KES Transactions on Sustainable Design and Manufacturing I
Sustainable Design and Manufacturing 2014 : pp.744-756 : Paper sdm14-109

Suitability of multi-layer perceptron neural network model for the prediction of roll forces and motor powers in industrial hot rolling of high strength steels

D. W. James*, A. J. P. Gater, N. P. Lavery, J. Sienz

Swansea University, College of Engineering, Swansea, SA2 8PP, UK

*Corresponding author: Tel +441792 606879, Email: d.w.james@swansea.ac.uk

Abstract

Process data from a hot rolling mill is analysed to determine the variation of the load requirements of the mill stands with temperature for 13 high strength steel grades. The relationship is shown to be highly non-linear making prediction of rolling loads very difficult. Feed forward multi-layer perceptron (MLP) neural networks are suggested as a means of prediction.

A methodology is presented for the generation of a system of MLP neural networks for offline prediction of required rolling forces and also motor powers during finishing rolling of a range of high strength steels. The networks are trained using approximately 12000 coils worth of process data, and prediction errors are less than 10% for load and power in over 90% of the data for most of the stands measured.

1. Introduction

In recent years there has been an increasing adoption of advanced high strength steels within the automotive sector [1]. These steels, such as dual phase, complex phase and boron added hot forming grades, require richer chemistries than traditional high strength low alloy (HSLA) or low alloy steels. Typical alloying additions include niobium, manganese, chromium, silicon, titanium and boron. These steels exhibit higher flow stresses at elevated temperatures, resulting in greater power and force requirements during processing.

The final process in production of high strength steels is hot rolling where the gauge of a strip is reduced to a specification required for market. The gauge is reduced by passing the heated strip through a series of finishing stands in a hot rolling mill. In each finishing stand the gauge of the steel strip is sequentially reduced by passing it through a set of rolls. Typically the strip enters the rolling mill with a thickness of 30–40mm which is reduced to a finishing gauge of 0.8-2.2mm (when cooled).

To optimise production it is desirable to maximise the width of the strip and speed at which it moves through the mill. These parameters are limited by two key factors – the roll separating force exerted by the strip on the rolls during processing and the power requirement of the roll motors. The roll separating force, forces the rolls apart reducing the reduction possible by each stand. Hydraulic presses are used to counter act the roll separation force so the reduction is limited by the load they can produce. The speed at which the strip can move through the mill is limited by the motor power.

In the development of new high strength steels and processing of existing grades it would be a major advantage to be able to predict rolling loads and powers required of the mill prior to production. Historically, a large volume of work has been undertaken, producing analytical relationships between flow stress and roll force. Orowan and Ing [2] developed a numerical method to calculate the distribution of pressure over the arc of contact between the roll and the strip. Local roll deformation is treated as a 2D plastic deformation. Sims [3] uses and extends the methods developed by Orowan and Ing [2] to calculate normal roll pressure, roll load and torque in hot rolling mills. A simplified version of Sims [3] model which gives similar results was presented by Ford and Alexander [4]. Attempts have also been made at predicting flow stress and rolling forces as a function of chemistry and processing history through a combination of numerical models, constitutive equations finite element analysis and empirical models. Misaka and Yoshimoto [5] develop an expression for mean flow stress by considering the effects of strain, strain rate and temperature independently. A semi-empirical model developed by Siciliano and Jonas [6] to calculate flow stresses in the strip based on Sims equations and data from a seven stand rolling mill for microalloyed Nb, multiply-Alloyed Cr-Mo, and plain C-Mn steels. In a following work Siciliano and Poliak [7] extend this model to include alloying and microalloying effects. Constitutive equations based on an idealized hot rolling process are used in a finite element model by Rudkins and Evans [8]. The constitutive equations were based on experimental plane strain compression tests. Serajzadeh [9] predicts the flow behaviour of rolled steel using on a 2D finite element model and a first order rate equation. The predicted roll forces are comparable with experimental data.

Due to the complexity of the hot rolling process, analytical treatments must make a number of assumptions regarding factors such as surface friction, inter-stand cooling, adiabatic heating in the roll bite and roll flattening. Once microstructure is incorporated a large volume of experimental data must also be produced to support the analysis of rates of recrystallization, grain growth, precipitate nucleation and growth, and the impact of grain size or dislocation density on flow stress. All of these factors are influenced by the composition of the steel. Bakkalogu [10] shows that the strength of HSLA steels can be improved by controlled rolling in the two-phase region. The increased strength is shown to be related to changes in microstructure. Jeong [11] demonstrates that hot rolling in the ferrite or austenite region affects the Lankford parameter of ultra-low carbon Ti-interstitial-free steel.

An alternative technique to analytical models is the use of neural network models for prediction and control. Neural networks are able to fit highly complex nonlinear functions and have been applied in the process automation of more than 40 rolling mills worldwide [12]. They have been utilized by a number of researchers for the prediction of rolling forces, in combination with other optimisation techniques such as genetic algorithms, on-line adaptive-control, or in a combinatorial fashion for the correction of mathematical models. Schlang et al [12] reviews the use of neural networks installed on steel mills worldwide. They are currently used to predict required rolling loads from previous process data. The author that development of artificial intelligence techniques will allow all aspects of steel production to be automated resulting in higher plant efficiencies. In an earlier work Schlang [13] describes a combined analytical/ neural network model where the analytical models are used as set points in the neural network training data. Son et al [14] use neural networks to predict the rolling force required in a hot rolling mill. The training data for the network is based on analytical models. A genetic algorithm is used to find an optimum network architecture. In a preceding paper Son et al [15] assess the practicality of training neural networks during hot mill operation.

The authors are only aware of two previous papers predicting power utilization in shape rolling, and by extension flat rolling, both produced by Behzadipour et al [16], [17] but the volume of experimental data in those studies was low.

In this work, rolling forces and motor powers for a seven stand four high finishing mill are predicted using neural networks. The geometry of the mill is shown in Figure 1 .These results are compared with measured industrial process data from a 7 stand 4 roll high hot rolling mill. The industrial process data covers approximately 12000 coils of rolled steel and 13 steel grades.

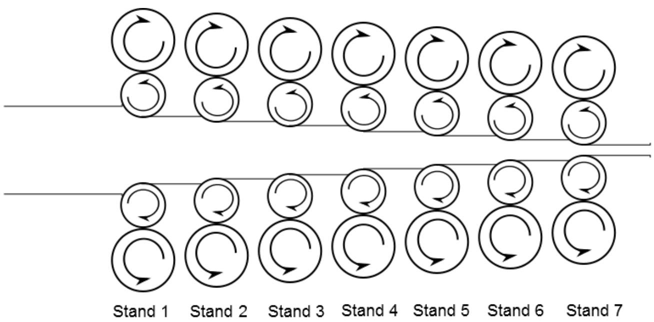

Figure 1 Mill geometry

In the first part of this work flow stress is calculated for 13 steel grades during reduction as a function of the measured roll separating force and reduction. By comparing the calculated mean flow stress in each stand with the mean temperature of strip moving through the stand non-linear behaviour is observed. This is shown to be consistent with dynamic recrystallization and accumulated strain effects [18].

In the second part of this work a methodology is described for producing a feed forward multi-layer perceptron (MLP) neural network model for offline prediction of rolling forces and motor powers. The networks are trained using the process data from all 13 steel grades provided and a validity of the predictions discussed.

2. Flow stress calculations

The roll separating force is dependent on the mean flow stress (MFS) of the steel. Following the work of Ford and Alexander [4] the MFS can be calculated from the mean shear yield stress (MSYS). The following expression is used to calculate the MSYS in a coil undergoing reduction:

$$MSYS = \frac{P}{\left[\sqrt{R'\Delta h}\right]\left[1.57 + \frac{\sqrt{R'\Delta h}}{h_0 + h_f}\right]} \qquad 1)$$

where P is the roll separating force(N), h_0 is the strip gauge entering the stand (m), h_f is the strip gauge exiting the stand (m), Δh is the gauge difference or draught (m) and R' is the deformed roll radius (m). The deformed roll radius is calculated using equation 2 derived by Hitchcock [19]

$$MSYS = \quad R' = R\left[1 + \frac{CP}{b(h_0 - h_f)}\right] \qquad 2)$$

where R is the original roll radius (m), b is the width of the strip (m) and C is given by

$$C = \frac{16(1 - v^2)}{\pi E} \qquad 3)$$

Where, E and v are the Young's modulus (Pa) and Poisson's ratio of the steel respectively. Finally the mean flow stress (MFS) is calculated by

$$MFS = \sqrt{3MSYS} \qquad 4)$$

Using these equations we have calculated the MFS for each of the samples in the process data based on the measured roll separation forces for P and assuming a value of E of 210GPa and v of 0.3. All other parameters are taken directly from the process data.

In a static recrystallization process a inverse linear relationship is expected between MFS and temperature is expected. To see if other recrystallization effects occur, the MFS was plotted against 1000/T for the 13 grades, as shown in Figure 2. The temperature of the steel entering and exiting the hot mill are recorded in the process data. It is assumed that the temperature of the steel slab decreases linearly as it goes through the rolling mill. This assumption is an over simplification but will provide qualitative information about the recrystallization effects of the process.

Figure 2 shows the mean flow stress vs 1000/T for a single steel grade. The results are plotted in 20K temperature bins for clarity.

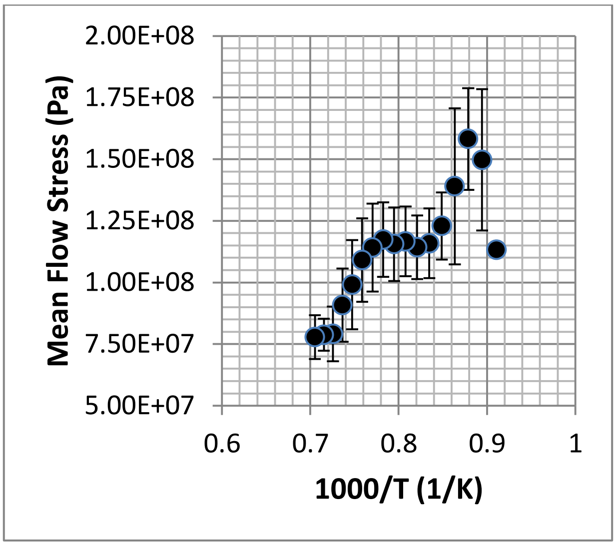

Figure 2 Mean flow stress vs 1000/T for a single steel grade

It can be seen that the flow stress does not increase linearly with inverse of temperature as would be expected for static recrystallisation (SRX). This is due to strain accumulation and dynamic recrystallisation, as demonstrated in Figure 3.

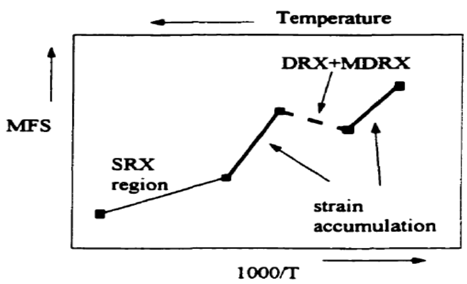

Figure 3 Schematic of partial recrystallization and strain accumulation in hot rolling [18]

At low temperatures (corresponding with the end of the hot mill) static recrystallisation takes place. There are two strain accumulation regions where the gradient of MFS vs 1000/T increases. Between these regions there is dynamic recrystallization (DRX) region where the MFS decreases with inverse T. The steel grade in Figure 3 exhibits all of these characteristics showing that the MFS of the strip as it passes through the hot mill is highly non-linear. Other grades examined in this study exhibit similar non-linearity. The point at which the recrystallisation changes regime will be dependent on the chemistry and processing conditions. This makes the prediction of roll separating loads extremely difficult.

Neural networks can be used as function fitting algorithm and are able to fit highly non-linear functions. A data set with known inputs and outputs is used to train the network to 'recognise' patterns between the input and output data. The network can then be used to predict outputs from a set of inputs that it has not seen before (the unseen input data should be of similar type and range as the training data). Given the complexity of roll load prediction, and the large amount of process data available, neural networks are an ideal tool to solve this problem.

In the next section the suitability of multi-layer perceptron neural networks as a tool to predict hot mill rolling loads and also the hot mill motor powers is presented.

3. Neural network parameters and inputs

Feed forward multi-layer perceptron neural networks were used throughout this study, utilising a tan-sigmoid transfer function at all hidden nodes and a linear transfer function on output nodes. All generated networks were fully connected, and trained in MatLab [20] using the Levenberg-Marquette back propagation algorithm. Inputs and targets in the data set were normalised to have zero mean and unity variance.

Early stopping was applied in the training of all of the investigated networks to reduce the risk of overfitting. The data was split into three sets, training, validation and testing sets. The network was trained using the training set. During each training epoch the network was exposed to the validation set and the MSE recorded. Typically at a certain epoch the MSE of the training set will continue to decrease, while the MSE of the validation set will stop decreasing and begin to rise, indicating overfitting. The training is stopped at this point and the network with the lowest validation set error is used. Finally the network was tested using the test set. In this work a data split of 70%, 15%, 15% of the existing data was used between the training, validation and testing sets respectively.

Network inputs were chosen from the full array of logged hot mill parameters, and were selected based on process experience and a consideration of the relevant parameters incorporated into analytical treatments on rolling. In the following, the terms F1 through F7 refer to the seven stands of the finishing train.

The intent of this study was to develop fourteen neural networks in total, one each per stand for rolling load and motor power respectively. The input parameters common to all of the stands are:

- Weight percentage of each alloying element in the steel grade (Carbon, Magnesium, Silicon Niobium, Titanium, and Nitrogen)

- Slab width
- Slab thickness
- Temperature of the slab entering the rolling mill
- Temperature of the slab exiting the mill

Parameters that are not common to all the networks are:

- The tangential velocity of the slab entering the stand
- Exit gauge of the stand
- Fractional reduction of the slab after going through the stand

Each network takes these values into account for the stand whose power and force it is predicting and for all stands preceding it.

For each finishing train stand the targets for the load applied to the steel strip (tons) and power required to drive the rollers (% of rated motor capacity) were taken from average data.

4. Development of the Neural Network Architecture

Three network paradigms were investigated in this study, utilizing five network configurations. The inputs to the networks are stand dependent and detailed in table 1.

1) Neural networks possessing a single hidden layer
2) Neural networks possessing two hidden layers, in which the first hidden layer maintains a constant size of either 2 or 5 neurons.
3) Neural networks possessing two hidden layers, in which the second layer maintains a constant size of either 2 or 5 neurons.

These architectures were chosen to see if the prediction error could be reduced by including an additional hidden layer and how it was affected by varying the number of neurons in each of the layers. Ideally a neural network should contain the minimum number of degrees of freedom possible, while sufficiently fitting the system to be modelled. In order to achieve this, a systematic search routine was used to find an approximately optimum architecture with the number of nodes in the network's hidden layers being varied. The number of nodes in the variable hidden layer was increased from two to fifty in steps of four. For each network structure, training was undertaken 25 times, and the average mean squared error in the test set across the 25 trainings was recorded.

While the Means Square Error (MSE) in the training data continued to fall with increasing numbers of hidden nodes, a state of diminishing returns was observed with respect to the error in the test set, beyond which any further increase in the number of hidden nodes resulted in a negligible reduction in the average MSE. The beginning of this plateau was chosen as an 'optimal' network configuration, given full connectivity. An example of plateauing of the MSE of the test and validation data with number of hidden nodes is shown in Figure 4.

A methodology has been presented by which a family of MLP neural networks can be generated to successfully describe the rolling forces and power requirements in the processing of a range of high strength steels. The level of prediction accuracy is reasonable given that average input and target values are used in training the networks and the variation in the load for a given coil can be of the order of hundreds of tons. While neural networks cannot be used to extrapolate, they can be used in an interpolative fashion provided measures are taken to ensure good generalisation, such as early stopping during training. The methodology presented in this paper could be used to guide the design of novel steel grades, which have chemistries falling within the range of the input parameters.

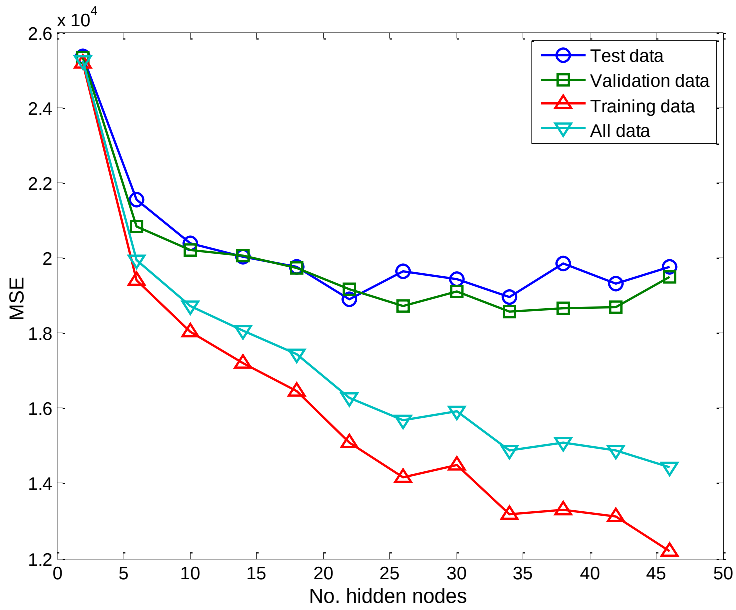

Figure 4 MSE of neural network predictions of load for a single hidden layer network against number of hidden nodes for a single stand

Once the optimum range of hidden nodes had been determined via the coarse stepping outlined above, the methodology was repeated in the range (n-3) to (n+3), where n is the number of hidden nodes in the best performing network during the first training round. A pseudo code of the training regime for all the networks considered is shown in Figure 5.

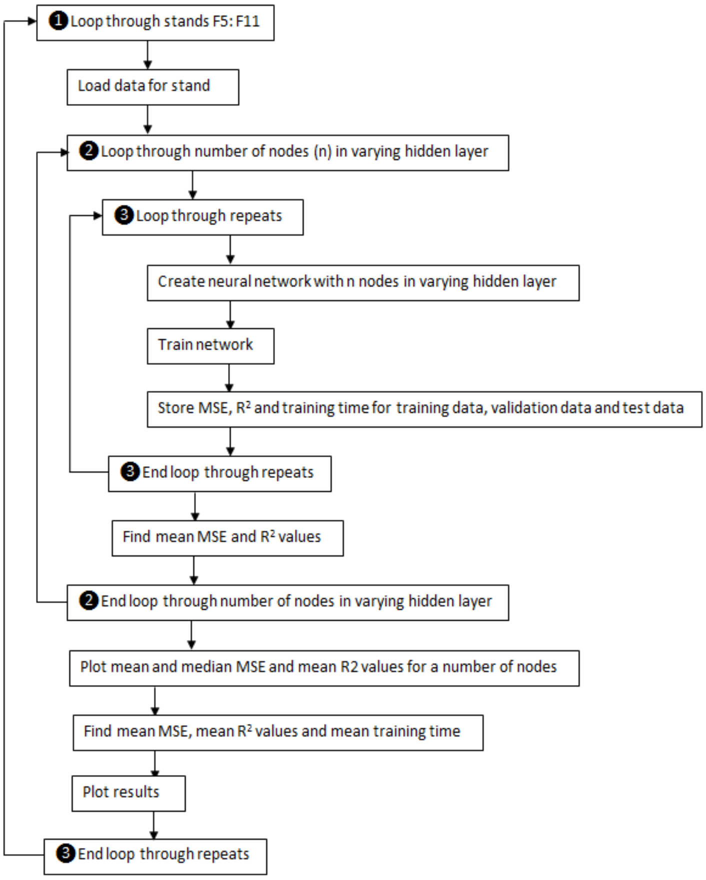

Figure 5 Neural network training routine

For each of the five network architectures considered, the network having the lowest MSE from the 25 training attempts was kept, and plots made of the MSE vs. stand number to identify the best architecture for this study (Figures 6 and 7).

Across all stands the networks utilising a single hidden layer were found to perform as well as, or better than, the two hidden layer architectures. Additionally the single hidden layer networks had a lower number of total degrees of freedom, considering all weights and biases.

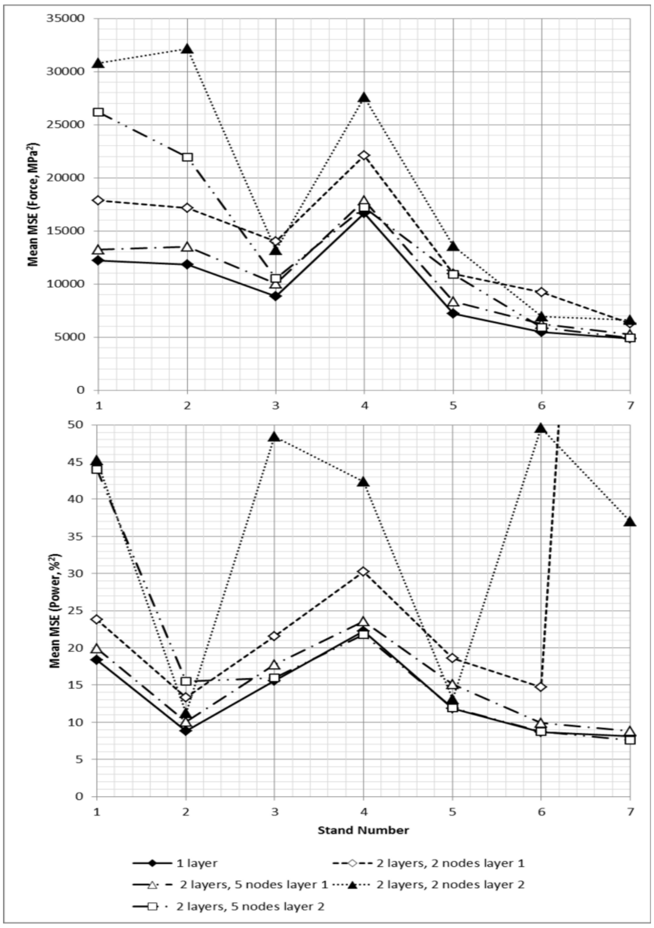

Figure 6 Percentage error of neural network predictions of power and load against measured values for different neural network architectures

5. Assessment of the model

Performance of the models has been assessed using a variation on "Predicted Ability of Model" (PAM) [15], where the PAM describes the percentage of predictions made falling within a given error range.

$$PAM = \frac{No\ correct\ predictions\ within\ x\%\ of\ target}{Total\ number\ of\ predictions} \times 100\%$$

Each network generated is run using the input parameters from the measured data. The loads and powers predicted are compared with the measured values. Using the PAM procedure confidence levels within 5%, 10% and 15% have been assessed. More than 95% of all predictions relating to the test set have been found to fall within 15% of the measured values, and typically greater than 90% of all predictions within 10% as shown in Figure 7.

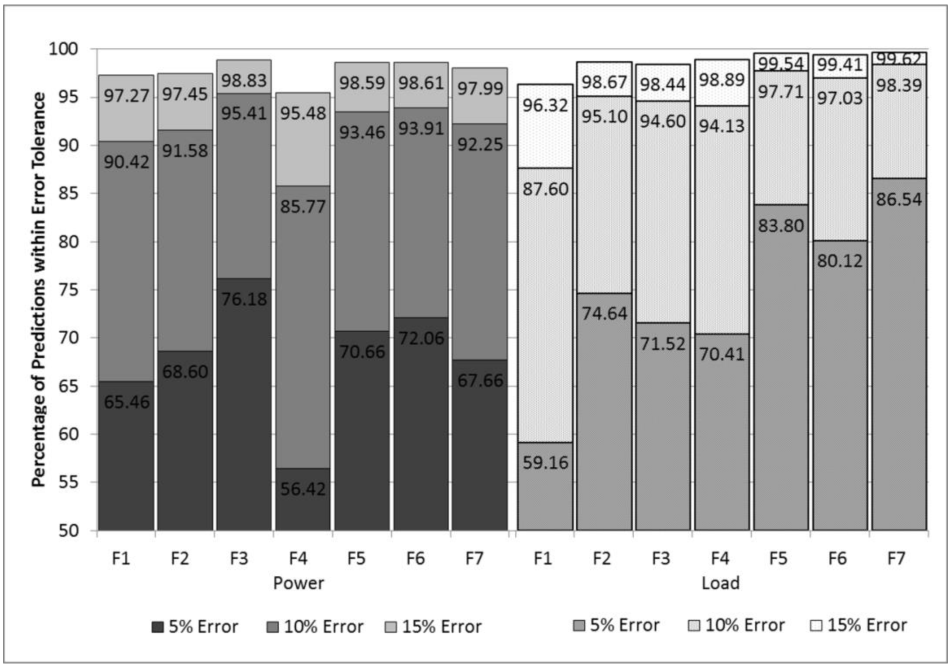

Figure 7 Percentage of neural network predictions for load and power within confidence levels

6. Conclusion and Discussion

The load and power requirements of a hot rolling mill are dependent on MFS of the steel being rolled. The MFS as a function of temperature is highly non-linear. In this work this non-linear behaviour from actual plant data has been shown.

Neural networks have been suggested as a tool to predict the load and power requirements of the hot rolling mill. Given a good resource of historical plant data it has been demonstrated that MLP neural networks can be used to successfully describe the rolling forces and power requirements in the processing of a range of high strength steels. The level of prediction accuracy is reasonable given that average input and target values are used in training the networks and the variation in the load for a given coil can be of the order of hundreds of tons. While neural networks cannot be used to extrapolate, they can be used in an interpolative fashion provided measures are taken to ensure good generalisation, such as early stopping during training.

In addition we have described a methodology which could be used to generate and train suitable neural networks from hot rolling mill data. The generated networks can then be used as a guide in the design of novel steel grades, which have chemistries falling within the range of the input parameters of the training data.

Acknowledgements

The work described in this paper was carried out as part of the Advanced Sustainable Manufacturing Technologies (ASTUTE) project (ref. numb. 80380) in collaboration with Tata steel, and the authors would particularly like to thank Richard Underhill and Douglas Gordon for their inputs. ASTUTE has been part-funded by the European Regional Development Fund through the Welsh Government, and the authors would like to acknowledge this funding.

8. Works Cited

[1] M. D. Tumuluru, “Resistance Spot Welding of Coated High-Strength Dual-Phase Steels,” *Welding Journal,* no. August, pp. 31-37, 2006.

[2] E. Orowan and Ing, “The Calculation of Roll Pressure in Hot and Cold Flat Rolling,” *Proc. Inst. Mech. Eng,* vol. 150, pp. 140-167, 1943.

[3] R. B. Sims, “The calculation of Roll Forces and Torque on Hot Rolling Mill,” *Proc AIME,* vol. 168, no. 6, pp. 191-200, 1954.

[4] H. Ford and J. Alexander, “Simplified hot rolling calculations,” *J Inst Met,* vol. 92, pp. 397-404, 1963.

[5] Y. Misaka and T. Yoshimoto, “Formularization of Mean Resistance to Deformation of Plain Carbon Steels at Elevated Temperature,” *J. Jpn. Soc. Technol. Plast,* vol. 8, pp. 414-422, 1967.

[6] F. Siciliano and J. Jonas, “Mathematical Modeling of the Hot Strip Rolling of Microalloyed Nb, Multiply-Alloyed Cr-Mo, and Plain C-Mn Steels,” *Metallurgical and Materials Transactions A,* vol. 31a, no. February, pp. 511-530, 2000.

[7] F. Siciliano and E. I. Poliak, “Modeling of the Resistance to Hot Deformation and the Effects of Microalloying in High-Al Steels under Industrial Conditions,” *Materials Science Forum,* Vols. 500-501, pp. 195-202, 2005.

[8] N. Rudkins and P. Evans, “Mathematical modelling of mill set-up in hot strip rolling of high strength steels,” 1998.

[9] S. Serajzadeh, “A mathematical model for evolution of flow stress during hot deformation,” *Materials Letters,* vol. 59, pp. 3319-3324, 2005.

[10] A. Bakkaloglu, “Effect of processing parameters on the microstructure and properties of an Nb microalloyed steel,” *Materials Letters,* vol. 56, pp. 200-209, 2002.

[11] W. C. Jeong, “Effect of hot-rolling temperature on microstructure and texture of an ultra-low carbon Ti-interstitial-free steel,” *Materials Letters,* vol. 62, pp. 91-94, 2008.

[12] M. Schlang, B. Lang, T. Poppe, T. Runkler and K. Weinzierls, “Current and future development in neural computation in steel processing,” *Control Engineering Practice,* vol. 9, no. 9, pp. 975-986, 2001.

[13] M. Schlang, B. Feldkeller, B. Lang, T. Poppe and T. Runkler, “Neural Computation in Steel Industry,” in *Proceedings of the European control conference '99*, Karlsruhe, 1999.

[14] J. S. Son, D. M. Lee, I. S. Kim and S. K. Choi, “A study on genetic algorithm to select architecture of a optimal neural network in the hot rolling process,” *Journal of Materials Processing Technology,* Vols. 153-154, pp. 643-648, 2004.

[15] J.-S. Son, D.-M. Lee, I.-S. Kim and S.-G. Choi, “A study on on-line learning neural network for prediction for rolling force in hot rolling mill,” *Journal of Materials Processing Technology,* Vols. 164-165, pp. 1612-1617, 2005.

[16] S. Behzadipour, A. Khajepour and J. G. Lenard, “Power estimation in shaoe rolling using nueral networks,” in *43rd Mechanical Working and Steel Processing Conference*, Charlotte, USA, 2001.

[17] S. Behzadipour, A. Khajepour, J. G. Lenard and J. Biglou, “A new shape change quantification for estimation of power in shape rolling,” *Journal of Materials Processing Technology,* vol. 148, no. 3, pp. 353-361, 2004.

[18] F. Siciliano, Mathematical Modeling of the Hot Strip Rolling of Nb Microalloyed Steels- PhD Thesis, Montreal: McGill University, 1999.

[19] J. H. Hitchcock, “Roll neck bearings,” ASME Res. Committee, 1935.

[20] MATLAB, version 7.12, Natick, Massachusetts: The Mathworks Inc, 2011.