

Recognizing Steel Plate Side Edge Shape automatically using Classification and Regression Models

Pekka Siirtola¹ Satu Tamminen¹ Eija Ferreira¹ Henna Tiensuu¹ Elina Prokkola² Juha Rönning¹

¹Biomimetics and Intelligent Systems Group, P.O. BOX 4500, FI-90014, University of Oulu, Oulu, Finland
{pekka.siirtola, satu.tamminen, eija.ferreira, henna.tiensuu, juha.roning}@ee.oulu.fi

²SSAB Europe, Raah plate mill, Finland

Abstract

In the steel plate production process it is important to minimize the wastage piece produced when cutting a mother steel plate to the size ordered by a customer. In this study, we build classification and regression models to recognize the steel plate side edge shape, if it is curved or not and the amount of curvature. This is done based on time series data collected at the manufacturing line. In addition, this information needs to be presented in a way that enables fast analysis and long-term statistical monitoring. It can then be used to tune the parameters of the manufacturing process so that optimal curvature can be found and the size of the wastage piece can be reduced. The results show that using the classification and linear regression methods, the side edge shape can be recognized reliably and the amount of curvature can be estimated with high accuracy as well.

Keywords: steel manufacturing, classification, regression, plate plan pattern control, plate side edge

1 Motivation

In the steel plate production process, the molten steel is first converted into slabs which still are hot and glowing. Then the thickness of the slab is reduced in the reversible multi-pass rolling process. Plate rolling can be divided in three main stages: sizing, broadside rolling and finishing rolling. In sizing stage the slab is rolled in longitudinal direction to produce required intermediate thickness. Before broadside rolling the slab is turned around 90 degrees and then rolled in transverse direction to obtain the required plate width. After broadside rolling the slab is turned again 90 degrees and rolled to final thickness (Ginzburg, 1989). After rolling the customer plates are cut from the mother plate. As the steel material and manufacturing are expensive, it is desired that the amount of cutting wastage is as small as possible. The uneven shapes at the plate end sides and lateral sides cause yield loss, amounting to about 5% to 6% of a total tonnage of slab used (Ruan et al., 2013). To minimize this loss, the shape of the rolled mother plates needs to be optimized.

During the rolling process, inhomogeneous plastic deformation occurs as width spread at the plate edge regions. There happens width spread at the plate side edge por-

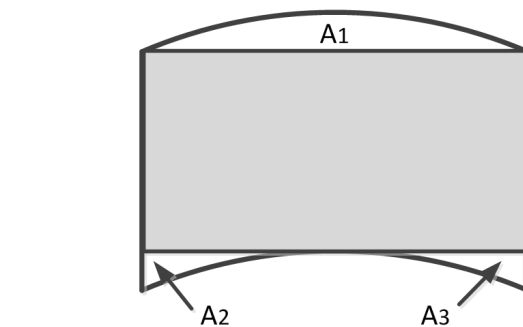


Figure 1. Wastage of concave side is much smaller than of convex side with same curvature but opposite direction, $A_1 > A_2 + A_3$.

tions, which forms convex shape at the plate ends. Laterally material spreads more at the both plate ends compared to central portion thus having a tendency towards concave plate side edge shape. However, because there is also broadside rolling phase in addition to the longitudinal rolling phase, the plate shape is dependent on the combination of the longitudinal rolling ratio and the broadside rolling ratio. The bigger the broadside rolling ratio is compared to longitudinal rolling ratio, the more convex the final plate side edge shape is formed. Width spread at the plate edge regions is the prime cause of the uneven shapes formed during the hot rolling process (Ruan et al., 2013). There are techniques designed to produce a true rectangular shape of rolled plate. One of the methods is MAS-rolling which was developed by Kawasaki Steel Corporation (Yanazawa et al., 1980).

However, due to rolling inaccuracies the rectangular shape cannot always be achieved even with MAS-rolling. There is width deviation not only between plates but also inside a plate. This results sometimes in convex and sometimes in concave plate side edge shape. However, the amount of wastage from convex side is much bigger than from concave side with same curvature, see Figure 1. In addition, concave shape increases the size of a rectangular shaped plate in camber shaped plates. Usually the width accuracy is as its best in the middle of the plate and weakens towards the ends. For this reason the deviations in the

ends are bigger and thus more extra material need to be designed there, which also favors concave shaped plate. Therefore, in order to minimize the amount of wastage, a slightly concave edge shape of a steel plate should be preferred. Using MAS-rolling, the target plate side edge shape can be designed slightly concave in order to almost completely avoid plates with convex edge shape. The desired amount of curvature can be defined by adjusting the process parameters. However, because of the uncertainty related to the steel plate production process, exactly the desired amount of curvature cannot be reached. In order to optimize the MAS-rolling parameters and thus curvature of a plate, the shape of the mother plates needs to be monitored.

Our study is made for SSAB Europe, Raahe plate mill, Finland, where the monitoring of the plate shape is currently done visually at the cooling banks by own eyes. There is no camera based monitoring system available and if plates that have already passed cooling banks needs to be viewed, plate shape can be visualized by means of collected process data. Plate shape information can be gathered from thickness gauge data, see Section 3 for more details. However this data is difficult to analyze making visual monitoring time consuming. In addition, conclusions made using visual monitoring are always based on subjective view. Furthermore, unlike our approach, this type of monitoring does not allow the modelling of statistical distribution of the curvature based on historical production data, which can be used to understand the uncertainty of the manufacturing process and to optimize process parameters to obtain the optimal curvature of plates. If the plate shape can be optimized, the mother plates can be manufactured by using smaller slabs which leads to better yield.

In this study, we build models to define the shape of a time series describing the steel plate side edge, if it is curved or not and the amount of curvature, in such form that they can be analyzed in a glance and enable long-term statistical monitoring. The article is organized as follows: related work is covered in Section 2 and the used data set is described in Section 3. Our method is introduced in Section 4, and the method is validated in Section 5. We will discuss our results in Section 6 and, finally, the conclusions are in Section 7.

2 Related work

Many aspects of steel manufacturing have been modelled and optimized using machine learning and data mining methods which have been used for instance to estimate impact toughness (Tamminen et al., 2010), to diagnose faults (Tian et al., 2015), to model the yield strength (Koskimäki et al., 2007) and to model the rolling temperature (Tiensuu et al., 2011).

In this study, we aim to define the curvature of a time series describing the steel plate side edge, and use this information to build statistical distribution model to visu-

alize what kind of curve shapes the studied data set includes and how the amount of curvature is distributed in the manufacturing process. In the literature the term *plate plan pattern control* is used to describe techniques and methods that are designed to ensure that produced steel plates have the desired shape and to minimize wastage (NIIR Board of Consultants & Engineers, 2006). Most of the studies related to plate plan pattern control concentrate on improving the rolling process by installing new equipment to the manufacturing line making them expensive, such as Zhang et al. (2015); Inoue et al. (1988), and there are not many studies where plate view pattern is improved by tuning the process parameters based on the data collected in the manufacturing line as we do in our study.

Lee et al. introduced a neural network based method to predict the width of the mother steel plate based on the size of a slab and different manufacturing parameters (Lee et al.). Their method can be used to select parameters to rolling process so that finished steel plate has the desired width. Juutilainen et al. predicted the steel strip's width rejection probability with statistical models and optimized the working allowance of the product based on the results, material and rejection costs (Juutilainen et al., 2012, 2015). While the width prediction of a steel plate is closely related to our study, our approach differs from it in a way that we study steel plates one side at a time. Therefore, our study gives better understanding about the plate plan pattern than studies which concentrate on the width of a plate and, therefore, study both sides simultaneously.

The studies by Ruan et al. (Ruan et al., 2014, 2015) are the closest to our study. In Ruan et al. (2014) regression analysis was used to build models to predict plate plan view pattern based on slab size and rolling parameters. Models were trained based on simulation results and validated using industry tests. The study concentrates on analyzing what causes convex and concave sides. In our study we are not trying to predict the shape of a side or find causes for certain shapes from the data collected during the rolling process. Instead, our aim is to get better understanding about the manufacturing process by monitoring the ratio of different shapes and the amount of curvature from the plain view patterns measured from finished steel plates. This enables better understanding about the uncertainty related to the manufacturing process and how process parameters should be selected to minimize the yield loss caused by uncertainty.

3 Data set and pre-processing

The data were collected from steel plate production line at SSAB Europe, Raahe plate mill, Finland. The measurements were done using a thickness measurement system from IMS Systems Inc. This measurement system contains three radiation sources and radiation detectors. When a steel plate moves through the production line, the amount of detected radiation can be used to measure not only the thickness of the plate but also its width. In this

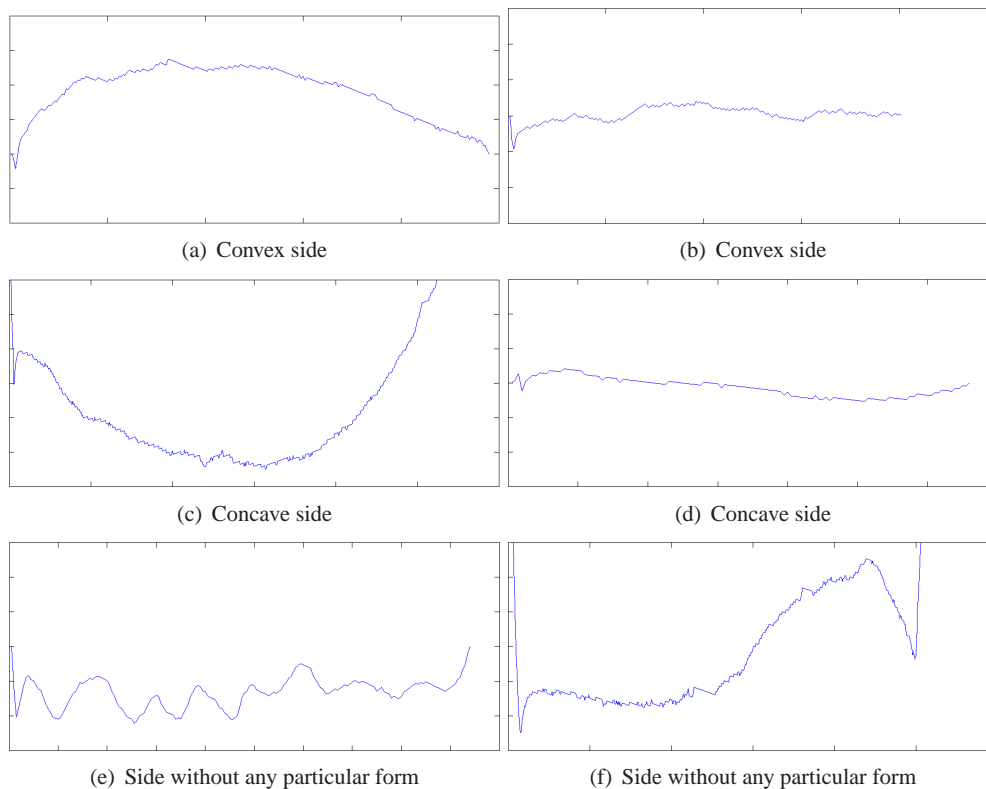


Figure 2. The data were classified into three classes: convex, concave, and other. This figure shows examples of members of different classes.

study these width measurements are considered as a time series that describes the shape of a steel plate side edge. The distance between two adjacent measurements is approximately 50 mm, so for an average plate of 20 m, each side can be described as a time series containing 400 measurements.

The data were collected from 399 plates, and therefore, from 798 sides. The length of the plates varied from 4.0 meters to 27.5 meters and width from 1.3 meters to 3.3 meters. Based on this data, the aim of this study was to build models to detect the direction of curvature and the amount of curvature of a curve describing the side of a steel plate. However, the data set did not include labels or correct values about the amount of curvature. Therefore, in order to build a model to recognize the shape of a steel plate side edge, it was necessary to label the data set by hand. When the data was visualized to label it, it was noted that there are many cases where it is difficult to say whether the curve describing the side is convex or concave because plates can be over 20 meters long but still width within the plate only differs a couple of millimeters. In addition, sides are not regular or symmetric, and they can contain features from both convex and concave shapes. It was also noted that all edges are not convex or concave, instead, due to uncertainty and errors in the manufacturing process sides can have other shapes as well. Eventually, the data were labeled into three classes: 'convex', 'concave', and 'other'. There were 160 convex sides, 556

concave sides and 82 instances classified as 'Other'. Examples of instances from each class are shown in Figure 2. Note that because the data is hand-labeled it can contain errors as in some cases it was difficult to conclude what is the correct class label for the side.

Before the data were analyzed, pre-processing was done. In the pre-processing stage, the measurements describing the head and tail of the side edge were removed by removing 10 percent of the measurements from both ends. This was done as typically the head and tail of the side are rounded, and therefore, they are not indicative of the shape of the rest of the side. In addition, when the width of the steel plate is measured, the plate is not necessarily positioned fully straight at the manufacturing line. This skewness caused by biased positioning was removed from the measurements in the pre-processing stage by straightening the time series describing the side of a steel plate based on the line described by the first and last measurement point of the side.

4 Defining the curvature of a steel plate

Recognition of the direction and the amount of curvature of a curve describing the steel plate side edge is basically a regression problem as the goal can be considered as one continuous value where the sign of the value tells the direction of curvature and absolute value describes the

amount of curvature. However, it is very difficult to visually estimate the amount of curvature and give an accurate continuous value of describing it. Therefore, the amount of curvature of time series T describing a side edge of a steel plate was estimated using the following equation:

$$curvature_{estim}(T) = \frac{abs(max(T) - min(T))}{\|T\|}. \quad (1)$$

This means that curvature estimation was done by calculating the amplitude of a time series T describing the steel plate side edge, and dividing this with the length of a steel plate. When this information was combined with class labels; defining whether the plate is convex, or concave; we build a response vector where the sign of a response tells the direction of the curvature, and the absolute value tells the amount of curvature. By applying regression analysis to this response vector and features extracted from the measurements it was possible to built a linear regression model to be used to define the shape of a plate, and to define the curvature of a plate.

In addition to convex and concave sides, there are also sides that are neither of these two options. Therefore, these need to be recognized from the data set before regression model can be trained. In this study, this was done using a binary classifier that classifies instances into two classes: the side is either curve (convex or concave) or not.

As a conclusion, the recognition of the direction and the amount of curvature of the side of a steel plate is divided into two tasks: (1) Classify instances into two classes: side is either an curve or not, and (2) Using data from convex and concave side's, train a regression model using estimation of the amount and direction of curvature as response.

In both tasks, the models were trained using the same feature set consisting of the following features: polynomials of degree 1, 2, and 4 were fitted to the time series describing the side and the obtained coefficients as well as the error of the fitting were used as features. Fitting was done not only to the whole time series but also different sizes of parts of it, for instance to the first and the second half separately. In addition, a straight line was fitted through the first and the last measurement of the time series and the ratio describing the number of points below and above this line was used as a feature. This was also done to different parts of the time series so that altogether 23 features of this type were extracted. Other features included minimum, maximum, and different percentiles of the values. The complete feature vector included 66 features.

5 Experiment

In this section, the method presented above is applied to the data set introduced in Section 3. As it was stated in the previous section, the recognition of the direction and the amount of curvature of a steel plate side edge is divided into two tasks: at first it was recognized whether the side was convex/concave or not. In the second phase, convex and concave sides were further studied to estimate the

Table 1. Recognizing non-curve side edges.

Classifier	Accuracy	Precision	Recall
QDA	94.6%	87.0%	82.5%
LDA	91.5%	77.1%	78.9%
C4.5	93.4%	86.2%	73.4%
kNN, k=1	93.9%	87.1%	76.0%
kNN, k=3	94.9%	91.5%	77.9%
kNN, k=5	95.0%	93.2%	78.8%

Table 2. Confusion matrix showing the classification results using QDA.

	Curve	Other
Curve	699	17
Other	26	56

amount of curvature as well as the direction of the curvature.

In the first part, the classification was done using four different methods to compare their performance. The classifiers used were kNN, LDA, QDA and C4.5. The idea of the k nearest neighbor classifier is to classify a data point into the class, to which most of its k nearest neighbors belong. In this study, k values 1, 3 and 5 were employed. Linear discriminant analysis (LDA) is used to find a linear combination of features that separate the classes best. The resulting combination may be employed as a linear classifier. QDA (quadratic discriminant analysis) is a similar method, but it uses quadratic surfaces to separate classes. C4.5 is a decision tree model that based on the difference in entropy partitions the space spanned by the input variables to maximize the score of class purity. This is done so that the majority of points in each cell of the partition belong to one cell (Hand et al., 2001).

In order to avoid over-fitting in the classification process, the data set were randomly divided into five parts and cross-validation was performed so that one part at the time was used for testing and the rest for training. The most descriptive features from the feature set were selected using a sequential forward selection (SFS) method.

The results are shown in Table 1. In the table, *accuracy* is defined as *true positives / all*, *precision* was calculated as *true positives / (true positives + false positives)* calculated for both classes and averaged, and *recall* stands for *true positives / (true positives + false negatives)* calculated for both classes and averaged. In addition, the results using QDA are shown in more detail in Table 2.

In the second part, a linear regression model was trained using the same feature set to estimate the direction and amount of curvature of the plate's side. Curvature values calculated using Equation 1 combined with information about the direction of the curvature were used as a response vector to this regression model. Table 3 demonstrates the accuracy of recognizing the direction of curve

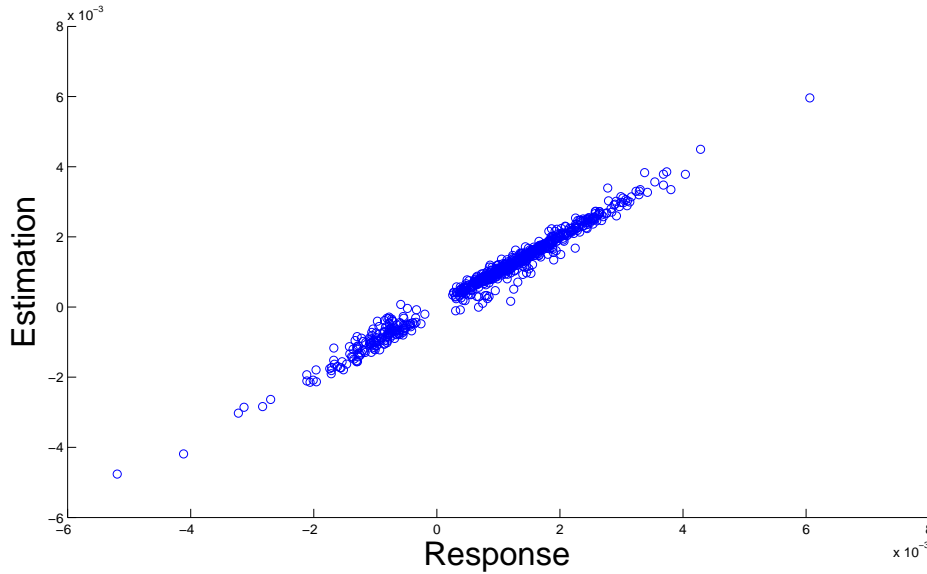


Figure 3. Using linear regression model the direction and curvature of a side can be estimated accurately.

Table 3. Three methods to detect the direction of the curvature of a side are compared.

Method	Accuracy
<i>Regression</i>	99.6%
<i>Intuitive</i>	96.5%
<i>Classification, QDA</i>	99.6%

curvature using the trained model. The results of regression model were compared to two alternative methods: *intuitive method* and *classification* using QDA. Intuitive method is simple, a line is drawn between the first and last observation of the time series, and the direction of curvature is decided based on which side of the line has more observations. In the case of QDA, 5-fold cross-validation was applied and side edges were classified as curve or not.

Finally, Figure 3 shows how well the estimation given by the regression model is comparable to the response calculated using Equation 1.

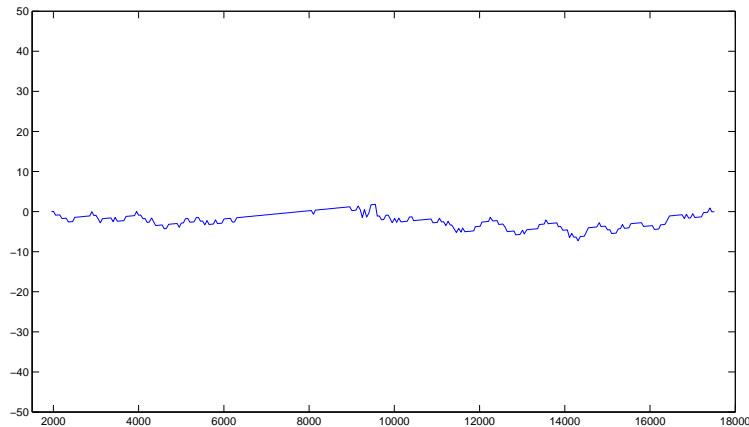
6 Discussion

In our approach, the recognition of the shape of the side of the mother steel plate was divided into two phases: at first we recognized whether the side was convex/concave or not. In the second phase, convex and concave sides were further studied to estimate the amount of curvature as well as the direction of the curvature. The recognition rates in Table 1 show that sides without a curve shape can be detected with high accuracy. All the tested classifiers performed well, only LDA had a bit lower recognition rates, although it could still recognize shapes with more than 90 percent accuracy. While the classification accuracy was

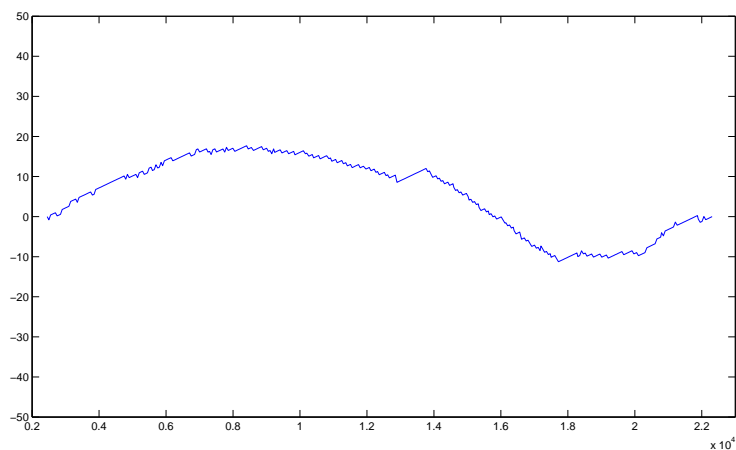
very high, the precision and recall were not. This is because class 'Other' had a lot lower recognition accuracy than class 'Curve'. This is partly due to the nature of the data set. It can be seen from Table 2 that the data set is very unbalanced. There are a lot more convex/concave sides than ones without any particular form. From manufacturing perspective this is of course good but from a model training perspective it is problematic because there is not the same amount of data from both classes.

Table 3 shows how well the direction of curvature can be detected. It can be seen that using linear regression model it can be detected almost perfectly, in fact in 714 cases out of 716 the curve direction was recognized correctly. Therefore, the method is very reliable. Detection of the direction of the curvature using the QDA classifier performs equally well, while the intuitive method gives a somewhat lower recognition accuracy. However, the advantage of the regression model compared to the classification is that it does not only detect the direction of curvature but it also estimates the amount of curvature, see Figure 3. However, the accuracy of the estimation of the amount of the curvature is more difficult to validate as the response used to train the regression model was also an estimation. In addition, the estimation calculated using Equation 1 is very vulnerable to anomalies of the data which can be caused for instance by errors in the measurements. In this sense, the estimation given by the regression model can be considered more stable and reliable as it is not based on only the measurements data of one side of a steel plate but on the whole data set including 798 sides. Therefore, anomalies of single sides do not have that much effect on the estimation given by the regression model.

Figure 4 shows two sides where the estimation given by the regression model differed the most compare to estimation obtained using Equation 1. In Figure 4(a) the amount



(a) Curvature of this side based on regression model is 1.1mm, and based on Equation 1, 9.1mm.



(b) Curvature of this side based on regression model is 17.4mm, and based on Equation 1, 26.7mm.

Figure 4. Sides where curvature estimations differ the most.

of curvature based on regression model is 1.1mm, meaning that thinnest point of the steel plate side edge is 1.1 mm narrower than the widest point. On the other hand, according to Equation 1 this difference is 9.1mm. In this case, Equation 1 has over-estimated the curvature because in this case there is a lot of variation in measurements making estimation based on amplitude inaccurate. Therefore, the estimation given by regression model, which suggests that side is almost straight, can be considered more reliable in this case. Also in Figure 4(b) the estimation given by regression model can be considered more reliable than the estimation obtained using Equation 1 which has over-estimated the curvature because of the drop at the end of the time series.

Our results show that using the presented method the shape of the side edge of a mother steel plate can be detected with high accuracy and the results can be presented in a way that enables analyzing a steel plate in a glance and enable long-term statistical monitoring. For example, our method to detect sides without curve shape can be used

to monitor the manufacturing process and detect if there are possible problems in the manufacturing process. Figure 5 shows the classified data set and sides without curve shape are drawn with higher spikes. Now based on this visualization it can be seen that there are intervals containing a lot of sides without curve shape, one for instance around x -axis value 40. These types of intervals can be a consequence of some unexpected change in the process, and monitoring enables a quick response to it. In addition, using the estimations obtained using the regression model, in Figure 6 the deviation of convex and concave sides is visualized. Negative value means that side edge is convex, bigger negative value stands for more convex side edge. Similarly, positive value means that side edge is concave and zero equals to straight side edge. Figure 6 shows that majority of the sides are concave but there is still quite many convex plates. Based on these statistics it is possible to define new parameter values for the rolling process to minimize the number of convex plates and to minimize the cutting wastage.

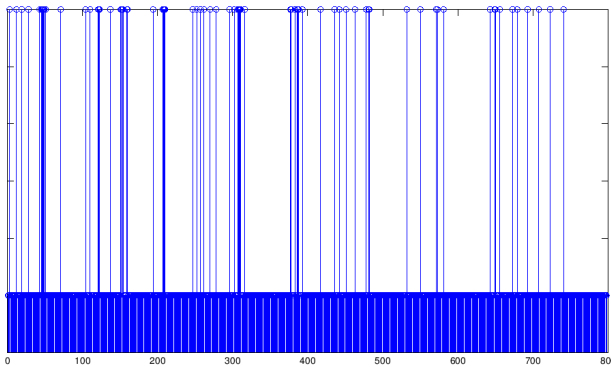


Figure 5. Many of the sides without curve form are centered at clusters. Short bars describe sides with curve form and long without.

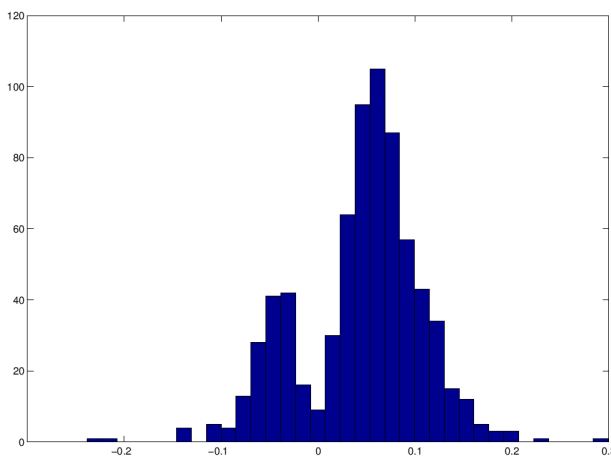


Figure 6. Curvature estimation can be used to optimize rolling parameters. In this case, rolling parameters should be tuned to reduce the number of convex sides.

7 Conclusions

In this study we developed a method to define the steel plate side edge shape based on data collected at the manufacturing line. The method consists of two phases: at first we recognize whether the side is a curve or not, and then we estimate the amount of curve curvature as well as the direction of the curvature. According to our experiments, classification methods can accurately be used in the first phase to detect if the side is a curve. In addition, our experiments show that linear regression model can be used to accurately recognize not only the direction of the curve curvature but also to estimate the amount of curvature.

Acknowledgment

The authors would like to thank SSAB Europe, Raahel plate mill, Finland for providing the data and their expertise for the study. Further acknowledgments are given to the Finnish Funding Agency for Technology and Innovation (TEKES) and Infotech Oulu for supporting this research.

References

- V. Ginzburg. *Steel-rolling technology: theory and practice*. Marcel Dekker, Inc., 1989.
- D. J. Hand, H. Mannila, and P. Smyth. *Principles of data mining*. MIT Press, Cambridge, MA, USA, 2001. ISBN 0-262-08290-X.
- M. Inoue, K. Ohmori, T. Orita, I. Okamura, S. Isoyama, and M. Tarui. Development of a process for manufacturing trimming free plates. *Transactions of the Iron and Steel Institute of Japan*, 28(6):448–455, 1988.
- I. Juutilainen, S. Tamminen, and J. Rönning. A tutorial to developing statistical models for predicting disqualification probability. In *Computational Methods for Optimizing Manufacturing Technology, Models and Techniques*, pages 368–399. IGI Global, 2012.
- I. Juutilainen, S. Tamminen, and J. Rönning. Density forecast based failing probability predictors in manufacturing. *European Journal of Industrial Engineering*, 9(4):432–449, 2015.
- H. Koskimäki, I. Juutilainen, P. Laurinen, and J. Rönning. Detection of the need for a model update in steel manufacturing. In *Proc. of the Fourth International Conference on Informatics in Control, Automation and Robotics (ICINCO 2007), Angers, France*, pages 55–59, 2007.
- D. Y. Lee, H. S. Cho, and D. Y. Cho. A neural network model to determine the plate width set-up value in a hot plate mill. *Journal of Intelligent Manufacturing*, 11(6):547–557.
- NIIR Board of Consultants & Engineers. *Steel Rolling Technology Handbook*. ASIA PACIFIC BUSINESS PRESS Inc., 2006.
- J. H. Ruan, L. W. Zhang, S. D. Gu, J. Zhang, W. He, and S. H. Chen. Establishment of models for plan view pattern control in heavy plate rolling process based on 3-d fem simulation. *International Journal of Materials and Product Technology*, 47(1-4):103–125, 2013.
- J. H. Ruan, L. W. Zhang, S. D. Gu, W. B. He, and S. H. Chen. Regression models for predicting plate plan view pattern during wide and heavy plate rolling. *Ironmaking & Steelmaking*, 41(9):656–664, 2014.
- J. H. Ruan, L. W. Zhang, Z. G. Wang, T. Wang, Y. R. Li, and Z. Q. Hao. Finite element simulation based plate edging model for plan view pattern control during wide and heavy plate rolling. *Ironmaking & Steelmaking*, 42(8):585–593, 2015.
- S. Tamminen, I. Juutilainen, and J. Rönning. Quantile regression model for impact toughness estimation. In *Advances in Data Mining. Applications and Theoretical Aspects*, pages 263–276. Springer Berlin Heidelberg, 2010.
- Y. Tian, M. Fu, and F. Wu. Steel plates fault diagnosis on the basis of support vector machines. *Neurocomputing*, 151, Part 1:296 – 303, 2015. ISSN 0925-2312.

- H. Tiensuu, I. Juutilainen, and J. Rönning. Modeling the temperature of hot rolled steel plate with semisupervised learning methods. In *Proc. 14th International Conference on Discovery Science, Lecture Notes in Computer Science*, volume 6926, pages 351–364. Springer, October 5-7 2011.
- T. Yanazawa, J. Miyoshi, K. Tsubota, T. Ikeya, H. Kikugawa, and K. Baba. Development of a new plan view pattern control system in plate rolling. *Kawasaki steel technical report*, pages 33–46, 1980.
- T. Zhang, B. Wang, Z. Wang, and G. Wang. Side-surface shape optimization of heavy plate by large temperature gradient rolling. *ISIJ International*, advpub, 2015.