JFE TECHNICAL REPORT
No. 26 (Mar. 2021)

Application of Machine Learning to Tandem Cold Mill Setup

YAMAZAKI Tatsuya[*1] MUKAIYAMA Akihiko[*2] YAJIMA Masahide[*3]

Abstract:

In response to increased demand for steel sheets that are difficult to roll (e.g. high-tensile strength steel sheets for automobiles) in recent years, improved accuracy has also been required in tandem cold mill setup systems. Therefore, a newly-developed setup system using machine learning was applied to tandem cold mills at JFE Steel Corporation's West Japan Works. This system includes multiple predictive models, such as draft setting, tension setting, rolling force, rolling torque, and forward slip models, and trains these models using a large volume of actual operational data. Improvement of the accuracy of each predictive model contributes to increased productivity.

1. Introduction

As the features of tandem mills, although the strip thickness of the base material and the target thickness in the product specification are given, there is a degree of freedom in the strip thickness at each mill stand, and the rolling force, rolling torque and lubrication condition of each stand affect the steel strip quality, i.e. the thickness, shape and surface appearance at each stand. Due to these features, various setup methods have been proposed for tandem mills, for example, setting the reduction ratio and the inter-stand tension of each stand [1–3).

In the continuous tandem mill, which is now the mainstream technology, flying gauge change is generally used to improve productivity. This is a technique in which the strip is rolled at low speed without a rolling stop when passing a weld point between coils, and the roll gap and roll speed of each stand are changed by feed-forward control when the weld point passes through the stand. Then, after completion of weld point passing, the roll gap and roll speed are controlled by feedback as the operation end of automatic thickness control or automatic tension control. Practical application of such advanced rolling control technologies today is largely due to advances in rolling theory. Specifically, the amount of roll gap change during weld point passing is calculated from the model prediction value of the rolling force and the estimated mill modulus, and the amount of roll speed change is calculated from the setting value of the rolling reduction of each stand and the model prediction value of forward slip.

From the standpoint of productivity, prediction of rolling torque is important because the rolling power for plastic working of the material is supplied by the rolling roll drive motors, and the achievable speed of the rolling roll is limited to the range in which rolling power (depending on rolling torque) does not exceed the rated output of the mill motors. Hence, the estimated maximum roll speed of a tandem mill is obtained from the prediction value of rolling torque, and is limited to the speed at which the predicted output of each motor is less than its rated output in all stands.

As mentioned above, various model calculations based on rolling theory are carried out in the tandem mill setup procedure in addition to setting reduction

[†] Originally published in *JFE GIHO* No. 45 (Feb. 2020), p. 48–53

[*2] Staff Assistant Manager,
Cold Rolling Technology Sec.,
Cold Rolling Dept.,
West Japan Works (Fukuyama),
JFE Steel

[*1] Staff Assistant Manager,
Cold Rolling Plant Control Sec.,
Plant Control Dept.,
West Japan Works (Fukuyama),
(currently, Staff Deputy Manager, Facilities Planning Sec., Planning Dept., West Japan Works)
JFE Steel

[*3] Assistant Manager,
Cold Rolling Plant,
Cold Rolling Dept.,
West Japan Works (Fukuyama),
JFE Steel

and tension, and the quality of those calculations greatly affects product quality and production efficiency.

Due to the recent increase in demand for difficult-to-roll materials (e.g. high-tensile strength steel sheets for automobiles and electrical steel sheets for electrical equipment), the variety of tandem mill products has diversified, making it necessary to upgrade setup techniques and predictive models for new varieties of products.

This paper describes a tandem mill setup system using machine learning which has been introduced in the cold rolling mills of JFE Steel Corporation's West Japan Works. This system is based on big data from past operation records and includes predictive models of the rolling force, rolling torque, forward slip and suitable setup values of reduction and tension. The system also automatically collects and accumulates operation records in its database and uses those records in training the above-mentioned models.

2. Conventional Tandem Mill Setup System

2.1 Setup Process

This section presents the outline of a conventional tandem mill setup process. **Fig. 1** shows the general flow of cold mill setup.

2.1.1 Setting of reduction and tension

First, the rolling reduction and tension between stands are set on the basis of the given manufacturing conditions (e.g. product thickness, width, type of steel, and rolling roll information).

To date, many setting methods have been proposed from various viewpoints. For example, methods for the purpose of productivity improvement include giving setting values from a table or maximizing the output of the rolling roll drive motors[1)]. On the other hand, as methods for operation stabilization and quality improvement, a method of minimizing an evaluation function consisting of prediction values for rolling force and motor output[2)] and a method using a neural network[3)] have been proposed.

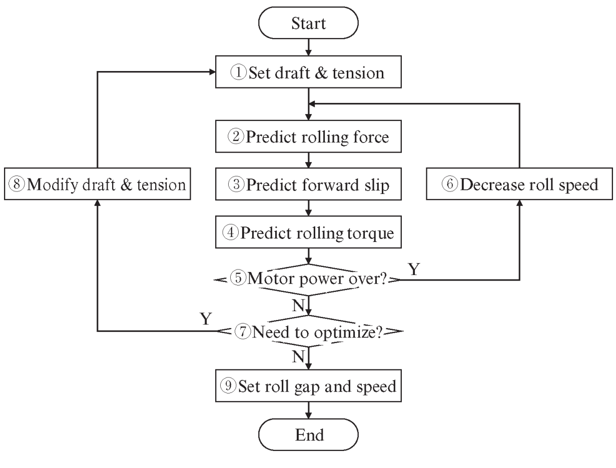

Fig. 1 Tandem cold mill setup flow chart

2.1.2 Rolling force, rolling torque and forward slip[4)]

When the reduction ratio and tension between stands are determined by the procedure in the preceding paragraph, the rolling force, rolling torque and forward slip can be predicted based on these setup values.

This section describes the prediction calculation using a model (hereinafter referred to as physical model) based on rolling theory. The physical model gives an approximate solution which is obtained by simultaneously solving the balance of the force in the roll bite and the elastic deformation of the rolling rolls.

The following describes Bland and Ford's rolling theory and Hitchcock's roll flattening model, which are widely used in cold rolling.

First, in Bland and Ford's theory, the rolling pressure can be expressed by the following equation by approximately solving the balance equation of force in the roll bite:

$$\left.\begin{aligned} p &= \begin{cases} \dfrac{kh}{h_0}\left(1-\dfrac{\sigma_0}{k_0}\right)e^{\mu H} \\[2ex] \dfrac{kh}{h_1}\left(1-\dfrac{\sigma_1}{k_1}\right)e^{\mu(H_1-H)} \end{cases} \\ H &= 2\sqrt{\dfrac{R'}{h_0}}\tan^{-1}\left(\sqrt{\dfrac{R'}{h_0}}\phi\right) \end{aligned}\right\} \quad \cdots\cdots (1)$$

where k is the deformation resistance, h is the strip thickness, σ is the tensile stress, μ is the coefficient of friction, R' is the radius of the flattened roll by Hitchcock and ϕ is the angle of the polar coordinate system as seen from the center of the flat roll, where "0" and "1" indicate the roll bite outlet and inlet, respectively.

For the deformation resistance k in cold rolling, the general method is to use a nonlinear regression equation obtained through experiments.

$$k = C(\varepsilon + d)^n \quad \cdots\cdots (2)$$

where ε is the equivalent strain and C, d, n are regression parameters.

In general, the friction coefficient is given by a table, back calculation from the above model equation, or a multiple regression calculation.

On the other hand, Hitchcock's roll flattening theory assumes that the shape of the roll surface remains circular even after roll elastic deformation and the radius R' is given under the elliptical distribution of rolling pressure.

$$R' = R\left[1 + \frac{16\left(1-\upsilon^2\right)}{\pi E}\frac{P}{h_1 - h_0}\right] \quad \cdots\cdots\cdots\cdots\cdots\cdots (3)$$

where R is the original roll radius before deformation, E, ν are the Young's modulus and Poisson's ratio of the roll and P is the rolling force per unit width.

Consequently, the above equations (1) and (2) show that the rolling force P is determined by the flattened roll radius R', and the flattened roll radius R' is determined by the rolling force P. Therefore, it is possible to calculate the rolling force and flattened roll radius by converging calculation using these two equations under appropriate initial values (e.g. P=0, R'=R).

Furthermore, the following physical models are well known for forward slip f and rolling torque G.

$$\left.\begin{aligned} f = \tan^2\Bigg(&\tan^{-1}\left(\sqrt{\frac{h_1-h_0}{h_0}}\right) \\ &-\frac{1}{4\mu}\log\left(\frac{h_1}{h_0}\frac{1-\sigma_0/k_0}{1-\sigma_1/k_1}\right)\sqrt{\frac{h_0}{R'}}\Bigg), \end{aligned}\right\} \quad \cdots\cdots\cdots (4)$$

$$G = \frac{R}{R'}\int pxdx + \frac{R}{2}\left(h_1\sigma_1 - h_0\sigma_0\right) \quad \cdots\cdots\cdots\cdots\cdots\cdots (5)$$

2.1.3 Motor power check[4)]

Using the predicted values of rolling torque and forward slip obtained in the previous section,

$$\left.\begin{aligned} W &= 2G\frac{\upsilon_R}{R} + \upsilon_0 h_0\sigma_0 - \upsilon_1 h_1\sigma_1 \\ &= 2G\frac{\upsilon_0}{(1+f)R} + \upsilon_0 h_0\sigma_0 - \upsilon_1 h_1\sigma_1 \end{aligned}\right\} \quad \cdots\cdots\cdots\cdots (6)$$

where v_R, v_0, v_1 respectively mean the roll speed, the strip speed at the roll bite outlet and the strip speed at the roll bite inlet.

When each roll speed is determined to be the motor top speed, if there is a stand satisfying $W > W_{\text{rated}}$ for the rated output W_{rated}, the roll speed is reduced by the exceedance ratio, and the calculation in the previous paragraph is then performed again.

2.1.4 Optimization of reduction and tension settings

In the discussion up to the preceding paragraph, the reduction ratio and the tension between stands were determined, and the predicted values of rolling force, etc. were obtained. Then, the reference values of the roll gap and the roll speed can be determined by the method in the following section 2.1.5. However, some of the previously proposed methods for setting the reduction ratio and the tension between stands include numerical optimization (e.g. optimization of the reduction ratio so as to maximize the sum of the estimated rolling powers of each stand). In this case, the procedure from 2.1.1 to 2.1.3 is repeated.

2.1.5 Roll gap and roll speed setting

Finally, the roll gap S is obtained from the relationship between the elastic deformation of the mill due to the rolling force, the reference value of the roll gap and the nominal value of strip thickness as follows:

$$S = h_0 - \frac{Pb}{M} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots (7)$$

where b is the strip width and M is the mill modulus.

The roll gap and roll speed calculated in the above flow are set as reference values to the hydraulic cylinder controller and the motor drive controller, respectively.

2.2 Overview of Developed System

Figure 2 shows the configuration of the newly-developed setup system. In this system, a machine learning server was newly installed and was interfaced with the Level-2 computer. A general-purpose product was selected as the hardware of the machine learning server.

The machine learning server collects operation records such as the rolling force obtained from Level-1 (PLC: Programmable Logic Controllers and DCS: Dis-

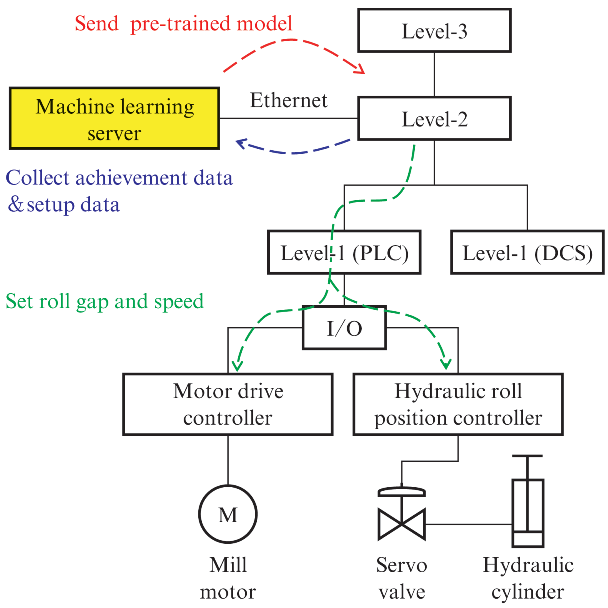

Fig. 2 System configuration diagram

tributed Control Systems), setup data from the Level-2 computer and production conditions such as the coil dimensions and rolling roll information from the Level-3 computer (business computer) through the Level-2 computer, and stores this information in the database.

The Level-2 computer predicts the suitable settings of rolling reduction and tension, rolling force, rolling torque and forward slip for each stand by using machine learning models. Concretely, the calculation is carried out using the prediction calculation logic described in the next chapter, and the prediction value of the conventional model is replaced by that of the machine learning models. The reference values of the roll gap and roll speed are set to the controllers through the Level-1 computer as change quantity data at the timing of flying gauge changes. In this way, the predictions of the new models are reflected in mill control.

In the machine learning server, model learning is carried out when the work rolls have been changed, and the learned parameters are transmitted to the Level-2 computer when learning has been completed. The Level-2 computer can update the models automatically by rereading the parameters appropriately.

The system configuration described above was adopted because it offers the following advantages.

(1) The model learning function can be improved without changing the Level-2 computer.
(2) Expensive learning calculations can be excluded from the Level-2 computer.
(3) Predictive calculation performance (speed) can be improved by execution in the Level-2 computer.

3. Model Construction by Machine Learning

This chapter presents an outline of the newly-developed prediction models used in the setup system.

The model learning method described below is an application of the error back propagation method[5] to a regression model composed of a multilayer perceptron, which is a kind of neural network, as shown in **Fig. 3**.

3.1 Prediction Step

This section describes the predictive calculation procedure by the multilayer perceptron. In this model, all objective variables are multidimensional (e.g. the number of objective variables in rolling force prediction is the same as the number of tandem mill stands).

First, the explanatory variable x is appropriately selected from the material conditions (e.g. strip thickness, strip width, strip hardness, etc.) and the roll conditions (e.g. roll radius, roughness, etc.) and is given as a n-th vector by the following equation:

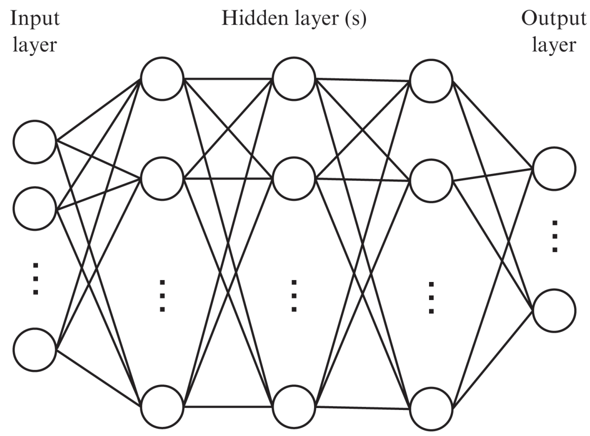

Fig. 3 Multi-layer perceptron network

$$x = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \quad \text{.............................. (8)}$$

The objective variable y is given as a m-th vector as follows:

$$y = \begin{bmatrix} y_1 & y_2 & \cdots & y_m \end{bmatrix} \quad \text{.............................. (9)}$$

For simplicity, the notation of the order of the weight parameter is omitted below, but it is assumed to be an appropriate size corresponding to the input and output.

Assuming the weight parameters are W_1, b_1 and the number of neurons in the first intermediate layer is n_1, the output vector h_1 of the first intermediate layer is denoted as:

$$h_1 = f_1\left(x \cdot W_1 + b_1\right) \quad \text{.............................. (10)}$$

where the function f_1 is an activation function ReLU (Rectified Linear Unit).

Further, by repeating the above procedure for the other intermediate layers, the n_i-order output vector h_i of the i-th intermediate layer is obtained by

$$h_i = f_i\left(h_{i-1} \cdot W_i + b_{i-1}\right) \quad \text{.......................... (11)}$$

Here, f_i is also ReLU (Rectified Linear Unit). Finally, the model prediction y is given assuming N as the total number of intermediate layers.

$$y = h_N \cdot W_{N+1} + b_{N+1} \quad \text{.............................. (12)}$$

Each of the above computations is easy to implement and quickly computable because it is a combination of simple matrix computations and activation functions. In this system, the prediction calculation logic explained in this section is implemented on the Level-2 computer, and the elapsed time for a predictive calculation is about 1 millisecond.

When producing a rare grade steel, a model by machine learning may calculate a largely deviated prediction value (due to a lack of performance data). Therefore, this system suppresses large deviations by performing threshold processing using existing conventional model calculation values.

Since the model construction method in this paper does not use unique knowledge of rolling theory (except in the appropriate selection step of the explanatory variable), it can be said that the effectiveness of this method is not limited to the rolling process.

3.2 Learning Step

The error back propagation method propagates the error from the output layer to the input layer so as to reduce the errors between the teacher data t and the model predicted value y, and updates the weight parameter of each layer. Although the specific procedure will be described below, in order to simplify the description, the total number of intermediate layers is assumed to be N=1. Then, the calculation procedure in the previous section is denoted as follows:

$$h = f\left(x \cdot W_1 + b_1\right), \quad y = h \cdot W_2 + b_2 \quad \cdots\cdots\cdots\cdots\cdots \quad (13)$$

First, the loss function L can be expressed by the following equation, where t is the teacher data and L is defined as the sum of squares error according to the problem formulation in this paper.

$$L = \frac{1}{2}\sum_{i=1}^{m}\left(y_i - t_i\right)^2 = \frac{1}{2}\left\|y - t\right\|^2 \quad \cdots\cdots\cdots\cdots\cdots\cdots \quad (14)$$

Then, its derivative is denoted as

$$\frac{\partial L}{\partial y} = y - t \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (15)$$

Next, the following equation is obtained from the chain rule of the differentiation.

$$\left.\begin{array}{l} \dfrac{\partial L}{\partial h} = \dfrac{\partial L}{\partial y} \cdot W_2{}^T, \\ \dfrac{\partial L}{\partial W_2} = h^T \cdot \dfrac{\partial L}{\partial y}, \\ \dfrac{\partial L}{\partial b_2} = \dfrac{\partial L}{\partial y} \end{array}\right\} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (16)$$

Here, the notation T means the transposition of the matrix and vector. Further, if $h_1=f(v)$, $v=xW_1+b_1$,

$$\left.\begin{array}{l} \dfrac{\partial L}{\partial v} = \begin{cases} \dfrac{\partial L}{\partial h} & \left(v_i \geq 0\right) \\ 0 & \left(v_i < 0\right) \end{cases}, \\ \dfrac{\partial L}{\partial W_1} = x^T \cdot \dfrac{\partial L}{\partial v}, \\ \dfrac{\partial L}{\partial b_1} = \dfrac{\partial L}{\partial v} \end{array}\right\} \quad \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \quad (17)$$

From the above discussion, the loss function is expected to be reduced by updating the weight parameters as follows:

$$W_i \leftarrow W_i - \alpha\frac{\partial L}{\partial W_i}, \quad b_i \leftarrow b_i - \alpha\frac{\partial L}{\partial b_i} \quad \cdots\cdots\cdots\cdots \quad (18)$$

where α is the learning rate of the stochastic gradient descent, and a small value (e.g. 0.001, 0.01, etc.) is selected.

A gradual reduction of the loss function is expected by repeating this learning procedure for a large amount of data. In other words, by repeating this procedure, the output y is expected to approach the teacher data t.

In this system, the learning procedure described above is installed in the machine learning server, and learning of each model is carried out using the operation data accumulated in its database.

4. Result of Application

4.1 Operation Status of Developed System

The setup system using machine learning described in the previous chapter has been applied to a cold rolling plant at this company's West Japan Works. At present, the five kinds of models shown in **Table 1** are used practically.

4.2 Verification of Prediction Accuracy

Table 2 shows the prediction error of the conventional model and model proposed in this paper. In the evaluation of prediction error in Table 2, RMSE (Root Mean Squared Error) of the prediction value and the actual value was calculated. First, RMSE in the conventional system (physical models based on rolling the-

Table 1 Objective variables of the models

Name	Number of objective variables
Draft setting	Equals to the number of mill stands.
Tension setting	Equals to the number of mill stands plus one.
Rolling force	Equals to the number of mill stands.
Rolling torque	Equals to the number of mill stands.
Forward slip	Equals to the number of mill stands.

Table 2 Evaluation of predictive errors of the models (Conventional system = 1)

Steel grade	Model name	Conventional system	Proposed system
A	Rolling force	1	0.41
	Rolling torque	1	0.45
	Forward slip (roll speed)	1	0.82
B	Rolling force	1	0.26
	Rolling torque	1	0.93
	Forward slip (roll speed)	1	0.81

Table 3 Evaluation of application of models

Expected effects of application	Main result
Less welded point breaks during flying gauge control	26.4% decrease
Increased productivity	2.4% increase in a typical grade

ory, as described in Chapter 2) was defined as "1," and that of the proposed system was evaluated relatively. Second, for the accuracy of forward slip, the predicted error of the ratio of the rolling roll speeds at each stand was evaluated because it is difficult to obtain actual results for forward slip. The evaluation period was approximately 3 months. Each model was evaluated by the prediction errors for two typical steel grades. The prediction errors of the reduction and tension settings cannot be defined because they are setting values, (i.e. suitable reference values). Therefore, the evaluation of these two items is omitted here.

Table 2 shows that the accuracy of each model was improved by the new system. It should be noted that the evaluation greatly depends on the steel grade because the accuracy of the conventional model differs substantially depending on the steel grade.

4.3 Effect of Application

Table 3 shows an evaluation of the effects of this system on operation stabilization and productivity improvement.

First, the frequency of occurrence of strip rupture at weld points was reduced by 26.4% by improving the accuracy of rolling force prediction.

Second, a 2.4% improvement in productivity was confirmed with a typical steel grade as a result of improvement of the prediction accuracy of rolling torque and forward slip and suitable setting of rolling reduction and tension.

5. Conclusion

This paper described a new tandem mill setup system using machine learning which is now in operation in the cold rolling mill of JFE Steel Corporation's West Japan Works.

(1) Improvement of the accuracy of various rolling models has contributed to operation stabilization and productivity improvement.
(2) This system also realized maintenance-free operation by an automatic learning process.

References

1) Kamata, M. Thin sheet continuous rolling. Chijinshokan, 1997, 308 p.
2) Murakami, A.; Nakayama, M.; Okamoto, M.; Sano, K. Path Schedule Optimization of a Tandem Cold Mill. Iron and Steel. 2004, vol. 90, no. 11, p. 953–957.
3) Toyofuku, T.; Takenokoshi, A.; Yamamoto, M.; Tanaka, H. Methods for Determining Draft Schedules in Tandem Mills Using Neural Networks. CAMP-ISIJ. 1991.
4) Matsumoto, H. Theory and Practice of Strip Rolling (Revised), chapter 2. ISIJ, 2000, 350 p.
5) Rumelhart, D. E.; Hinton, G. E.; Williams, R. J. Learning internal representations by error propagation. Parallel Distributed Processing: Exploration in the Microstructure of Cognition. 1986, vol. 1, p. 318–362.