



# From Measurements to Knowledge - Online Quality Monitoring and Smart Manufacturing

Satu Tamminen<sup>1</sup>(✉), Henna Tiensuu<sup>1</sup>, Eija Ferreira<sup>1</sup>, Heli Helaakoski<sup>2</sup>,  
Vesa Kyllönen<sup>2</sup>, Juha Jokisaari<sup>3</sup>, and Esa Puukko<sup>4</sup>

<sup>1</sup> University of Oulu/BISG, Oulu, Finland

[satu.tamminen@oulu.fi](mailto:satu.tamminen@oulu.fi)

<sup>2</sup> VTT, Oulu, Finland

<sup>3</sup> SSAB Europe, Raahe, Finland

<sup>4</sup> Outokumpu Stainless, Tornio, Finland

<http://www.oulu.fi/bisg>

**Abstract.** The purpose of this study was to develop an innovative supervisor system to assist the operators in an industrial manufacturing process to help discover new alternative solutions for improving both the products and the manufacturing process.

This paper presents a solution for integrating different types of statistical modelling methods for a usable industrial application in quality monitoring. The two case studies demonstrating the usability of the tool were selected from a steel industry with different needs for knowledge presentation. The usability of the quality monitoring tool was tested in both case studies, both offline and online.

**Keywords:** Data mining · Smart manufacturing · Online monitoring  
Quality prediction · Knowledge representation · Machine learning

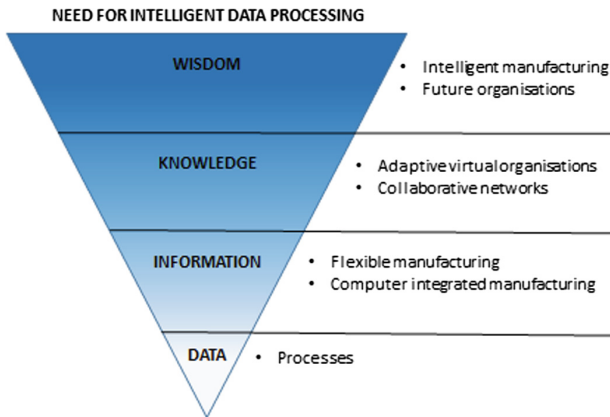
## 1 Introduction

Knowledge can be considered to be the most valuable asset of a manufacturing enterprise, when defining itself in the market and competing with others. The competitiveness of today's industry is built on quality management, delivery reliability and resource efficiency, which are dependent on the effective usage of the data collected from all possible sources. The risk is that the operators with the limited capacity to process the incessant information flow miss the essential knowledge within the data. Recent advances in statistical modelling, machine learning and IT technologies create new opportunities to utilize the industrial data efficiently and to distribute the refined knowledge to end users in right time and convenient format.

Manufacturing has benefited from the field of data mining in several areas, including engineering design, manufacturing systems, decision support systems,

shop floor control and layout, fault detection, quality improvement, maintenance, and customer relationship management [1]. While the amount of data expands rapidly, there is a need for automated and intelligent tools for data mining. Statistical regression and classification methods have been utilized for steel plate monitoring [2]. Decision support systems (DSS), for example, become intelligent when combined with AI tools such as fuzzy logic, case-based reasoning, evolutionary computing, artificial neural networks (ANN), and intelligent agents [3, 4].

Knowledge engineering and data mining have enabled the development of new types of manufacturing systems. Future manufacturing is able to adapt to demands of agile manufacturing, including a rapid response to changing customer requirements, concurrent design and engineering, lower cost of small volume production, outsourcing of supply, distributed manufacturing, just-in-time delivery, real-time planning and scheduling, increased demands for precision and quality, reduced tolerance for errors, in-process measurements and feedback control [5]. Smart manufacturing will bring solutions to existing challenges, but the current industry utilizes generally the information from its environment and in best cases only the first level of knowledge (Fig. 1). The progress in industrial data utilization is enabled with novel intelligent data processing methods.



**Fig. 1.** The evolution of data to knowledge requires novel methods for intelligent data processing that enable the shift to smart manufacturing.

Bi *et al.* state that every major shifting of manufacturing paradigm has been supported by the advancement of IT. Internet of Things (IoT) may change the operation and role of many existing industrial systems in manufacturing. The integration of sensors, RFID tags and communication technologies into the production facilities enables the cooperation and communication between different physical objects and devices [6]. One of the technical challenges in IoT research is the question how to integrate IoT with existing IT systems. When the massive amount of real-time data flow is to be analysed, currently strong big data analytics skills are needed from the end user [7]. However, the employment of

experts concentrate on the core area of the industry, which in its turn, generate a demand for intelligent tools for decision support.

Information presentation is a complex task in manufacturing, as the amount of quality parameters that need to be linked with even a larger number of process parameters is difficult to process with capabilities of a human being. Akram *et al.* show how statistical process control (SPC) and automatic process control (APC) can be integrated for process monitoring and adjustment [8]. Statistical models bring wider possibility to produce information with their capability to predict the future outcome, which enables the process and production planning. The challenge is how to enable the communication between people, how to get information that they need from the process or the product, if the information transfer is enabled between the work posts or manufacturing facilities, or how to provide information about the malfunction or decreased quality of the products. The information should be presented clearly, solutions for the problem, also warnings if automatic corrective actions are enabled. As a whole, the information chain should be supported with a tool that enables the knowledge based conversation within the company.

When product quality improvement is pursued, Kano and Nakagawa suggest that the process monitoring system should have at least the following functions: it should be able to predict product quality from operating conditions, to derive better operating conditions that can improve the product quality, and to detect faults or malfunctions for preventing undesirable operation. They have used soft sensors for quality prediction, optimization for operating conditions improvement, and multivariate statistical process control (MSPC) for fault detection in steel industry application [9]. From these objectives, the derivation of better operating conditions may be the most difficult one to reach; even the definition for better conditions can be challenging to draw, as the conditions are often a compromise of least harmful and cost efficient practices.

In this article, we propose a method for online quality monitoring during a manufacturing process with two application cases in steel industry. Our tool links together the statistical models for prediction of quality properties based on the process settings and variables, and presents the results with easily interpretable visualisations. This paper is organized as follows. Section 2 describes the requirements and specifications for online quality monitoring tool for industrial use. Section 3 presents the choice of the modelling method for quality monitoring purposes. The quality prediction based tools for decision support in two case studies are presented in Sect. 4. Section 5 concludes the quality monitoring development.

## 2 Developing a Quality Monitoring Tool for Industrial Use

### 2.1 The Domain Requirements and Requests

We launched the development of the quality monitoring tool (QMT) by surveying the requirements set by business, end users and IT environment of the

process industry. We arranged a series of workshops with participating companies, development partners and other stakeholders.

When specifying the requirements of the QMT, we selected the key user groups for the first development step. This way, the tool could be demonstrated with users having different needs. The selection had the highest impact on the user interface and information presentation.

Technical specifications of the quality tool were stable and reliable applications, performance, maintainability, scalability: adding new modules, features, methods or algorithms should be easy, security, authentication, recoverability, standards and tools (programming languages) and accessibility (web application) are also needed.

The QMT prototype is illustrated in Fig. 2. The transfer of the information from the manufacturing process to the end users is presented in the following four steps that are (1) data acquisition, (2) data storage, (3) information analysis and (4) information delivery. In most advanced visualizations in our tool, the information has been refined to knowledge with automatic interpretation of the results.

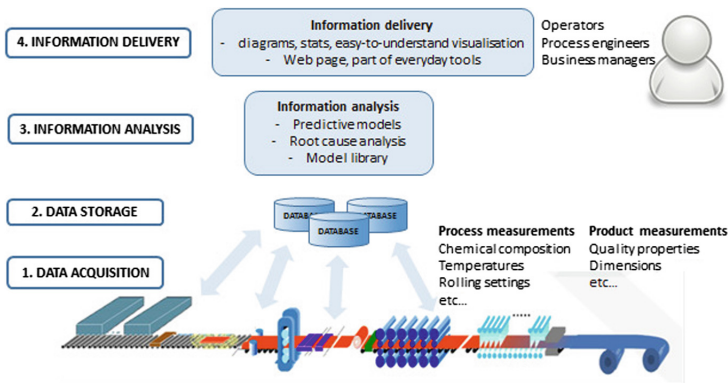


Fig. 2. The prototype of QMT.

## 2.2 The Specifications for the Tool

The quality information in the QMT is based on the statistical prediction models implemented in R language and equations and rules implemented in C++Mathematical ExpressionToolkit Library (ExprTk). R is a free and open source language for statistical computing. R is integrated into QMT with RServe module, which allows other programs to use facilities of R. R scripts can be written standalone and integration into QMT is straightforward. ExprTk is a mathematical expression parsing and evaluation engine. It is integrated into QMT by including it directly to the source code.

QMT server side is implemented in C++ language. Server side functionality of QMT includes data access and integration into models. Online data access

for QMT is accomplished by reading data from a database. Typically, selected database views are created for accessing data from a database. Some data pre-processing is needed before data can be used for model calculation. For example, a valid range for all model input variables has been defined, and if these limits were violated, model result may not be reliable which is shown with a question mark in the QMT user interface.

The QMT user interface is web based, and in typical use, it provides an overview of process quality as a starting point. In the quality overview, colour coded bars present the quality status of different process phases for each product during the selected time span. Typically, red colour indicates process failure or malfunction, yellow a warning for a process failure and green normal operation. Additionally, white colour indicates that quality information could not be calculated for some reason. Figure 3 illustrates a screen shot from the user interface. The overview to the process shows the predicted quality based on several quality models at different process steps. It hides mathematical models and all process variables from which the quality information is composed. The user can define the relevant quality models to be presented, and if any specific product looks interesting, the tool provides a possibility to analyse it further just by clicking the corresponding bar. Naturally, different user groups require different kinds of views to QMT based on their needs.

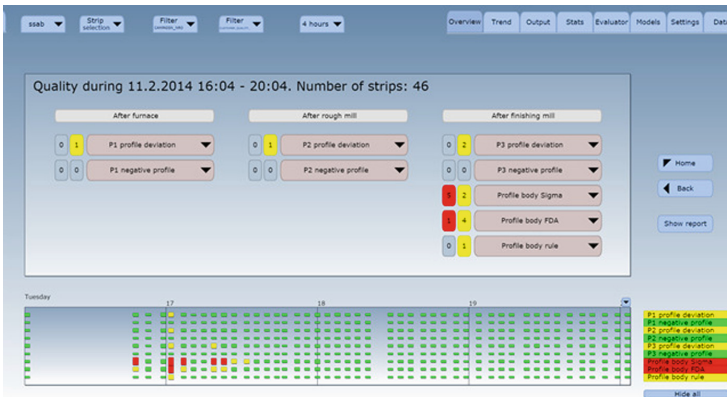


Fig. 3. The user interface of QMT. (Color figure online)

### 3 Statistical Quality Models

In industrial applications, high nonlinearity and many interactions between process settings challenge the performance of the models. Furthermore, the information about the relations between the predicted variable and the explanatory variables should be available, as for the user, the prediction itself is as valuable as the information about the effects of the process variables to the predicted quality

property. With an online system, the functionality of the tool would suffer, if the models were not capable of processing observations with missing data.

During the last two decades, neural networks have been a popular method for modelling data with complex relations between variables [10–12]. Lately, ensemble algorithms have risen to challenge them with equal accuracy, faster learning, tendency to reduce bias and variance, and they are more likely to over-fit. Seni and Elder state that ensemble methods have been called the most influential development in data mining and machine learning in the past decade [13]. Gradient boosting machines are a family of powerful machine learning techniques that has been successfully applied to a wide range of practical applications [14]. Boosted regression trees are capable of handling different types of predictors and accommodating missing data, there is no need for prior transformation of variables, they can fit complex nonlinear relationships, and automatically handle interactions between predictors [15]. For QMT, the generalized boosted regression models (GBM) were selected, and details of this model can be found in [16]. Juutilainen *et al.* presents in detail how to build models for rejection probability calculation in industrial applications [17].

## 4 Quality Monitoring in Manufacturing

Two case studies from steel industry were selected to demonstrate the use of the QMT. In case 1, a typical end user is a process engineer with an interest in detailed information about the process and with a need to find root causes for decreased quality. In case 2, a typical end user is an operator with a need for simple and easy-to-interpret presentations about the possibilities of how to improve the quality online.

### 4.1 Case 1: Strip Profile

A steel strip profile is a quality property for which the product development and the customer set a target value. This information is also essential for the following process steps; especially a negative profile can be very harmful during cold rolling. The target for profile locates typically between 0.03–0.08 mm. Because during the rolling schedule, the target can vary from product to product and strip to strip, profile adaptation is not possible, it is more difficult to hit the target every time. With prediction models, it is possible to design products that more likely fulfil the requirements, as well as to find root causes for the failure. In our QMT, the user can select between the profile and the deviation from the target profile models, depending on the needs.

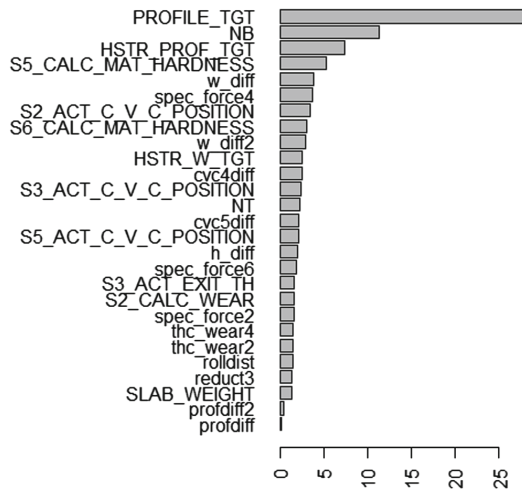
A typical user could be a process engineer, who wants to learn more about the process and improve it by designing new process practices or product types. The user would expect to have the following outputs that would assist him/her in decision making:

- colour-coded predicted quality during a selected time span for each product

- for a selected product, details about the related process parameters
- information if the model is extrapolating, e.g. some parameter values exceed the training data, and thus, the prediction may be less reliable
- details and visual information about the parameters in the model; what are the most important factors affecting the quality and how do they affect it
- if the product is predicted to have lower quality, how does it differ from the good ones and what could be done differently.

The information flow can get easily overwhelming, and the customization of the result presentation becomes crucial. It is important that the user can find the preferred analysis tools easily and the automated interpretation of the results is provided to speed up the decision making.

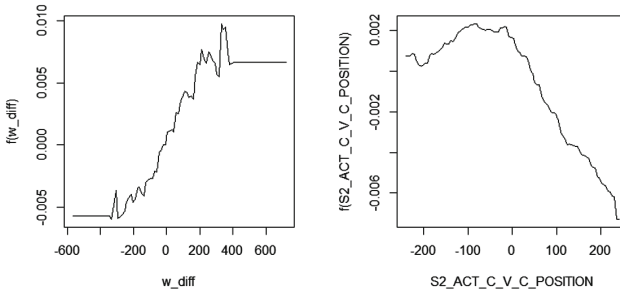
By observing the quality prediction model, the user can learn more about the quality property itself and how different process parameters affect it. The strength of the influence for each variable in the model correlates with the actual impact on the quality, and when the quality needs to be improved, the strongest variables are the first candidates to be considered. Figure 4 presents the relative influence of the variables in the profile deviation model. For example, process factors that relate to strength of the steel have a high impact on the profile deviation risk.



**Fig. 4.** The visualized variable importance in the GBM model for quality prediction.

The GBM model enables the visual inspection of the effect of each variable in the model. Thus, the user can learn to understand the manufacturing process better. Figure 5 presents two variables from a profile deviation model. The desired value for the property would be zero, and it can be seen that changes in strip width from product to product will increase the risk of profile

deviation to both directions, whereas the large positive values of roll position ( $S2\_ACT\_C\_V\_C\_POSITION$ ) will increase the risk of negative profile deviation.



**Fig. 5.** The visualized effects of two variables in the GBM model for profile deviation prediction.

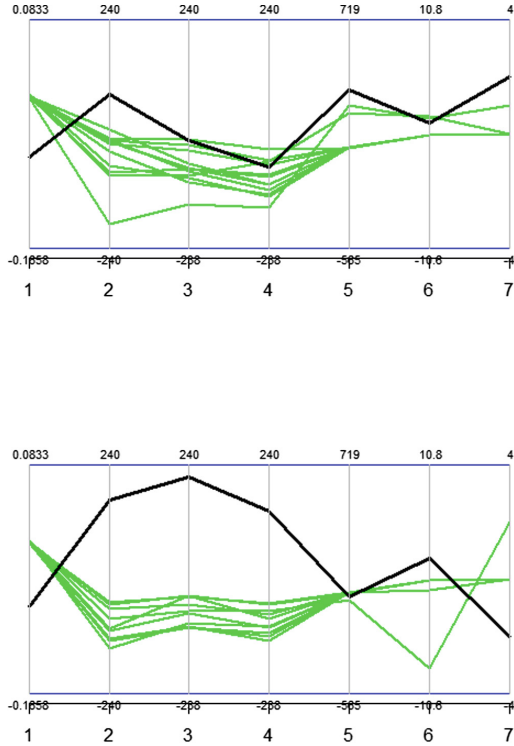
While analysing the prediction results of a product with a risk of a failure, the user typically asks what the difference between the observed one and good products is. Visually the difference can be effectively presented with parallel coordinates [18]. The comparison is not simple though, because, the determination of a similar good products is a complex procedure. There is not only one way of making each product, but it is always a compromise forced by the status of the production line. Furthermore, there can be hundreds of different products with small modifications depending on the customer. In this application, the weight and height of a strip were used as similarity measures when fetching the best products from a pool of good products that can be used as examples of good production practices. Figure 6 presents two examples of products with negative predicted deviation from the profile target (black) and their comparison with similar good products (green). In the upper case, the observed product seems to have a bit higher value for parameters 2 and 7, but no clear candidates for quality improvement cannot be determined. In the lower case, variables 2, 3, 4 and 7 have a significant difference to the good products, and thus, the user will learn that those settings possess a higher risk for failure.

The comparison can be done also by calculating the distances between the curves. This way, a threshold can be set for showing high enough deviations from the good products. In Fig. 7, the distances of an observed product are visually compared with the good ones. The customisation of the QMT has been made possible by offering several visualisation tool choices for the user.

## 4.2 Case 2: Strip Roughness

The central roughness of a stainless steel strip is a defect that appear after cold-rolling and surface treatments. The tendency to suffer from this defect type depends on the chemical composition and mechanical properties of the product,





**Fig. 6.** The parallel coordinates visualize the difference between the good products (green) and the observed product (black) having an increased predicted risk for a failure. (Color figure online)



**Fig. 7.** The deviations of the model variables from the good products for a selected observation.

but also cold-rolling process parameters have a high impact on the surface. The QMT provides the user an easy to follow overall view to the process, and the user gets simple suggestions for improving the process, if an increased risk of defect occurred.

A typical user could be a process operator, who needs to concentrate on various information sources simultaneously. The presentation of the predicted quality have to be simple and it should support decision making when there is a limited time to react. The user would expect to have the following outputs that would assist him/her in decision making:

- colour-coded predicted quality during a selected time span for each product
- clear visualization of recommended actions for a product with defect risk.

It is important that only relevant information is presented, or the user might start to neglect it. The chemical composition may have a high influence on the product quality, but at this point, the process operator have no possibility to modify it, and thus, the information is meaningless for the user. Instead, the process engineer is responsible for the improvement of the whole manufacturing process.

Figure 8 presents the information provided for the process operator, when the observed product has a risk of surface defect. It is easy to select which parameter should be adjusted, when no distracting information is present.

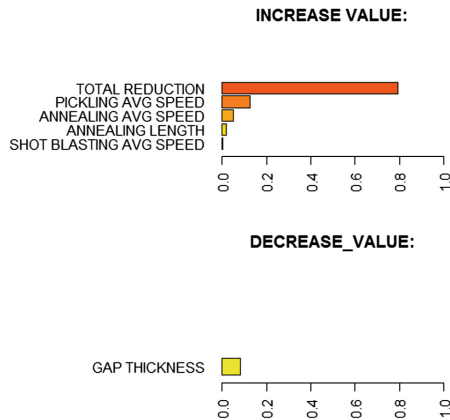


Fig. 8. Recommendations for process improvement.

## 5 Conclusion and Perspectives

This paper presented an online quality monitoring tool for information acquisition and sharing in manufacturing. The web based tool provides decision support for users in different roles in manufacturing. Furthermore, the user can find root-causes for the reduced quality and learn how to improve the process.

Statistical quality models predict the quality of each product during the manufacturing, and the results are colour-coded to easily interpreted visual presentations. When the user notices a deviant product or a period of defected products, it is easy to fetch more information about the product by selecting suitable actions. The provided visualizations will help to understand the model and factors that affect the prediction, and thus, the predicted quality as well. More advanced methods link the observed products with successful similar products and highlight the differences in production. It can also recommend actions for quality improvement.

The QMT is having an online test period at both participating factories. The user feedback will provide us valuable information for further development of the tool. New user groups with different needs for information presentation will be included in the tool later. In its current version, the selected product can be compared with good ones fetched from a saved data set that has a large presentation of different product types. Later, the dynamicity will be improved by allowing the QMT to fetch an up to date comparison data from the online data base. As a result, it will be faster to find process settings that may be causing quality issues in constantly changing environment.

## References

1. Harding, J., Shahbaz, M., Srinivas, Kusiak, A.: Data mining in manufacturing: a review. *J. Manuf. Sci. Eng.* **128**, 969–976 (2006)
2. Siirtola, P., Tamminen, S., Ferreira, E., Tiensuu, H., Prokkola, E., Rönning, J.: Automatic recognition of steel plate side edge shape using classification and regression models. In: *Proceedings of the 9th Eurosim Congress on Modelling and Simulation (EUROSIM 2016)* (2016)
3. Phillips-Wren, G.: Intelligent decision support systems. In: *Multicriteria Decision Aid and Artificial Intelligence*. Wiley, Chichester (2013)
4. Logunova, O., Matsko, I., Posohov, I., Luk'ynov, S.: Automatic system for intelligent support of continuous cast billet production control processes. *Int. J. Adv. Manuf. Technol.* **74**, 1407–1418 (2014)
5. Dumitrache, I., Caramihai, S.: The intelligent manufacturing paradigm in knowledge society. In: *Knowledge Management. InTech*, pp. 36–56 (2010)
6. Bi, Z., Xu, L., Wang, C.: Internet of things for enterprise systems of modern manufacturing. *IEEE Trans. Ind. Inf.* **10**(2), 1537–1546 (2014)
7. Xu, L., He, W., Li, S.: Internet of things in industries: a survey. *IEEE Trans. Ind. Inf.* **10**(4), 2233–2243 (2014)
8. Akram, M., Saif, A.W., Rahim, M.: Quality monitoring and process adjustment by integrating SPC and APC: a review. *Int. J. Ind. Syst. Eng.* **11**(4), 375–405 (2012)
9. Kano, M., Nakagawa, Y.: Data-based process monitoring, process control, and quality improvement: recent developments and applications in steel industry. *Comput. Chem. Eng.* **32**(1–2), 12–24 (2008)
10. Bhadesia, H.: Neural networks in materials science. *ISIJ Int.* **39**(10), 966–979 (1999)
11. Tamminen, S., Juutilainen, I., Rönning, J.: Quantile regression model for impact toughness estimation. In: Perner, P. (ed.) *ICDM 2010. LNCS (LNAI)*, vol. 6171, pp. 263–276. Springer, Heidelberg (2010). [https://doi.org/10.1007/978-3-642-14400-4\\_21](https://doi.org/10.1007/978-3-642-14400-4_21)

12. Tamminen, S., Juutilainen, I., Rönning, J.: Exceedance probability estimation for quality test consisting of multiple measurements. *Expert Syst. Appl.* **40**, 4577–4584 (2013)
13. Seni, G., Elder, J.: Ensemble methods in data mining: improving accuracy through combining predictions. In: *Synthesis Lectures on Data Mining and Knowledge Discovery*. Morgan & Claypool, USA (2010)
14. Natekin, A., Knoll, A.: Gradient boosting machines, a tutorial. *Front. Neurobot.* **7** (2013)
15. Elith, J., Leathwick, J., Hastie, T.: A working guide to boosted regression trees. *J. Anim. Ecol.* **77**, 802–813 (2008)
16. Friedman, J.: Stochastic gradient boosting. *Comput. Stat. Data Anal.* **19**, 367–378 (2002)
17. Juutilainen, I., Tamminen, S., Rönning, J.: A tutorial to developing statistical models for predictiong disqualification probability. In: *Computational Methods for Optimizing Manufacturing Technology, Models and Techniques*, pp. 368–399. IGI Global, USA (2012)
18. Inselberg, A.: Visual data mining with parallel coordinates. *Comput. Stat.* **13**(1), 47–63 (1998)