Extended Abstracts
ISBN 978-952-5183-56-6

Henna Tiensuu*, Satu Tamminen, Olli Haapala, Juha Röning

Intelligent methods for root cause analysis behind the center line deviation of the steel strip

Extended abstract: This article presents a statistical prediction model-based intelligent decision support tool for center line deviation monitoring. Data mining methods enable the data driven manufacturing. They also help to understand the manufacturing process and to test different hypotheses. In this study, the original assumption was that the shape of the strip during the hot rolling has a strong effect on the behaviour of the steel strip in Rolling, Annealing and Pickling line (RAP). Our goal is to provide information that enables to react well in advance to strips with challenging shape. In this article, we show that the most critical shape errors arising in hot rolling process will be transferred to critical errors in RAP-line process as well. In addition, our results reveal that the most critical feature characterizes the deviation better than the currently used criterion for rework.

In Tornio, the stainless steel is cold rolled with an integrated rolling, annealing and pickling line, which is called a RAP-line. The center line deviation of the steel strip in the RAP-line is a major quality factor that can produce serious problems if the strip did not stay in place during RAP processing. In the worst case, a diverged strip position can stop the whole production and brake the devices. Hot rolling strip center line deviation at the previous process step is a commonly used and easily available measurement of the strip shape. It could be possible to simulate the behaviour of the strip with a certain shape in the RAP-line during different individual phases. Ilmola et al present a very exact picture of the rolling process of the steel strip in two parts, the hot rolling and heat transfer and microstructure formation during water-cooling [1]. However, in this case, the incoming strips at the RAP-line include a large amount of different shapes, and to get the big picture of the whole production, all of them require simulating. In addition, there are hundreds of products behaving differently because of their chemical composition or the mechanical properties, and thus that the simulation work would be huge if not practically impossible. With data mining methods, the big picture of the process can be achieved quickly. These methods can process a large set of different process settings and products data simultaneously and take into account the variation caused by the measurement error or actual realization related to the features of the products and the process parameters [2], [3], [4]. Powerful machine learning methods are capable of modeling highly nonlinear process parameter dependencies and enable the effective use of the process data.

In this application, the center line deviation is predicted using gradient boosting methods (GBM) [5]. The idea of this machine learning algorithm is to form a strong learner by combining together the set of iteratively estimated weak learners. The model is able to treat efficiently the complex and nonlinear relationships within the data set, which is mandatory with industrial applications. Other advantages are that the method is capable of processing observations with missing information, and contrary to neural networks, it works also with smaller data sets. Due to the complexity of the center line deviation assessment, a lot of attention has been paid to the selection of the response variable that describes the position deviation best. We ended up considering the 90 m of the inner circle of the strip coil and calculated the mean of the process variable, ST6 super position for describing the center line position of the strip. Next, we applied an average filtering (n=10) for ST6 super position values of the selected length and calculated the difference between the mean of the ST6

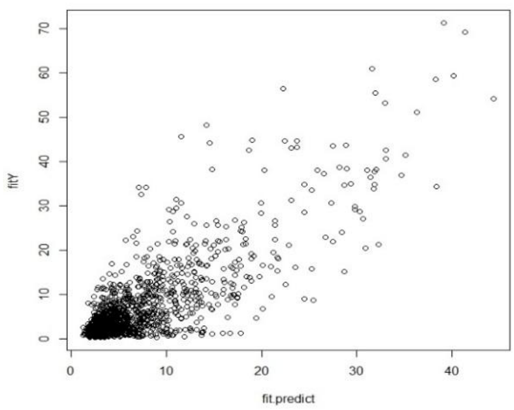

Figure 1. Predicted ST6 super position (x-axis) vs. measured values (y-axis) in the test set. A product group with inferior performance is highlighted (black).

super position and the filtered curve. The formed response variable was predicted with GBM-model that was implemented with R-program. The overall correlation between the measured and predicted values of the test set was 0.74, the root mean square error was 7.2 and the mean absolute error was 4.6. The scatterplot between the predicted response variable and the measured values in the test set is shown in Figure 1.

The modeling results indicate that the currently used criterion for rework is not the best candidate to characterize the strip behavior. Instead, the most important parameter is the slope_max, which describes the asymmetrical shape of the strip during hot rolling. From Partial dependence plot (PDP) in Figure 2, it can be clearly seen that the higher the slope_max is during hot rolling, the higher the center line deviation is in RAP process correspondingly. The increase in deviation is especially large after 40 units. This result proved the hypothesis that the shape errors in hot rolling forecast problems also in RAP-line. The thickness of the strip has an impact on the response variable as well. Our results reveal also the less critical shape errors, which are the wedge shape and the longitudinal deviation measured from the middle of the strip.

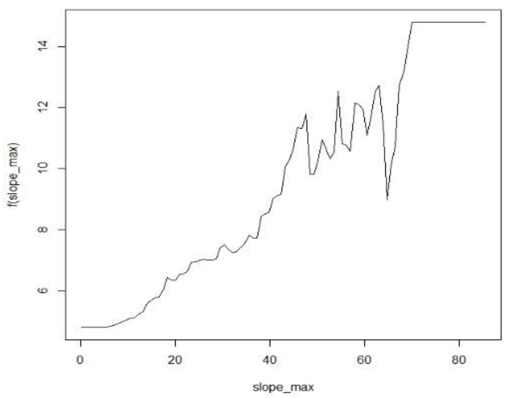

Figure 2. The PDP for slope_max reveals the feature's increasing impact on the response. Especially, after 40 units, the feature has a hazardous effect on the center line position.

It is important to reveal the reasons behind the predictions, especially, when a product has a higher predicted risk for the failure. In this application, SHapley Additive exPlanation (SHAP) values for each product were calculated based on the product group [6]. Thus, the products were compared to the average performance within the group of similar products. We selected one example to demonstrate the usage of the method. The product with bad prediction is shown in Figure 3. As can be seen, the prediction for center line deviation is 31.24 with the bad one, when the average prediction is 11.54 for this steel type in general. The phi value describes the strength of the feature value contribution in the prediction. The strongest candidate behind a poor prediction is slope_max. Also, the membership in the thinner strips class increases the risk.

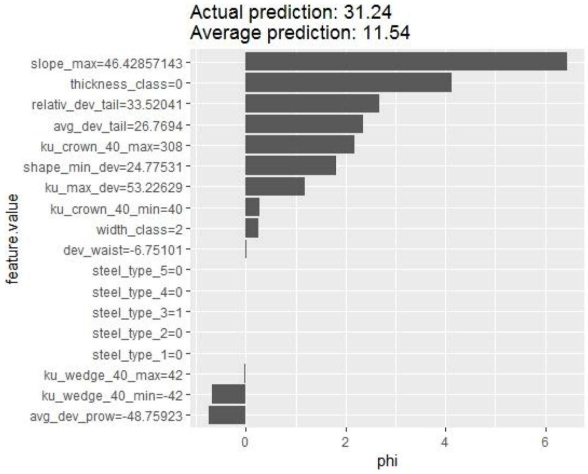

Figure 3. SHAP values for a bad product with a prediction of 31.24, while the average prediction is 11.54 inside the same steel group.

As a result of this research, we found the best response variable that describes the center line deviation. Our root cause analysis behind the center line deviation gives essential knowledge to the process development workers at the Tornio mill. We were able to define a clear critical limit value 40 for slope_max. In the first step, this information can be used for redirecting the strips which exceed the limit to repair treatment before RAP-line. It is also possible to implement the model for online use and to predict more efficiently the products that need to be repaired. The individual information behind the prediction for each product provides useful guidance for the rework.

The developed model enables the user to understand better the quality of the products, how the process works, and how the quality model predicts and performs. The results can be developed further to a smart decision support tool that helps to find out the best way of dealing with the critical products. The tool enables the identification of the quality problems of the steel strips at the earliest possible moment, which leads to the reduction of the rejection risk and to increased profits for the producer. Because of the laborious and manual data collection, the current model covers the main product groups. The performance of the model can be improved by collecting more data from the product groups with a smaller number of representatives, but online learning methods could provide a more serviceable solution in the long run.

Keywords: smart decision support, data driven manufacturing, artificial intelligence, data mining, steel strip rolling, GBM

*Henna Tiensuu: University of Oulu / BISG, E-mail: henna.tiensuu@oulu.fi

Satu Tamminen: University of Oulu / BISG, E-mail: satu.tamminen@oulu.fi

Olli Haapala: Outokumpu Stainless Oy, Tornio, E-mail: Olli.Haapala@outokumpu.com

Juha Röning: University of Oulu / BISG, E-mail: juha.roning@oulu.fi

References

[1] Ilmola J., Pohjonen A., Seppälä O., Leinonen O., Larkiola J., Jokisaari J., Putaansuu E., Lehtikangas P., Coupled multiscale and multiphysical analysis of hot steel strip mill and microstructure formation during water cooling. Procedia Manufacturing 15, proceedings of the 17th International Conference on Metal Forming, METAL FORMING 2018, Toyohashi, Japan, 2018, 65-71

[2] Tamminen S., Juutilainen I., Röning J., Exceedance probability estimation for a quality test consisting of multiple measurements. Expert Syst. Appl. 40(11):4577–4584, 2013

[3] Juutilainen I., Modelling of Conditional Variance and Uncertainty Using Industrial Process Data. Acta Universitatis Ouluensis: Technica, University of Oulu, 2006

[4] Engel J., Modelling variation in industrial experiments. Applied Statistics 41:579–593, 1992

[5] Friedman J. H., Stochastic gradient boosting. Comput. Stat. Data Anal. 38(4):367–378, 2002

[6] Lundberg S. M., Lee S. I., A Unified Approach to Interpreting Model Predictions. In: Advances in Neural Information Processing Systems 30, edited by Guyon I, Luxburg UV, Bengio S, Wallach H, Fergus R, Vishwanathan S, Garnett R, pp. 4765–4774. Curran Associates, Inc. 2017