Machine Learning with Applications 7 (2022) 100245

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

# Multistep networks for roll force prediction in hot strip rolling mill

Shuhong Shen [a], Denzel Guye [a,*], Xiaoping Ma [b], Stephen Yue [a], Narges Armanfard [c,d]

[a] Department of Mining and Materials, McGill University, 3610 rue University, Montréal, QC, Canada, H3A 0C5
[b] Algoma Steel Inc., 105 West Street, Sault Ste. Marie, Ontario, Canada, P6A 7B4
[c] Department of Electrical and Computer Engineering, McGill University, 3480 rue University, Montréal, QC, Canada, H3A 0E9
[d] Mila - Quebec AI Institute, 6666 St Urbain St, Montréal, Quebec H2S 3H1, Canada

ARTICLE INFO

*Keywords:*
Machine learning
Multistep neural networks
Hot rolling
Roll force prediction

ABSTRACT

Hot rolling processes consist of multiple single rolling stand operating at high temperature and speed to achieve desired steel shapes and superior properties, via exerting roll forces that need to be accurately predicted by a model. The currently used model of the mill of this study shows prediction instability and is unable to accurately accommodate changes in steel grade. In this paper, we propose a machine learning based framework to establish a model that accurately predicts roll forces at each mill stands of the hot strip rolling mill. In contrast to the traditional models, the proposed expert system considers an individual model for each rolling stand and employs rolling history when predicting roll forces. The proposed model includes both steel chemistry and physical process parameters for its predictions. Our experimental results demonstrate that the proposed framework improves both prediction accuracy and stability by 40%–50% over the currently used mill model. The enhanced prediction accuracy will greatly improve dimensional and microstructural control, as well as ensuring the avoidance of mill overloads.

## 1. Introduction

The increasing market demand for high quality steel products that are economically viable requires increasingly tighter control at every stage of materials processing. Hot rolling is a key process step where a number of rolling stands (aka passes) reduce the thickness of hot steel slab via exerting tremendous roll forces. Conventional mathematical models (Hodgson & Gibbs, 1992; Maccagno, Jonas, Yue, McCrady, Slobodian, & Deeks, 1994; Siciliano & Jonas, 2000) predict the rolling forces taking into consideration of various process parameters, such as roll speed, reduction ratio, temperature, etc. and descriptions of metallurgical phenomena like recrystallization, precipitation and grain growth etc. These models are semi-empirical and rely on manually fitting parameters in physics-based equations using previously obtained rolling data.

Fig. 1 shows the configuration of a 7-stand hot strip mill set up at Algoma Steel. The rolling mill is equipped with a slab reheat furnace, followed by an edger for width control before a single roughing mill stand. The pre-deformed strip is then de-scaled and is transported into 6 stands of 4-high rolling mill for stepwise thickness reduction, and finally enters cooling water wall and coilers. This rolling mill is capable of producing 1.5 to 16 mm thick strips depending steel grade by altering the gauge (gap between the rolls) setup in its online pass schedule design and control model. For process control, the slab dimensions and temperature are measured before the slab arrives at the first rolling stand. These measured parameters and the desired strip properties are given to the mill model to set up the roll gap and rolling speed for each stand, based on predictions of temperature, and roll force, torque for each rolling stand from R1 to F6. These estimations are then fine tuned based on the production performance, to create a final strip with desired properties.

According to the literature on metallurgical models, all existing roll force prediction models assume that a single common model can be applied to all stands when predicting forces – i.e., they tend to treat all rolling passes in the same way, without explicitly distinguishing the different levels of impact of the process parameters, which we hypothesize is the cause of various limitations. Furthermore, they ignore the effect of rolling history, over the successive passes, when predicting forces.

In this paper, to provide more stable and accurate roll force predictions, we propose an expert machine learning based system that considers an individual model for each rolling stand and employs rolling history when predicting roll forces. The proposed system considers both chemistry (e.g., C, Mn, Si, etc.) and process parameters (such as slab width, thickness reduction, roll speed, etc.) for its predictions. Among various machine learning techniques, artificial neural networks (ANN) are known for their flexibility in model design, high noise tolerance and exhibited high prediction accuracy (Singh & Chauhan, 2009).

* Corresponding author.
*E-mail addresses:* shuhong.shen@mail.mcgill.ca (S. Shen), denzel.guye@mail.mcgill.ca (D. Guye), xiaoping.ma@algoma.com (X. Ma), steve.yue@mcgill.ca (S. Yue), narges.armanfard@mcgill.ca (N. Armanfard).

https://doi.org/10.1016/j.mlwa.2021.100245
Received 19 March 2021; Received in revised form 15 December 2021; Accepted 27 December 2021
Available online 29 December 2021

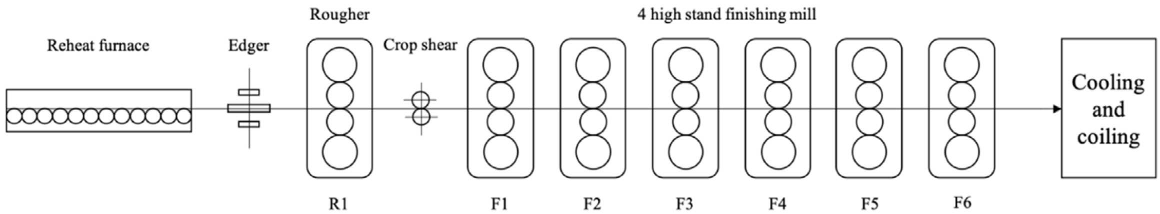


**Fig. 1.** Configuration of a 7-stand hot strip mill.

Inspired by the multi-pass rolling process, the proposed framework is a multi-step network (MSN) where the input and structure of each step is customized for a corresponding specific rolling pass. Three variations of MSN are devised to investigate the effect of process parameters and rolling histories. Experimental results demonstrate that MSN achieves a high prediction accuracy and stability over the current traditional mill mode predictions.

The remainder is organized as follows: related works are presented in Section 2. Section 3 presents the proposed multistep network roll force prediction. In Section 4, experimental results, which demonstrate the performance of the proposed method are presented. Conclusions are drawn in Section 5.

## 2. Related works

Portmann, Lindhoff, Sorgel, and Gramckow (1995) integrated Artificial Neural Network (ANN) into conventional physics-based algorithmic models as a correction factor that is learned from measured actual roll force and adjusts the computation formulae based on the prediction error. Martinetz, Protzel, Gramckow, and Sörgel (1995) proposed another hybridized algorithmic model with neural network as a part of process optimization system to control roll force, width and strip profile, considering process parameters including strip chemistry, roll speed, temperature, width, etc. Cho, Cho, and Yoon (1997) have proposed single stand substitutive and corrective neural networks taking both calculated values from mechanistic models and measured process parameters as input to predict roll force in cold rolling. They found that a committee of substitutive networks that directly predicts the roll force showed the greatest improvement over mathematical models. Larkiola, Myllykoski, Korhonen, and Cser (1998) summarized a series of applications of machine learning techniques in both process control and prediction of coil mechanical properties. High accuracy and fast computation speed were featured using ANN in the presence of large quantity of process data. A neural network based model developed by Niu, Sun, and Karjalainen (2000) was capable of accurately predicting austenite grain size of microalloyed steel within ± 4% error in lab studies and showed a significant improvement over regression technique. Lee and Choi (2004) used on-line adaptable neural networks in a plate mill to predict the ratio of conventionally modeled roll force to actual values as a correction factor in the algorithmic model. A significant decrease in mean prediction error and standard deviation in both offline training and online trials was attributed to the short-term learning taking rolling-condition-dependent errors and unexpected disturbances into account. Their network input features considered both the effect of major chemical compositions, namely carbon (C), manganese (Mn) and silicon (Si), and predicted roll force ratios in the previous two passes to account for systematic error. Son, Lee, Kim, and Choi (2005) further proposed an online learning network to account for both long-term learning with lot change, and short-term learning without new data input, coupled with mathematical roll force prediction model. Yang, Linkens, and Talamantes-Silva (2004), on the other hand, developed a standalone roll force predicting network ensemble. However, the model was trained using data generated from an orthogonal array experiment design combined with finite element simulation, as opposed to rolling mill data. Moreover, the input features only considered mechanical parameters, including heat transfer coefficient and friction, which are not measurable in real-life production. Later, Öznergiz, Özsoy, Delice, and Kural (2009) established three individual single-output neural networks predicting roll force, torque and slab temperature for one rolling stand of a reversible mill and reported a notable improvement on existing prediction models. Neural networks were also adopted in more recent studies attempting to predict numerous strip properties in hot rolling process, such as flow stress (Aghasafari, Abdi, & Salimi, 2014) and bending force (Wang, Gong, Li, Li, & Zhang, 2017), with high prediction accuracy.

Yan wu et al. proposed extreme learning machine (ELM) for the prediction bending force for control of strip profile during hot rolling (Wu, Ni, Li, Luan, & He, 2021). Results show a two-hidden layer optimized ELM model optimized by simulated annealing and genetic algorithm is the best candidate for rolling production. Jingyi Liu et al. previously used ELM optimized by genetic algorithm and particle swarm optimization for the prediction of rolling load for roughing and 7 passes of finishing on a dataset of 134 strips (Liu, Liu, & B. Le, 2019). 25% of data was used as test data. 30 features were used as inputs to predict the rolling loads for the 7 finishing passes. Input features were the width of incoming material, the thickness of incoming material, the entry and exit thickness of each roller, the entry and exit tension, the rolling temperature, and the rolling speed. R. Hwang et al. utilized deep neural networks to predict roll force of the Steckel Mill based on meta-features calculated from physics based equations (Hwang, Jo, Kim, & Hwang, 2020). Roll force for each pass was predicted using pass specific information and information from previous passes.

The application of ANN in rolling is therefore divided between hybridized models, where an ANN is used to train a compensation parameter in conventional mathematic models, and standalone ANN models completely separated from physics-based models, but in an isolated fashion (i.e. only variables at a single stand were predicted). The hybridized model, however, is far more restricted to changes in rolling conditions as the parameters in physics equations were predetermined and manually fitted over a different set of data. On the other hand, the standalone ANN model is able to fully utilize the process data and adapt to the setup of the studied rolling mill after learning. So far, few studies have been carried out to consider differences in rolling behavior at each pass and the effect of rolling histories. Simple network hyperparameter tuning is not sufficient in unveiling the hidden non-linear relationship between input features and the roll force.

## 3. Proposed method

Conventionally, physics-based metallurgical models tend to treat all rolling passes in the same way, without explicitly distinguishing the different levels of impact of the process parameters, which we hypothesize is a cause of various limitations. To individualize each rolling pass artificial neural networks (ANN) are selected to establish a coherent and accurate roll force prediction model. In ANN, weights are assigned between neurons, which dominate the contribution of each input to the final output, and are updated through exposing the ANN to the training data. In this study, a multi-layered feed-forward network was used for its capability of describing nonlinear regression problems and adjustable input–output mapping elements (Svozil, Kvasnicka, & Pospichal, 1997).

| | R1_Force | F1 Force | F2_Force | F3_Force | F4_Force | F5_Force | F6_Force |
|---|---|---|---|---|---|---|---|
| R1_Force | 1.000000 | 0.755391 | 0.825420 | 0.728427 | 0.671265 | 0.568995 | 0.520823 |
| F1 Force | 0.755391 | 1.000000 | 0.834844 | 0.670525 | 0.662059 | 0.561431 | 0.500945 |
| F2_Force | 0.825420 | 0.834844 | 1.000000 | 0.780816 | 0.768044 | 0.538665 | 0.493184 |
| F3_Force | 0.728427 | 0.670525 | 0.780816 | 1.000000 | 0.852524 | 0.550948 | 0.372477 |
| F4_Force | 0.671265 | 0.662059 | 0.768044 | 0.852524 | 1.000000 | 0.623473 | 0.451178 |
| F5_Force | 0.568995 | 0.561431 | 0.538665 | 0.550948 | 0.623473 | 1.000000 | 0.520344 |
| F6_Force | 0.520823 | 0.500945 | 0.493184 | 0.372477 | 0.451178 | 0.520344 | 1.000000 |

**Fig. 2.** Pearson correlation of roll forces.

Analysis of the roll force data shows that roll force in latter passes are related to the roll force in earlier passes, Fig. 2. Observing the correlation between stands it is apparent that for each stand correlation is not necessarily highest with those in its direct vicinity although this appears to be the prevalent case. Metallurgical theory suggests that roll force reflects a material's resistance to deformation and acts as a snapshot of the material's properties in time. We hypothesis that knowledge of material's previous conditions will improve the prediction of roll force at the subsequent pass. Observing a small subset of the training data, Fig. 3, where roll force at later stands increase or decrease in response to a change in rolling load at an earlier stand. Some more apparent instances of this are shaded in Fig. 3. This response should be material specific.

In the following subsections, three multi-step roll force prediction networks (MSN) are proposed: MSN (MSN-S) predicts roll force for each rolling pass based on the process parameters corresponding to that specific pass. MSN with feedbacks (MSN-F) uses the roll force value predicted for the previous pass as an input for prediction in the subsequent pass, acknowledging the effect of rolling history; MSN with all previous parameters (MSN-P) considers all process parameters related to the previous and current passes when making force prediction for the current pass. All three versions of MSN consider a specific model for each pass; MPN-F and MPN-P consider rolling history in their force predictions. Performance of MSN-S, MSN-F and MSN-P are shown on a real-world dataset and compared with the related mill model.

A list of features that are metallurgically recognized to contribute to roll forces are roll temperature, slab width, reduction percentage, travel time from previous stand, and roll speed; in addition, the most metallurgically significant alloying elements of steels (Siciliano & Jonas, 2000) were also used for the prediction: i.e. Carbon, Manganese (Mn), Silicon (Si), Aluminum (Al), Niobium (Nb), Titanium (Ti), Nitrogen (N), Molybdenum (Mo), and Vanadium (V).

**MSN-S:** It treats each rolling stand individually independent from other stands. A three-layer fully connected neural network is trained for each stand; the first layer takes features (described above) as input and the output layer provides an estimation for the required roll force at that specific pass. Due to the relatively small number of input features, one hidden layer of neurons is considered to incorporate nonlinear relationship between input and output. We propose to design and train individual networks for each stand — this is to reflect the potential differences at each stand.

**MSN-F:** Metallurgically, the hot deformation characteristics of the steel in the subsequent stands depend on the incoming microstructure, which are quantitatively reflected in the previous rolling load. MSN-F introduces the influence of the rolling history for prediction. It simulates the continuous process of hot rolling by connecting rolling stands together through feeding the predicted roll force from the previous stand as an extra input feature to the subsequent ANN corresponding to the next stand. A flowchart of MSN-F is shown in Fig. 4.

**MSN-P:** In MSN-F, when predicting the roll force for the current stand, we assume that all the relevant history is encoded in the previous stand's predicted force. In MSN-F, we go one step further and consider all the pass-specific features of the previous stands as rolling history representators and feed them as extra features when predicting the current stand roll force. The flowchart of MSN-P is shown in Fig. 5. By direct use of the raw information corresponding to the rolling history, any systematic prediction error that was contained by previous roll force and carried over between networks (see MSN-F) is eliminated, while maintaining the effect of rolling history. The input features of each pass now are composed of the fixed features, namely chemical composition and slab width, and the pass-specific features, that is, roll speed, slab travel time, reduction and temperature of the current and all previous passes. As is shown in Section Experimental Results, such long history inclusion when predicting roll forces, significantly improves prediction performance in the later stands.

## 4. Experimental results

Performance of the proposed expert system for roll force prediction is demonstrated on data recorded at Algoma Steel Inc. which is an integrated primary steel producer in Canada. The data are collected over 2 years of operation totaling 4460 effective entries across 7 different steel grades, ranging from V–N steels to Nb containing HSLA steels. Performance of the proposed MSN models is tested against the current Algoma mill model.

### 4.1. MSN implementation details

Every feature variable has been transformed to its z-score values. The normalized values mostly fall in the range of [−3,3]. Bayesian regularization is chosen as the backpropagation algorithm (Burden & Winkler, 2009). The cost function is mean squared error (MSE). A maximum of 250 epochs are allowed in training each network while the performance goal is $10^{-6}$. An 80-20 random split between training and test set is used. During training, the network weights are firstly randomly initialized and updated when the network is exposed to the training set. The number of hidden neurons for each pass is defined through 5-fold cross validation on the training data Activation function is set to Log-sigmoid. Each network consists of 20 hidden neurons and a single output neuron. The number of inputs varied with the individual networks for each MSN. 10 features are common to every pass and are included as input to every network while 4 features are pass specific, 2 of which are unavailable for the first pass.

### 4.2. Evaluation metrics

The predictive ability of model (PAM) is used in this study as the quantification metric for the prediction performance, which is defined by Eq. (1) (Poliak, Shim, Kim, & Choo, 1998):

$$PAM = \frac{NCP}{NP} \times 100\% \tag{1}$$

where NP is the number of total predictions and NCP is the number of correct predictions. NCP is defined as in Eq. (2) where $\hat{y}_j$ is the predicted roll force and $y_j$ is the true (i.e., final measured) roll force value.

$$\text{NCP} = \left| \frac{y_j - \hat{y}_j}{\hat{y}_j} \right| \leq 10\% \tag{2}$$

Any outlier that presents an error beyond ±10% would result in PAM to be less than 100%. To evaluate the predicted variance on a percentage basis, VAF (Variance accounted for) defined in Eq. (3) by (Babuska, 1998) is used, where $\boldsymbol{y}$ is the vector of measured values and $\hat{\boldsymbol{y}}$ includes the model predicted values. Predictions with less scattering lead to a higher value of VAF percentage.

$$VAF = \left[ 1 - \frac{var\,(\boldsymbol{y} - \hat{\boldsymbol{y}})}{var\,(\boldsymbol{y})} \right] \times 100 \tag{3}$$

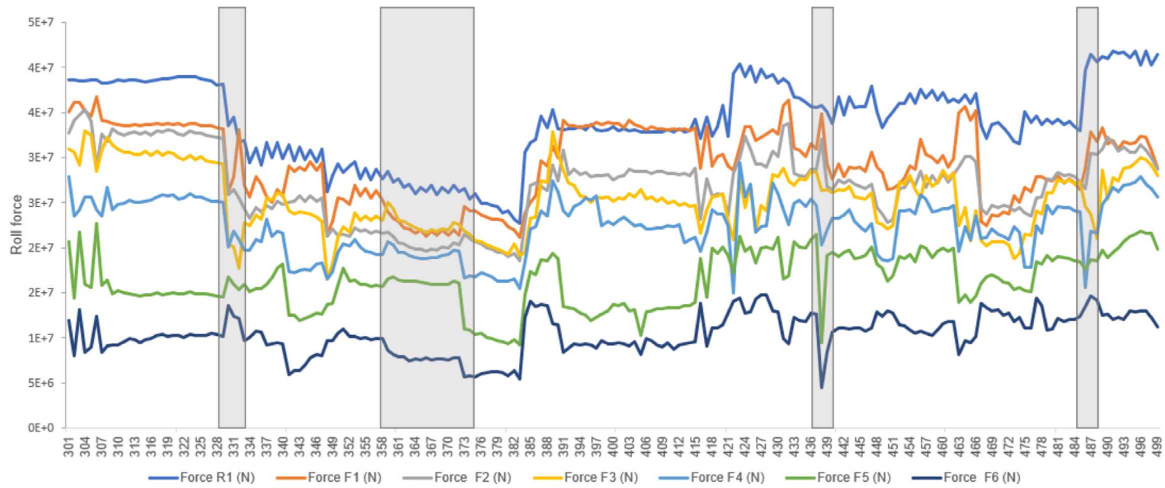


**Fig. 3.** Plot of 200 examples of roll force from the training data.

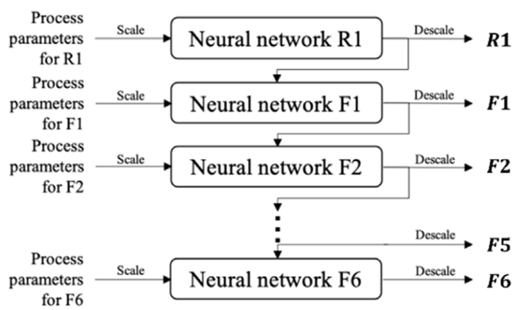


**Fig. 4.** Flowchart of connectivity of MSN-F where the predicted roll force of the previous stand is included for the subsequent prediction.

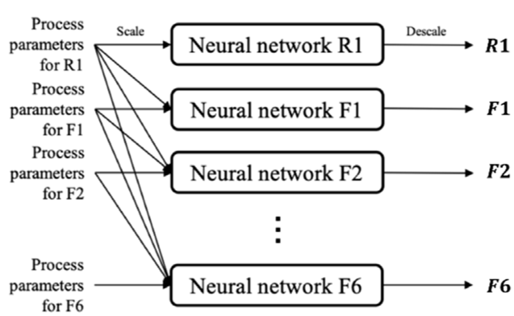


**Fig. 5.** Flowchart of the proposed MSN-P where both the current and previous rolling parameters make up the input features of each pass.

### 4.3. Algoma mill model

In the Algoma mill model, roll force is computed following the form described in Eq. (4) where $K_{\mathrm{m}}$ is material hardness without roll flattening, $P_{\mathrm{m}}$ is the roll flattening factor, $R_{\mathrm{m}}$ is the thickness reduction ratio in the present rolling stand and $\theta$ is a tuning parameter.

$$F_{mill} = K_m \times P_m \times R_m \times \theta, \tag{4}$$

The detailed analytical description of computation of $K_{\mathrm{m}}$, $P_{\mathrm{m}}$ and $R_{\mathrm{m}}$ are omitted here for confidentiality.

### 4.4. Prediction performance

4450 examples from the Algoma rolling mill are used to evaluate each method. PAM and VAF for the mill model and the proposed MSN-S and MSN-F, on the test data, at each pass are plotted in Figs 1 to 6. Both metrics for mill model show a similar trend where they plummet at F1, reaching 80.5% and 76%, respectively. They climb back up at F2 and peak at F3, followed by a decreasing performance at the final three passes. On the other hand, both proposed MSN-F and MSN-S models show a high test set performance in the first four passes, averaging 98% PAM and 94% VAF. However, a similar decrease also occurs for the final three passes, with a significant drop to 92% PAM and 82.9% VAF at F6. This inconsistency is a sign of error propagation and is potentially caused by a combination of input features with insufficient information to generalize the roll force. This is confirmed by the almost identical performances from the MSN-S, suggesting that previous roll force is not an adequate summarization of the rolling history. Despite the fact that the MSN-F and MSN-S models achieved a 9.5% net increase in PAM and 10.8% in VAF compared to the mill model, the large discrepancies in the final passes motivate MSN-P.

To quantify the superiority of the proposed MSN-P, the PAM and VAF percentages of all three models MSN-S, MSN-F and MSN-P are shown in Fig. 7.

Overall, the proposed MSN-P demonstrated a higher prediction accuracy and stability over the MSN-F and MSN-S, especially in the final three passes. This indicates the effectiveness of directly incorporating the raw previous rolling parameters. All three models performed well on F2 and F3. On the final three passes, further increase in both PAM and VAF from MSN-P is witnessed compared to a diminishing performance of MSN-S and MSN-F, which further confirmed the validity of representing previous rolling history directly via process parameters. An improvement of 6.5% in PAM and 13.6% VAF on the test set of the last pass F6 was found.

To visualize improved performance of the proposed MSN-P mode, extra analysis is conducted that compares predicted and measured (i.e., true) roll force for each of the 7 rolling passes. The results corresponding to the Algoma Mill model and MSN-P are respectively shown in Figs. 8 and 9.

In each of the graphs, the $x$-axis represents the measured roll force and $y$-axis is the roll force values predicted. The center line represents perfect predictability where the prediction would exactly match the measured values, with two lines above and below the center line being 10% error line. The regression prediction band is also plotted as an indicator of predictability, which is the line that best fits the predicted values.

Overall, the mill model prediction is highly scattered in all of the 7 passes except in Pass 4. Substantial underprediction is seen in Passes 1, 2 and 6 where the trendline largely falls below the center line. Specifically, in Pass 2, where the slab experiences a long transport time after the roughing mill, a significant variance and underestimation are found. In Pass 5 and Pass 7, a dramatic increase in outliers,

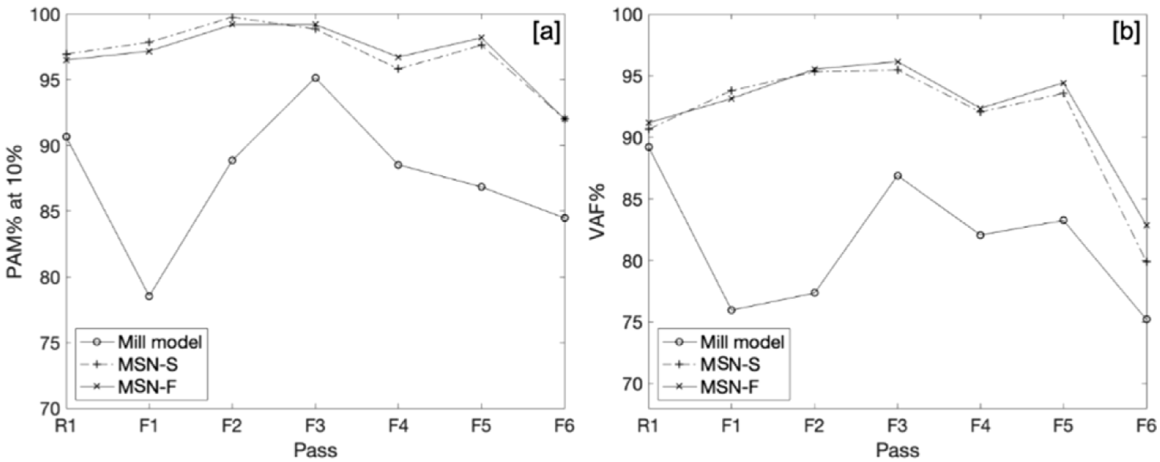


**Fig. 6.** [a] PAM percentage and [b] VAF percentage for mill model, MSN-F and MSN-S at each rolling pass on the unseen test set.

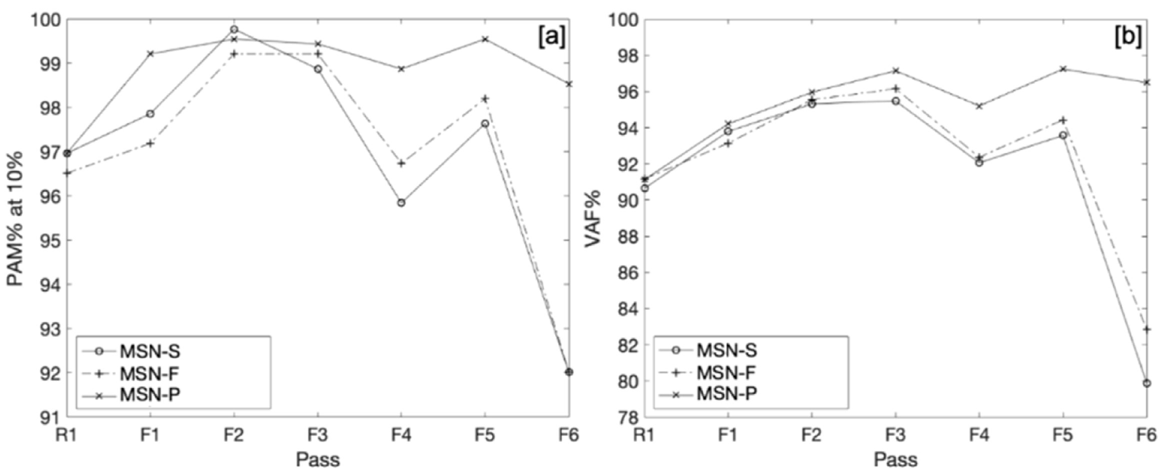


**Fig. 7.** [a] PAM percentage and [b] VAF percentage for all three neural networks model MSN-P, MSN-F and MSN-S at each rolling pass on the reserved test set.

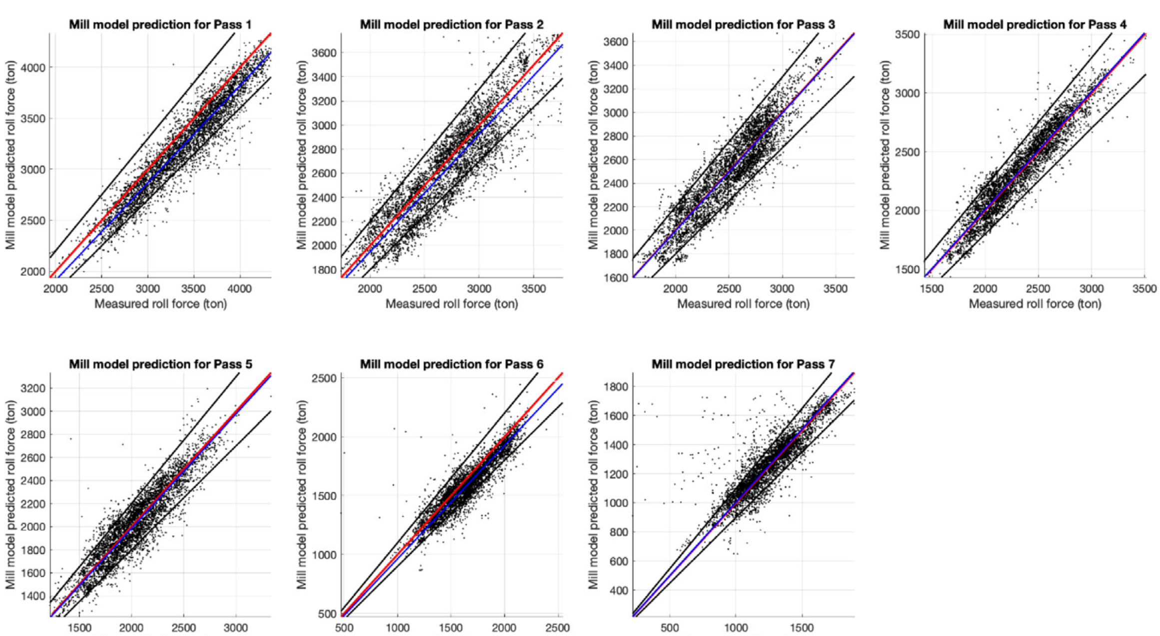


**Fig. 8.** Roll force predictability of mill model for each of the 7 rolling passes.

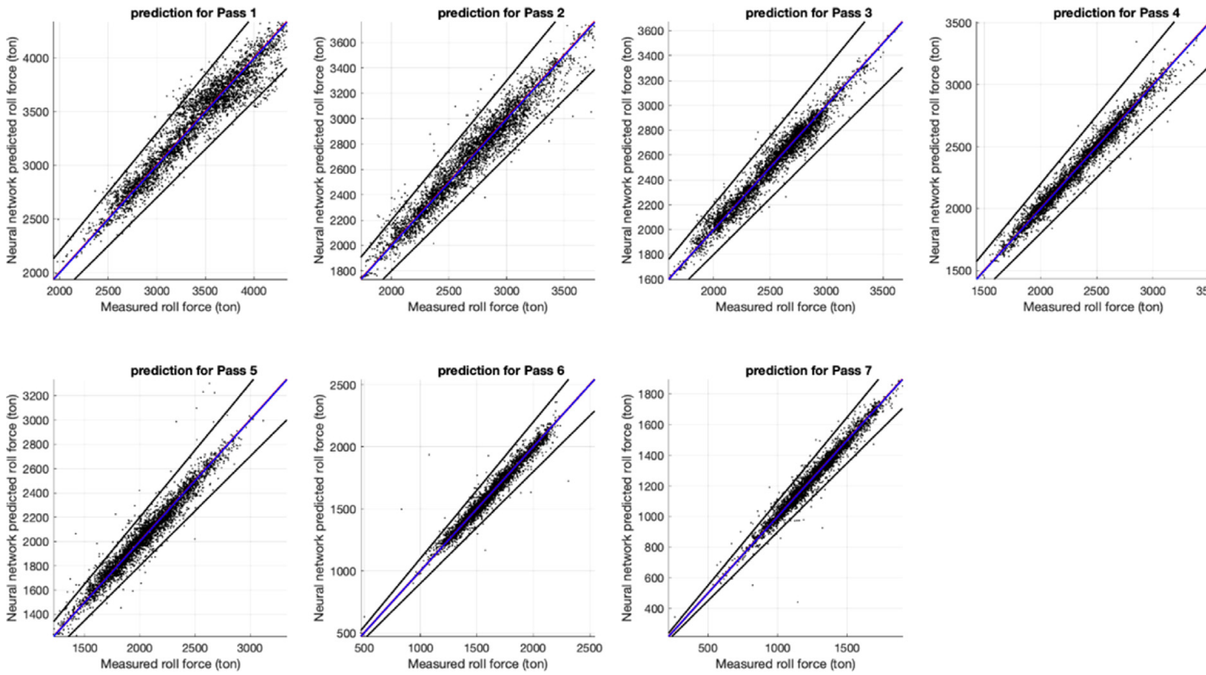


**Fig. 9.** Roll force predictability of MSN-P for each of the 7 rolling passes.

bearing more than 10% prediction error, is present, showing loss of uniformity. In general, the relatively unsatisfactory performance of the mill model stems from (1) lack of consideration of any metallurgical factors that could affect rolling (such as alloy composition), (2) using a common prediction model for all passes, (3) ignoring rolling history in predictions. The single tuning constant is insufficient to accommodate the effect that different rolling setups impose on the roll force, using physics-based formulae. Many equation parameters are fixed while they could be largely affected by the change of rolling conditions. These confirm that a more adaptable roll force prediction model is necessary.

Comparing Figs. 8 and 9 shows that all seven roll force values predicted by MSN-P achieved significant improvement over the mill model currently used in Algoma. On the test set, the mill model has a mean absolute error of 117.1 tons with standard deviation of 92.1 tons, compared to mean absolute error of 54.6 tons and standard deviation of 49.6 tons for the MSN-P, which translates to a net improvement of over 50% in both prediction accuracy and stability. Additionally, underestimation that occurred in the first two passes and instability in the final pass F6 in the mill model is largely prevented.

Roll force is in part a function of a material's response to deformation. From a metallurgical perspective it is reasonable that the prediction of roll force at latter passes would be more dependent on rolling history. The combination of processing conditions at each stand can result in full recrystallization, fractional recrystallization, or no recrystallization with the implication for the material's response to deformation being different in each case. During the rolling process, the temperature of the strip at each rolling pass decreases while the time between rolling stands simultaneously decreases. Under these conditions full recrystallization is possible at the earliest passes with the possibility for fractional recrystallization and no recrystallization increasing at later passes. The resistance to deformation of fully recrystallized austenite is a function temperature and strain at the current rolling pass. Under this condition MSN-S, MSN-F, and MSN-P should perform equally well and is observed in the prediction of R1, F2 and F3. However, in the absence of complete recrystallization the resistance to deformation is dependent on the unrelieved stress from previous passes as well as the temperature and strain. Information about unrelieved stress form earlier rolling passes is best provided by MSN-P and could explain the improved prediction in F4–F6.

These results confirm that treating each rolling pass differently and including rolling history contributes largely towards improving the roll force predictability. Using feedback loops as in MSN-F is not as effective as direct input of raw information as in MSN-P, in the form of propagated systematic error, and insufficient information to represent previous rolling history. The significant improvement on the final pass, simply by incorporating all previous process parameters, coincides with the complex nature of continuous rolling process and validates the importance of rolling history.

## 5. Conclusion

In this study, we proposed an expert system to construct a roll force prediction model to exceed the performance of currently used industry mill model. A multistep prediction model based on neural networks has been established including the effect of steel chemistry, physical rolling parameters and rolling histories for all rolling stands in an industry hot strip mill. The proposed model is trained using industry mill logs and its design is validated by comparing prediction performance with industry mill model. The proposed MSN-P showed a net improvement of 11.3% on prediction accuracy and 13.9% on stability, achieving an over 50% improvement compared to the mill model. The enhanced prediction accuracy largely helps with the mill setup to prevent excessive rolling load on the work rolls and a more precise shape control. This study successfully showed that standalone neural networks detached from physical model have great potential in process control and design of rolling schedules.

In this instance opportunities exist to improve the performance of the individual predictors. In the future this method will be augmented to incorporate complementary information about the materials' final mechanical properties to improve the prediction of rolling load. This is a good problem for multi-task learning where roll forces are a key process parameter for maintaining equipment integrity while mechanical properties are required product quality metrics. Applying the prediction of mechanical properties and roll force repeatedly provides a method to control the mechanical properties of the material while remaining within equipment constraints. This improved control of mechanical properties results in a higher value product.

## CRediT authorship contribution statement

**Shuhong Shen:** Methodology, Software, Validation, Formal analysis, Writing – Original Draft, Visualization. **Denzel Guye:** Methodology, Writing – Review & Editing. **Xiaoping Ma:** Resources, Writing – review & editing. **Stephen Yue:** Conceptualization, Writing – review & editing, Supervision. **Narges Armanfard:** Conceptualization, Methodology, Writing – review & editing, Supervision.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgments

The authors would like to thank Algoma Steel Inc., Canada for their financial and data support, as well as permission to publish this work. This research is generously funded by Mitacs Accelerate fund, Canada.

## References

Aghasafari, P., Abdi, H., & Salimi, M. (2014). Artificial neural network modeling of flow stress in hot rolling. *ISIJ International*, *54*(4), 872–879.

Babuska, R. (1998). Fuzzy modeling for control.

Burden, F., & Winkler, D. (2009). Bayesian regularization of neural networks. In D. J. Livingstone (Ed.), *Artificial neural networks: Methods and applications* (pp. 23–42). Totowa, NJ: Humana Press.

Cho, S., Cho, Y., & Yoon, S. (1997). Reliable roll force prediction in cold mill using multiple neural networks. *IEEE Transactions on Neural Networks*, *8*(4), 874–882.

Hodgson, P. D., & Gibbs, R. K. (1992). A mathematical model to predict the mechanical properties of hot rolled C-Mn and microalloyed steels. *ISIJ International*, *32*(12), 1329–1338. http://dx.doi.org/10.2355/isijinternational.32.1329.

Hwang, R., Jo, H., Kim, K. S., & Hwang, H. J. (2020). Hybrid model of mathematical and neural network formulations for rolling force and temperature prediction in hot rolling processes. *IEEE Access*, *8*, 153123–153133. http://dx.doi.org/10.1109/ACCESS.2020.3016725.

Larkiola, J., Myllykoski, P., Korhonen, A., & Cser, L. (1998). The role of neural networks in the optimisation of rolling processes. *Journal of Materials Processing Technology*, *80*, 16–23.

Lee, D. M., & Choi, S. G. (2004). Application of on-line adaptable neural network for the rolling force set-up of a plate mill. *Engineering Applications of Artificial Intelligence*, *17*(5), 557–565. http://dx.doi.org/10.1016/j.engappai.2004.03.008.

Liu, J., Liu, X., & B. Le, T. (2019). Rolling force prediction of hot rolling based on GA-MELM. *Complexity*, *2019*, Article e3476521. http://dx.doi.org/10.1155/2019/3476521.

Maccagno, T. M., Jonas, J. J., Yue, S., McCrady, B. J., Slobodian, R., & Deeks, D. (1994). Determination of recrystallization stop temperature from rolling mill logs and comparison with laboratory simulation results. *ISIJ International*, *34*(11), 917–922. http://dx.doi.org/10.2355/isijinternational.34.917.

Martinetz, T., Protzel, P., Gramckow, O., & Sörgel, G. (1995). *Neural network control for steel rolling mills*. London.

Niu, J. T., Sun, L. J., & Karjalainen, P. (2000). *Neural network-based model for prediction of hot-rolled austenite grain size and flow stress in microalloy steel, vol. 13*.

Öznergiz, E., Özsoy, C., Delice, I., & Kural, A. (2009). Comparison of empirical and neural network hot-rolling process models. *Proceedings of the Institution of Mechanical Engineers, Part B (Management and Engineering Manufacture)*, *223*(3), 305–312.

Poliak, E. I., Shim, M. K., Kim, G. S., & Choo, W. Y. (1998). Application of linear regression analysis in accuracy assessment of rolling force calculations. *Metals and Materials*, *4*(5), 1047–1056. http://dx.doi.org/10.1007/BF03025975.

Portmann, N. F., Lindhoff, D., Sorgel, G., & Gramckow, O. (1995). Application of neural networks in rolling mill automation. *Iron and Steel Engineer*, *72*(2), 33–36.

Siciliano, F., & Jonas, J. J. (2000). Mathematical modeling of the hot strip rolling of microalloyed nb, multiply-alloyed Cr-Mo, and plain C-Mn steels. *Metallurgical and Materials Transactions A*, *31*(2), 511–530.

Singh, Y., & Chauhan, A. S. (2009). Neural networks in data mining. *Journal of Theoretical & Applied Information Technology*, *5*(1).

Son, J.-S., Lee, D.-M., Kim, I.-S., & Choi, S.-G. (2005). A study on on-line learning neural network for prediction for rolling force in hot-rolling mill. *Journal of Materials Processing Technology*, *164-165*, 1612–1617. http://dx.doi.org/10.1016/j.jmatprotec.2005.01.009.

Svozil, D., Kvasnicka, V., & Pospichal, J. í. (1997). Introduction to multi-layer feed-forward neural networks. *Chemometrics and Intelligent Laboratory Systems*, *39*(1), 43–62. http://dx.doi.org/10.1016/S0169-7439(97)00061-0.

Wang, Z.-H., Gong, D.-Y., Li, X., Li, G.-T., & Zhang, D.-H. (2017). Prediction of bending force in the hot strip rolling process using artificial neural network and genetic algorithm (ANN-GA). *International Journal of Advanced Manufacturing Technology*, *93*(9–12), 3325–3338.

Wu, Y., Ni, H., Li, X., Luan, F., & He, Y. (2021). Prediction of bending force in the hot strip rolling process using multilayer extreme learning machine. *Mathematical Problems in Engineering*, *2021*, Article e6682660. http://dx.doi.org/10.1155/2021/6682660.

Yang, Y. Y., Linkens, D. A., & Talamantes-Silva, J. (2004). Roll load prediction—data collection, analysis and neural network modelling. *Journal of Materials Processing Technology*, *152*(3), 304–315. http://dx.doi.org/10.1016/j.jmatprotec.2004.04.372.