediction.

The data mining process on which this approach is based is the knowledge discovery in databases (KDD) process [12]. In this context data mining is defined as the “application of data analysis and discovery algorithms that produce a particular number of patterns (or models) over the data” [13]. The KDD process is divided in nine steps [12]:

1. Developing an understanding of the application domain
2. Creating a target data set, selecting a data set
3. Data Cleaning and preprocessing
4. Data reduction and transformation
5. Choosing the data mining task
6. Choosing the data mining algorithms
7. Data mining
8. Evaluating output of step 7
9. Consolidating discovered knowledge.

The IMPC concept presented in this section follows these steps. Steps 1 and 2 were executed while developing the IMPC, whereas the remaining steps 3 to 9 are automated and can be found in the modules described below.

The developed IMPC approach consists of different functional modules for distinct purposes. The core modules are (1) Data Acquisition and Storage (DAS), (2) Data Monitoring (DM), (3) Inline Quality Prediction (IQP) and (4) Process Parameter Optimization (PPO). Modules (2–4) represent the different process control stages and form the IMPC. While DAS is a prerequisite for implementing the IMPC, the IMPC itself can be realized to two different extents. Firstly, including the PPO-module, which allows an online optimization of process parameters depending on the current state of the production process in order to compensate previous process deviations. And secondly, just including the DM and IQP-module in terms of a decision support system. This type of process control relies on generating real-time quality estimates for the intermediate product. If inferior quality is predicted, the production process is halted. In contrast to the previous option, the decision on further actions to be taken to improve the products quality is not made

by the IMPC, but relies on expert knowledge. This modular design of the entire process control approach allows an on demand integration in existing process control infrastructures. Due to the early stage of the research project, the paper's scope is limited to this approach and therefore to the first three modules.

### ***3.1 Data Acquisition and Storage***

The extent to which the DAS module has to be introduced into a production system is determined by the complexity of production that has to be supervised. Obviously, complexity in this context reflects the number of distinct influencing parameters. These can range from environmental factors, e.g., temperatures, humidity, noise or vibrations caused by other machinery, to machine states like applied torque, engine speed or execution time for a predefined task. This short overview illustrates that a general concept for DAS cannot be given, as it rather depends on the actual process the entire IMPC concept is applied to.

In case of the rolling mill case study, processing parameters as well as machine states of all major processing stations are relevant. Examples for these are zone temperatures, material temperatures and time spent in each zone at the rotary hearth furnace, or rolling speed and forces, material temperatures at each forming step and grooves used at each of the rolls. A comprehensive description of relevant factors as well as information on the actual implementation of the DAS is given in [5, 6].

### ***3.2 Intelligent Manufacturing Process Control***

The IMPC concept as implemented in the case study can be interpreted as a separate building block in a company's process control landscape (see Fig. 3). The IMPC does not affect the process of production planning, i.e., it generates no advice on optimized production sequences. Instead, the focus is on analyzing processing states in real-time aiming at forecasting the intermediate product's quality properties. This knowledge is used for either deriving recommendations on whether the product should be processed any further or for optimizing the next processing step's parameters so that the required product quality can be obtained.

In both cases the IMPC requires information on the current processing parameters which are gained from the DAS module. Moreover, information on historic processing data, merged in a quality prediction model has to be available for assessing the intermediate product's quality correctly. When it comes to optimizing processing parameters, even more knowledge is demanded. In this case it is mandatory to have a material forming simulation, i.e., a finite element method (FEM) simulation, at hand which verifies the impact of altered processing parameters on the product. Figure 4 summarizes the required knowledge and infoData Monitoring and Inline Quality Prediction. Based on the real-time

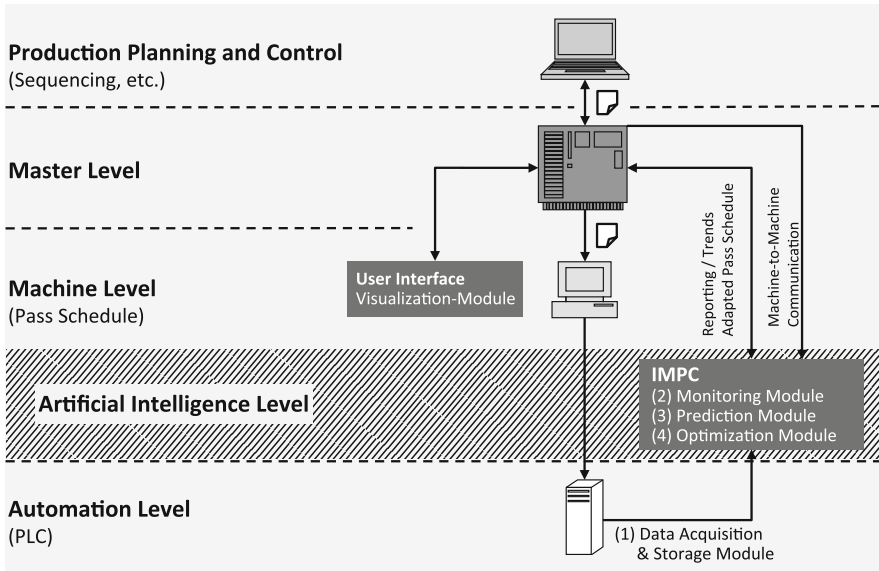


Fig. 3 Integration of artificial intelligence in operational process control

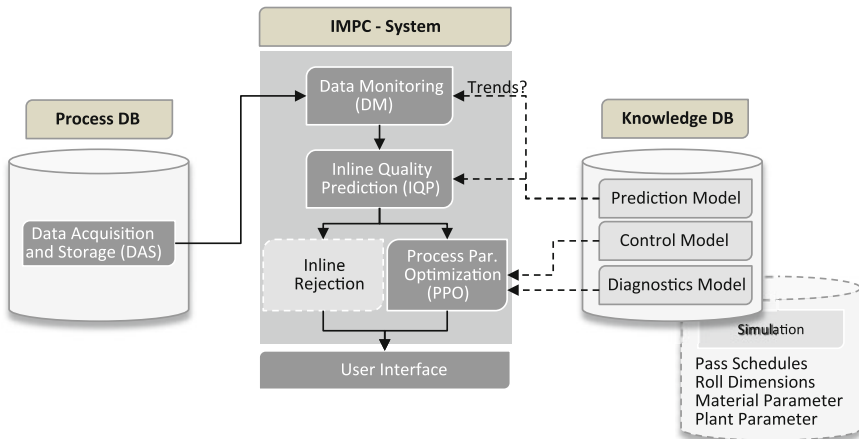


Fig. 4 Intelligent manufacturing process control model (IMPC)

processing of data available from the DAS module, the DM module generates all statistics and information demanded by the IQP module. These tasks also include all necessary pre-processing steps on the data gathered: e.g., depending on the type of data missing values are replaced either by a default value or average values in the case of time-series data. Additionally, for time series data statistics such as maximum and minimum values, standard deviation and the maximum gradient are computed. These characteristics are used for a first evaluation of the current

processing step before an in-depth analysis is conducted. All results in the DM module are then fed into the IQP module.

The IQP module translates all available information irrespective of its actual meaning into a quality assessment. This is done by means of data mining, i.e., supervised learning models [5, 14]. These models are trained on labeled historic datasets so that these algorithms can judge on the currently processed quality based on actual processing parameters. Additionally, the results of the final quality checks at the process chain's end are merged into the continuously growing database. By this, the prediction accuracy of the data mining models can be increased continuously.

The result of the IQP module can then be applied to decision rules, specifying whether an intermediate product is ejected from the process or processed further depending on the quality prediction and its position in the process chain. In a second step, the IMPC can be extended by the implementation of the PPO module, which also makes recourse on the quality estimates generated in the IQP module.

**Process Parameter Optimization.** The PPO module is the final step in implementing the IMPC model. As described above, it aims at adjusting parameters of upcoming process steps in such a way that identified quality deviations caused by previous processing steps are compensated in the remainder of the process chain. To accomplish this, the PPO module needs comprehensive access to information on process parameters and especially on FEM simulation results to validate the impact of altered parameters on product quality properties. This module is obviously the most challenging one in the IMPC concept. It relies on completely implemented DM and IQP modules. Due to the early state of the research project, the PPO module is still a theoretical concept and not yet defined on a detailed level. Consequently, the remainder of the paper focusses on the IMPC model consisting of the DM and IQP modules.

## 4 Quality-Levels and Decision-Rules in Rolling Mills

### 4.1 *Quality-Levels for Intermediate Products*

The most important prerequisite for the IMPC concept is the definition of quality in the context of intermediate products. Final products, i.e., steel rods in the case of the rolling mill case-study, are assessed according to various distinct quality parameters. The intermediate product's quality, i.e., the quality of the steel bar, which is predicted by the IMPC is not assessed physically at any point in the process chain. In this special case the intermediate product's quality cannot be related to the final product's quality one by one as each intermediate product (each steel bar) results in a variable number of steel rods depending on the profile specified in the customer's order. Hence, the intermediate product's quality has to be composed of the quality of the different final products.



Due to these facts, intermediate product quality is determined according to following sum function to aggregate the final products' quality properties to one embracing quality label for the training of the quality prediction models:

$$Q_b = 1 - \frac{1}{R_b} \sum_{r=1}^{R_b} \sum_{p=1}^P \left( w_p \cdot \frac{\lambda_{p,r,b}}{\lambda_{p,max}} \right); \forall b \in B, \text{ where} \quad (1)$$

$Q_b \in [0, 1]$ , quality level of bar  $b$

$w_p \in [0, 1]$ ,  $\sum_{p=1}^P w_p = 1$ , weight of quality property  $p$

$\lambda_{p,r,b} \in [1, \lambda_{p,max}]$ , value of quality property  $p$  of rod  $r$  in bar  $b$

$b \in \{1, \dots, B\}$ , steel bar  $b$  in set of allbars  $B$

$r \in \{1, \dots, R_b\}$ , steel rod  $r$  in set of all rods resulting from bar  $b$

$p \in \{1, \dots, P\}$ , quality property  $p$  in set of all properties

$B, R_b, P \in \mathbb{N}$

$Q_b$  represents the quality level of a certain steel bar (see Formula 1). It does not represent the actual physical quality according to some specified criterion such as "type of error", but allows an aggregated estimate on the overall quality of each bar. The quality level is composed of various different quality properties ( $\lambda_{p,r,b}$ ) that are normalized on a 0–1 interval by dividing each value by its maximum value ( $\lambda_{p,max}$ ). The quality level represents the weighted average of these distinct quality properties.

Thus, a quality level of 1 implies that the current bar does not show any quality deviations at all. On the other hand, a quality level of 0 emphasizes that all rods from this bar obtained the maximum error value. Hence, all values between zero and one have to be interpreted more carefully. The aggregation of various quality properties from different products to one single label leads to an information loss so that it cannot be judged whether one single rod has a quality level of almost zero or all rods show a slightly minor quality.

When it comes to predicting the intermediate product's quality, this loss of information can be neglected, because the resulting decision in the IMPC on whether the steel bar should be processed further or ejected is the same in both cases: assuming that one rod out of a bar is defective it is advantageous to completely process the bar and to discard the defective bar. On the other hand, if all rods have a slightly minor quality the customer might accept the rods despite of the quality deviations or the rods can be assigned to a different customer order which requires lower quality properties.

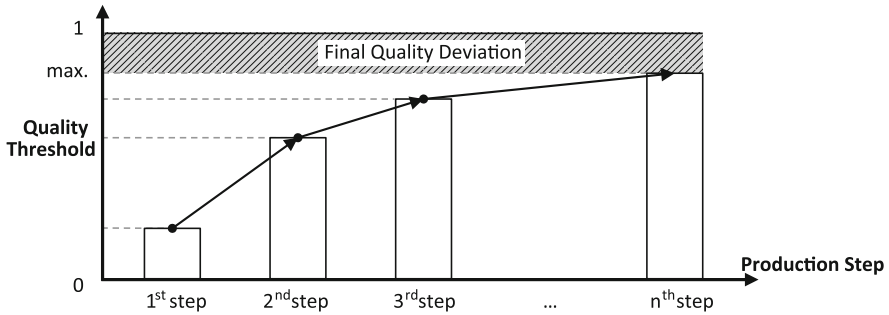


Fig. 5 Required quality thresholds at different production steps

## 4.2 Application of Quality-Levels in the IMPC Concept

The quality level of a certain intermediate product is assessed by the IMPC at each processing step. In order to meet customer requirements, the quality level has to meet the defined target at the end of the process chain. Moreover, quality assessments are becoming more precise, with increasing processing information on the intermediate product. Consequently, the tolerated quality deviations at the end of the process chain will be the smallest.

The IMPC accounts for these characteristics, as it applies a threshold method for deriving recommendations for the decision on ejecting the intermediate product or not. The threshold at the beginning of the process chain is rather low as on the one hand quality level predictions might be somewhat fuzzy and on the other hand the quality level can be influenced positively to a significant degree in the remainder of the processes. Thus, the threshold increases with the number of processing steps completed. At the last process step the difference between threshold and  $Q_b = 1$  corresponds to quality deviations accepted by the customer and the remaining fuzziness of the quality prediction model (see Fig. 5).

The quality level thresholds at all processes have to be determined by means of expert knowledge as well as data mining and multivariate statistics. Analyzing the quality related significance of each process's parameters reveals its impact on the quality level of the final product. The following chapter presents the analysis conducted on the rotary hearth furnace's process parameters.

## 5 Fundamental Process Parameter Studies and First Results

This chapter focusses on statistical analyses conducted on the effects of process parameters of the rotary hearth furnace. The analyses are performed on a sample of about 800 steel bars of one certain material and size.

**Table 1** Correlation coefficients of quality level and characteristic parameters

	Min.	Max.	Max.	Std. Dev.	Max. Grad.
<b>Furnace temperature</b>					
Qual. level	-0.045	0.065	0.032	-0.099	-0.093
Significance	0.106	0.037	0.189	0.003	0.005
<b>Temperature at top of steel bar</b>					
Qual. Level	-0.094	0.025	0.081	-0.146	-0.115
Significance	0.005	0.247	0.012	0.000	0.001
<b>Temperature at top in core bar</b>					
Qual. Level	-0.048	0.069	0.103	-0.124	-0.119
Significance	0.094	0.027	0.002	0.000	0.000
<b>Temperature at bottom of steel bar</b>					
Qual. Level	-0.093	0.067	0.101	-0.145	-0.119
Significance	0.005	0.031	0.003	0.000	0.000

### ***5.1 Correlation Analyses on Process Parameters of Rotary Hearth Furnace***

The DAS module at the furnace gathers information on four parameters: the temperature in the furnace and the temperature at the bottom, top and core of each steel bar. Each of these parameters represents time series data, as they are collected with a frequency of up to 100 Hz [8], while the bar is in the furnace. For analyzing the effect and significance of these factors regarding the quality level, the time series data is preprocessed in order to determine characteristics that describe each of the four value series. For each time series of each steel bar passing the furnace the maximum, minimum and average temperatures are computed as well as standard deviation and largest gradient.

Analyzing the correlation coefficients (see Table 1) shows that the standard deviation and the maximum gradient have the biggest impact on the quality level and influence it negatively in all analyses. This result shows that not the temperatures themselves but their variance and the speed of temperature changes during heating are the most important determinants of quality levels. Process parameters having a significance level  $\leq 0.05$  have a statistically significant effect on the quality level.

### ***5.2 Regression Analyses on Process Parameters of Rotary Hearth Furnace***

On the sample set regression analyses are conducted in order to judge on the size of each parameter's impact and its significance for the prediction of quality levels. As the residuals of the linear regression are not distributed normally a logistic regression is carried out on the sample set [15, 16]. The required binning of the dependent variable, i.e., the quality level, is performed according to expert knowledge leading to two classes

**Table 2** Regression coefficients and significances from logistic regression

	Regression coefficients	Significance
<b>Furnace temperature</b>		
Min.	-0.006	0.052
Max. Grad.	0.039	0.367
Avg.	-0.005	0.173
Std. Dev.	-0.023	0.008
Max. Grad.	0.291	0.365
Constant	-34.972	0.536
<b>Temperature at top of steel bar</b>		
Min.	-0.012	0.000
Max. Grad.	-0.090	0.154
Avg.	0.002	0.438
Std. Dev.	-0.030	0.000
Max. Grad.	2.324	0.018
Constant	118.763	0.116
<b>Temperature in core of steel bar</b>		
Min.	-0.007	0.104
Max. Grad.	-0.053	0.039
Avg.	0.011	0.000
Std. Dev.	-0.012	0.183
Max. Grad.	-0.076	0.940
Constant	60.616	0.025
<b>Temperature at bottom of steel bar</b>		
Min.	-0.012	0.000
Max. Grad.	-0.049	0.195
Avg.	0.009	0.001
Std. Dev.	-0.028	0.004
Max Grad.	1.872	0.151
Constant	61.134	0.149

of quality levels, those in  $[0; 0.75]$  labeled 0 for insufficient quality and those in  $[0.75; 1]$  labeled 1 for good quality. Instead of a comprehensive regression on all parameters, four separate analyses are performed as the process parameters analyzed are interdependent. The results of the regression models are shown in Table 2. The significances state, that the standard deviation is the major determinant of the quality level in most cases. Moreover, it has the strongest influence on quality levels of all significant parameters. That the steel bar's core is the only case in which the standard deviation of temperatures has no significant effect on the quality level can be explained with the heating process itself. The temperatures at the core are increasing markedly slower and more steadily, as the surrounding material attenuates the temperature differences. The same reasoning can be applied for the maximum gradient's values. Instead of the standard deviation the average temperature influences the quality most in this scenario.

The results for the minimum temperatures are the most striking ones in this analysis. The regression coefficients show that lower minimum temperatures, i.e., starting temperatures in the heating process, add to final quality. This effect is marginal yet

mostly significant. This result can be explained by the sample for minimum temperatures which is not representative as it is skewed towards lower minimal temperatures. Actually it is expected, that higher minimum temperatures lead to higher quality. This can also be reasoned using standard deviation, as an increased minimum temperature will lead to a lower standard deviation in a continuous heating process as the difference between minimum and maximum temperatures decreases.

### ***5.3 Conclusion on Statistical Tests***

The correlation analysis shows that especially the standard deviation of temperatures is critical for producing good quality as it is highly significant and besides this has the highest negative correlation coefficients. The same is true for the maximum gradient. The regression analyses conducted confirm the relevance of the standard deviation, whereas the maximum gradient is found to be insignificant. This can be explained by the correlation of these two parameters. The relevance of average bar temperatures at all three locations in the regression model can be proven by the same facts. Developing the learning algorithm for quality predictions these factors with high significances have to be considered in the model.

The results presented of the logistic regression are in fact quite similar to those results generated by the linear regression analysis regarding input parameter significances and relative impact of each parameter. Even though these models include twenty parameters computed from only four separate time series gathered at the very first step in the process chain, the coefficient of determination (Nagelkerkes- $R^2$ ) indicates that the bar's models account for 7–9 % and the furnace's model for 4 % of the overall variability. The remaining 92 % have to be accounted for by models of the remaining processes.

This result directly influences the decision on the quality threshold of the rotary hearth furnace. Knowing all prediction models' coefficients of determination sheds a light on the distribution of quality thresholds. The higher the percentage of variation accounted for at a certain point in the process, the higher this process's quality level threshold has got to be. Thus, the quality threshold at the furnace will be rather low, compared to the remaining ones, which is a result that has been expected by the process experts, who assume the heating process not to be the most critical one. Nevertheless, a smooth heating-up process, avoiding high temperature changes, will have a positive effect on the product's quality.

## **6 Summary and Future Work**

This paper proposes a production control concept based on data mining techniques. This concept offers the chance to reduce waste of energy and material resources as well as reworking and rejection costs resulting from producing, if

inferior quality is not identified in the process. The proposed IMPC concept identifies those products by assessing their quality based on a prediction model. At each processing step a certain quality threshold has to be attained in order to permit further processing. To select quality relevant parameters for prediction, first statistical analyses were conducted on the data gathered at the first process step which show that the standard deviation of the temperature value series data recorded in the furnace and at three points of the steel bar has the most significant effect on quality levels. Moreover, the statistical models account for 7–9 % of variability, which is a satisfactory result due to the models' complexity and coinciding with experts' expectations.

The statistical analyses have to be conducted for all remaining steps in the process chain. Based on the results of these, the relevant parameters for the IQP module of the IMPC concept have to be identified. Thresholds for each process will be derived from both models' coefficients of determination and expert knowledge. Given this information local quality prediction model will be trained for each process. Once this is done, the IMPC can be applied in practice and helps to reduce waste due to processing products of inferior quality.

**Acknowledgments** This work has been supported by the DFG, Collaborative Research Center 876 "Providing Information by Resource-Constrained Data Analysis", project B3 "Data Mining in Sensor Data of Automated Processes". <http://sfb876.tu-dortmund.de>

## References

- Otte, R., Otte, V., Kaiser, V.: Data Mining für die industrielle praxis (Data Mining for Industrial Application). Hanser, Munich (2004)
- Alvarez, E.G.: Advanced process control to meet the needs of the metallurgical industry. World Metall. ERZMETALL **58**(3), 123–128 (2005)
- Morik, K., Bhaduri, K., Kargupta, H.: Introduction in data mining for sustainability. Data Mining and Knowledge Discovery, vol. 24, No. 2, pp. 311–324, Springer (2012)
- Ohno, T.: Toyota Productions System, pp. 6–8. Productivity Press, Portland (1982)
- Morik, K., Deuse, J., Faber, V., Bohnen, F.: Data mining in sensordaten verketteter prozesse (data mining in sensor data of interlinked processes). ZWF **105**(1–2), 106–110 (2010)
- Lieber, D., Konrad, B., Deuse, J., Stolpe, M., Morik, K.: Sustainable interlinked manufacturing processes through real-time quality prediction. In: Leveraging Technology for a Sustainable World: Proceedings of the 19th CIRP Conference on Life Cycle Engineering. Springer, Berkeley (2012) (accepted for publication)
- Haapamäki, J., Tamminen, S., Röning, J.: Data mining methods in hot steel rolling for scale defect prediction. In: International Conference on Artificial Intelligence and Applications, Innsbruck, Austria, pp. 90–94 (2005)
- Stolpe, M., Morik, K., Konrad, B., Lieber, D., Deuse, J.: Challenges for data mining on sensor data of interlinked processes. In: Next Generation Data Mining Summit: Ubiquitous Knowledge Discovery for Energy Management in Smart Grids and Intelligent Machine-to-Machine (M2 M) Telematics, Athens, Greece (2011). Available at: <http://www.kd2u.org/NGDM11>
- Oh, S., Han, J., Cho, H.: Intelligent process control system for quality improvement by data mining in the process industry. In: Braha, D.: Data Mining for Design and Manufacturing, Academic Publishers, Norwell pp. 289–309 (2001)

10. Dittmar, R., Pfeiffer, B.-M.: Modellbasierte prädiktive Regelung (Modell-based Predictive Control), pp. 1–4. Oldenbourg, Munich (2011)
11. Seborg, D.E., Edgar, T.F., Mellichamp, D.A.: Process Dynamics and Control, 2nd edn, pp. 411–414. Wiley, Hoboken (2004)
12. Fayyad, U.M.: Data mining and knowledge discovery: making sense out of data. *IEEE Expert* **11**(5), 20–25 (1996)
13. Fayyad, U.M., Piatetsky-Shapiro, G., Smyth, P.: From data mining to knowledge discovery in databases. *AI Magazine* **17**(3), 37–54 (1996)
14. Stolpe, M., Morik, K.: Learning from label proportions by optimizing cluster model selection. In: Gunopulos, D., Hofmann, T., Malerba, D., Vazirgiannis, M. (eds.) *ECML PKDD 2011, Part III*, vol. 6913, pp. 349–364, Springer, Berlin, Heidelberg (2011)
15. Menard, S.: Applied logistic regression analysis. 2nd edn. Sage University Papers Series on Quantitative Applications in Social Sciences 07–106, Sage, Thousand Oaks (2001)
16. Sethi, I.: Data mining: An introduction. In: Braha, D.: *Data Mining for Design and Manufacturing*, Kluwer Academic Publishers, Norwell pp. 1–40 (2001)