AIMS Mathematics, 10(11): 27058–27072.
DOI: 10.3934/math.20251189
Received: 01 July 2025
Revised: 28 July 2025
Accepted: 15 August 2025
Published: 21 November 2025

https://www.aimspress.com/journal/Math

Research article

Prediction model of strip rough rolling width based on the fusion of mechanism and data

Shengyue Zong[1,*] **and Xiaolong Li**[2]

[1] National Engineering Research Center for Advanced Rolling Technology and Intelligent Manufacturing, University of Science and Technology Beijing, Beijing 100083, China

[2] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

* **Correspondence:** Email: zongsy@nercar.ustb.edu.cn.

Abstract: In the rough rolling production process of strip steel, the width change is affected by the coupling of multiple factors. Accurate prediction of the width expansion is of great significance for improving the yield rate and product quality. This paper proposes a strip rough rolling width prediction model that integrates the mechanism model and the data-driven method. First, the width expansion mechanism model is established based on the Shibahara equation. After linearization by Taylor expansion, the least squares method and linear support vector regression algorithm are used to estimate the main mechanism parameters. Second, the extreme gradient boosting learning mechanism is introduced to correct the deviation after model parameter identification, thereby improving the robustness and accuracy of the prediction. Finally, the parameter convergence of the proposed method is verified in the simulation. The result shows that the proposed method is superior to the traditional modeling method in terms of prediction accuracy and convergence speed. This study provides a high-precision modeling idea for rolling width control that integrates physical constraints and data learning.

Keywords: width prediction; parameter identification; least squares method; support vector regression; extreme gradient boosting
Mathematics Subject Classification: 35Q93, 62J05, 68T05

1. Introduction

With the continuous development of intelligent and high-end manufacturing, the steel industry, as an important representative of process industry, is gradually evolving its production control level towards digitalization and model-driven, intelligent prediction. As one of the key links in the hot rolling process, the width control problem in the rough rolling process of strip steel has long been of concern to

the industry [1–3]. Width deviation not only directly affects the size matching of downstream finishing and shearing processes, but may also cause product scrapping, energy waste, and equipment damage in extreme cases. Inaccurate width prediction may lead to mismatched shearing dimensions, resulting in reduced yield and material waste. In severe cases, insufficient or excessive rolling compensation may cause equipment damage. Therefore, accurately predicting the width change trend in the rough rolling stage is of great significance for improving the yield rate, ensuring product quality and realizing closed-loop control [4–6].

The widening behavior of strip steel is a typical multi-factor coupling and strongly nonlinear dynamic response process. Its changes are affected by complex factors such as the thermal deformation characteristics of the material, the structural configuration of the rolling equipment, the rolling parameters (such as the amount of reduction, the roll gap size, the rolling rhythm), and the edge stress field distribution [7, 8]. In actual production, these factors often work together in a highly coupled and time-varying uncertain manner, making the prediction of widening face great challenges. In recent years, the width prediction model has been gradually applied in hot rolling systems, especially in the roughing mills of major steel companies. In order to characterize the widening law during the rough rolling of strip steel, researchers have proposed many empirical equation models based on physical mechanisms. Then, the Shibahara equation has become one of the most widely used width prediction models in engineering practice due to its good physical interpretation and wide industrial applicability [9]. By introducing the concepts of "natural widening" and "dog bone widening", the model systematically analyzes the main factors affecting the widening, and constructs a clearly structured calculation expression, which can reflect the deformation characteristics during the rough rolling process to a certain extent. The advantages of this type of method are clear physical interpretation and good generalization ability, especially when the parameters are controllable and the working conditions are stable. However, its structure is usually complex, the parameters are highly coupled, and it relies on a lot of prior knowledge and experimental data for parameter adjustment. Once faced with realistic factors such as changes in material properties, equipment wear, or fluctuations in process rhythm, the predictive performance of static models will drop rapidly, and there is a lack of an online adaptive update mechanism [10, 11].

Unlike the modeling method based on physical mechanism, the pure data-driven method does not rely on any explicit physical model in strip width prediction, but directly learns the mapping relationship between input (such as inlet width, thickness, rolling force, and rolling temperature) and output (width expansion) based on historical rolling data. Commonly used machine learning algorithms include multi-layer perceptron (MLP) [12], deep stochastic configuration neural network (DeepSCN) [13], support vector machine (SVM) [14], random forest (RF) [15], gradient boosting decision tree (GBDT) [16], and extreme gradient boosting (XGBoost) [17]. These methods have good modeling capabilities when dealing with high-dimensional and strongly nonlinear problems. Neural network models represented by MLP can automatically extract deep features and adapt to complex process conditions. SVM has good generalization ability in small sample modeling, and ensemble learning methods such as XGBoost effectively improve prediction accuracy and stability by integrating multiple weak classifiers [18, 19]. These methods usually have high modeling efficiency, are easy to implement, and have achieved certain results in industrial data scenarios.

Pure data-driven methods also have obvious limitations when applied to strip width prediction. First, due to the lack of physical mechanism constraints, such methods have a strong dependence on

the distribution of training data, and it is difficult to guarantee the generalization performance of the model under unseen working conditions or extreme disturbances. Especially when production conditions migrate, the model is prone to mismatch, resulting in prediction bias or anomalies. Second, the black box nature limits its interpretability, and the decision-making mechanism within the model is often difficult to analyze, resulting in operators being unable to accurately trace the causes when faced with prediction anomalies, thereby affecting diagnostic efficiency or misleading adjustment operations, reducing the model's availability and reliability. In addition, some deep models have high data requirements, and the training process relies on a large number of samples and computing resources, which still faces challenges in actual deployment [20, 21]. Although these algorithms have shown strong flexibility and modeling capabilities in complex system modeling, the models are highly dependent on the distribution of training data and are difficult to maintain generalization performance under unseen working conditions. In addition, such models generally have the problem of poor interpretability. Therefore, in recent years, modeling strategies that integrate physical mechanisms with data learning have become a research hotspot, which can improve prediction accuracy while ensuring model reliability [22, 23]. However, existing hybrid modeling methods mostly focus on static combination, lack a layer-by-layer residual decomposition mechanism, and have unstable performance in strong nonlinear scenarios.

Based on the above research, this paper proposes a width prediction modeling method that integrates mechanism and data, aiming to take into account the interpretability of physical models and the high precision advantages of data models. The main research contributions of this paper are as follows:

- A hybrid modeling framework that integrates physical mechanism models and machine learning algorithms is proposed to realize the system modeling path from structural modeling to residual correction.
- The joint mechanism of the least squares (LS) method, support vector regression (SVR), and XGBoost is used to compensate for the nonlinear errors at different levels in the model in stages, thereby improving the overall prediction accuracy.
- In simulation and comparative experiments, it is verified that the proposed method is superior to traditional modeling methods in terms of parameter convergence speed, reflecting good engineering adaptability and promotion potential.

2. Mechanism model selection

During the rough rolling process of strip steel, the width of the strip steel will change significantly due to the effects of material plastic deformation, rolling force distribution, roller structure, and edge bulge. The widening effect is reflected in both the natural widening caused by the overall deformation and the "dog bone" shaped local widening formed by the accumulation of materials on both sides. The Shibahara equation is an empirical model summarized based on the above phenomenon, which was proposed by Japanese scholars in the 1980s. It is widely used in hot rolling systems because of its clear structure and physical parameters. It has been adopted by many steel mills and embedded in actual control systems. It decomposes the widening amount into two parts: natural widening and dog bone widening. It can more accurately reflect the physical laws of the rough rolling process [9], and

its main calculation equation is as follows:

$$\Delta w_s = w_a \left[\left(\frac{H_e}{H_o} \right)^{d_s} - 1 \right], \tag{1}$$

$$\Delta W_f = z_s \cdot d_e \cdot \left[1 + \frac{\Delta w_s}{w_a} \right], \tag{2}$$

$$d_e = \exp \left[d_1 \cdot q^{d_2} \left(\frac{W_e}{L} \right)^{d_3 \cdot q} \left(\frac{H_e}{R_h} \right)^{d_4 \cdot q} \right], \tag{3}$$

$$z_s = \exp \left[z_1 \cdot \left(\frac{d_e}{W_0} \right)^{z_2} \left(\frac{H}{R_e} \right)^{z_3} \left(\frac{R_e}{W_e} \right)^{z_4} \left(\frac{W_0}{W_e} \right)^{z_5} \right], \tag{4}$$

where Δw_s is the natural width (mm), w_a is the entrance width value (mm), H_e is the strip thickness before pressure measurement (mm), H_o is the strip thickness after horizontal rolling (mm), d_s is the coefficient of thickness variation and edge bulge degree of vertical roller rolling, ΔW_f is the dog bone width (mm), z_s is the dog bone width coefficient, d_e is the horizontal width coefficient, q is the geometric parameter, W_e is the width of the workpiece after the vertical roller is pressed down (mm), L is the arc length of contact with the horizontal roller when the roller is flattened (mm), R_h is the horizontal roller radius (mm), W_0 is the width of the roller (mm), H is the thickness of the rolled product at the entrance of the stand (mm), R_e is the radius of the vertical roller (mm), and z_1-z_5 *and* d_1-d_4 are unknown model parameters.

The core idea of the Shibahara equation is to decompose the widening behavior into the function product of multiple influencing factors, such as rolling reduction rate, roll radius, strip inlet width-to-thickness ratio, and edge bulge degree, to construct a mathematical expression with certain physical meaning. Its mathematical form contains multiple nonlinear terms, especially the exponential function form, which increases the complexity of model calculation. In order to simplify the parameter identification process and adapt to the needs of linear modeling, this paper uses the method of logarithmic transformation and Taylor series expansion to linearize the model. According to Eqs (1)–(4), we can obtain

$$\begin{aligned} \Delta W_f = & \exp \left[z_1 \cdot \left(\frac{d_e}{W_0} \right)^{z_2} \left(\frac{H}{R_e} \right)^{z_3} \left(\frac{R_e}{W_e} \right)^{z_4} \left(\frac{W_0}{W_e} \right)^{z_5} \right] \\ & \times \exp \left[d_1 \cdot q^{d_2} \left(\frac{W_e}{L} \right)^{d_3 \cdot q} \left(\frac{H_e}{R_h} \right)^{d_4 \cdot q} \right] \times \left[1 + \frac{w_a \left[\left(\frac{H_e}{H_o} \right)^{d_s} - 1 \right]}{w_a} \right]. \end{aligned}$$

Taking the natural logarithm of both sides of the equation, we have

$$\begin{aligned}\log \Delta W_f =& z_1 \cdot \left(\frac{d_e}{W_0}\right)^{z_2} \left(\frac{H}{R_e}\right)^{z_3} \left(\frac{R_e}{W_e}\right)^{z_4} \left(\frac{W_0}{W_e}\right)^{z_5} \\ &+ d_1 \cdot q^{d_2} \left(\frac{W_e}{L}\right)^{d_3 \cdot q} \left(\frac{H_e}{R_h}\right)^{d_4 \cdot q} + \log\left(\frac{H_e}{H_o}\right)^{d_s} \\ =& \exp\left[\log z_1 + z_2 \log \frac{d_e}{W_0} + z_3 \log \frac{H}{R_e} + z_4 \log \frac{R_e}{W_e} + z_5 \log \frac{W_0}{W_e}\right] \\ &+ \exp\left[\log d_1 + d_2 \log q + d_3 q \log \frac{W_e}{L} + d_4 q \log \frac{H_e}{R_h}\right] + d_s \log\left(\frac{H_e}{H_o}\right).\end{aligned} \tag{5}$$

The model is linearized using the Taylor expansion method, and we can get

$$\begin{aligned}\log \Delta W_f \approx& 2 + \log z_1 + z_2 \log \frac{d_e}{W_0} + z_3 \log \frac{H}{R_e} + z_4 \log \frac{R_e}{W_e} + z_5 \log \frac{W_0}{W_e} \\ &+ \log d_1 + d_2 \log q + d_3 q \log \frac{W_e}{L} + d_4 q \log \frac{H_e}{R_h} + d_s \log\left(\frac{H_e}{H_o}\right).\end{aligned} \tag{6}$$

In the derivation of Eq (6), Taylor series expansion is applied to approximate the exponential terms, retaining the first-order terms and ignoring the higher-order terms. However, although the higher-order residual terms are simplified in the mathematical derivation, they still exist in practical applications and may accumulate, thereby introducing systematic deviations and affecting the prediction accuracy of the model. In particular, when the parameters of the strip rolling process fluctuate greatly and the measurement error cannot be ignored, the influence of the higher-order residual terms is more obvious.

3. LS basic format

In order to effectively compensate for the potential errors caused by the approximation of high-order terms and improve the robustness and accuracy of parameter identification, this paper refers to references [4, 10] and adopts the LS method combined with the machine learning algorithm for parameter identification. Based on the above-mentioned equation, d_e, q, and d_s are known parameters, ΔW_f, W_o, H, R_e, W_e, L, H_e, R_h, H_e, and H_o are measurable quantities, and the width prediction model is linearized into the following standard LS format by transforming and deriving Eq (6):

$$y(t) = \Phi^T(t)\theta + \varphi(t), \tag{7}$$

where $y(t) = \log \Delta W_f - 2 - d_s \log\left(\frac{H_e}{H_o}\right)$ is the measurement output,

$$\Phi(t) = \left[1,\ \log \frac{d_e}{W_0},\ \log \frac{H}{R_e},\ \log \frac{R_e}{W_e},\ \log \frac{W_0}{W_e},\ 1,\ \log q,\ q \log \frac{W_e}{L},\ q \log \frac{H_e}{R_h}\right]^T,$$

where $\theta = [\log z_1,\ z_2,\ z_3,\ z_4,\ z_5,\ \log d_1,\ d_2,\ d_3,\ d_4]^T$ is the parameter to be identified, and $\varphi(t)$ is the measurement noise, which obeys the Gaussian distribution with mean 0.

Remark 1. *It should be noted that in practical applications, all measurable physical quantities (such as strip width, thickness, roller diameter, etc.) have varying degrees of measurement errors. Therefore, the LS model should not only consider the linear relationship between input and output, but also fully consider the impact of noise characteristics on the final parameter estimation.*

The control and measurement in actual industrial systems are based on discrete time frameworks, and data is usually acquired at a fixed sampling period. In order to match such actual sampling characteristics, this study discretizes and reconstructs the continuous time system and converts it into a discrete time expression, thereby establishing a mathematical model compatible with the digital sampling method of the industrial site. Discretizing Eq (7) yields:

$$y_k = \Phi_k^T \theta + \varphi_k, \tag{8}$$

where k represents the sampling time.

4. Model parameter estimation

Since the second-order and higher-order terms are ignored after the Taylor expansion of Eq (5), directly using LS to estimate the structural parameter θ will bring systematic deviations, while SVR can tolerate some model errors through ε–insensitive bandwidth, but if SVR is used alone, the physical interpretability of LS is lost. Therefore, it is considered to let LS be responsible for capturing most of the first-order mechanism effects, and use SVR to compensate for the systematic deviation of LS. The two are learned alternately through residuals, and finally a parameter estimation that maintains both physical structure and robustness is obtained. The specific steps of the algorithm are as follows:

Step 1: Based on Eq (8), the LS algorithm [24] is used to estimate the parameters in model (6) online. The recursive form is as follows:

$$K_k = P_{k-1}\Phi_k\left[\Phi_k^T P_{k-1}\Phi_k + 1\right]^{-1}, \tag{9}$$

$$\hat{\theta}_k = \hat{\theta}_{k-1} + K_k\left[y_k - \Phi_k^T \hat{\theta}_{k-1}\right], \tag{10}$$

$$P_k = \left[I - K_k\Phi_k^T\right]P_{k-1}, \tag{11}$$

where $\hat{\theta}_k$ represents the parameter estimate at the kth moment, and the initial value $\hat{\theta}_0 = 0$. K_k is the gain matrix at the current moment, and P_k is the covariance matrix. The specific implementation steps of the LS are shown in Algorithm 1.

Step 2: Use SVR to model high-order residuals, and its iterative form is as follows:

1) Treat the high-order residuals as a function of the input feature x_k, that is,

$$r_k = y_k - \Phi_k^T \theta \approx f(x_k),$$

where $f(\cdot) \in \mathcal{H}_{\text{linear}}$, $\mathcal{H}_{\text{linear}}$ represents the reproducing kernel Hilbert space corresponding to the linear kernel.

2) According to $\hat{\theta}_k$ estimated by the LS algorithm, the residual is calculated as:

$$\hat{r}_k = y_k - \Phi_k^T \hat{\theta}_k. \tag{12}$$

Algorithm 1: The estimation of parameters of width prediction model using the LS method

Given an initial value: $\hat{\theta}_0$, P_0
Input: measurement dataset $\{\Phi_0, \Phi_1, \ldots, \Phi_k\}$ and measurement output set $\{y_0, y_1, \ldots, y_k\}$
for $i = 1, 2, \ldots, k$ **do**
  Based on Eq (9) to update K_i
  Based on Eq (10) to update $\hat{\theta}_i$
  Based on Eq (11) to update P_i
Output: The parameter estimation $\hat{\theta}_k$

3) SVR training: On the training sample set $\{(x_1, \hat{r}_1), (x_2, \hat{r}_2), \ldots, (x_k, \hat{r}_k)\}$, solve the following optimization problem:

$$\min_{w_k, b_k, \xi, \xi^*} \frac{1}{2}\|w_k\|^2 + C\sum_k (\xi_k + \xi_k^*)$$

$$\text{s.t.} \begin{cases} \hat{r}_k - \langle w_k, x_k\rangle - b_k \le \varepsilon + \xi_k \\ \langle w_k, x_k\rangle + b_k - \hat{r}_k \le \varepsilon + \xi_k^* \\ \xi_k, \xi_k^* \ge 0 \end{cases}$$

where w_k is the parameter vector to be learned by the linear SVR, b_k is the bias term, C is the penalty factor, ε is the insensitive bandwidth, ξ_k, ξ_k^* are slack variables used to process samples with errors greater than ε, representing the errors when the predicted value is higher or lower than the true value, respectively.

4) Output the final residual model as follows:

$$f_k(x) = \langle w_k, x\rangle + b_k. \tag{13}$$

Step 3: Re-estimate the LS algorithm with residual correction, that is, use $z_k = y_k - f_k(x)$ to construct the corrected measurement output, and substitute it into Algorithm 1 for the estimation update. When $\|\hat{\theta}_{k+1} - \hat{\theta}_k\| < \delta$ or the maximum number of iterations is reached, the final estimated parameter $\hat{\theta}$ is output.

In summary, the parameter estimation method based on LS and SVR is shown in Algorithm 2.

Algorithm 2: The parameter estimation based on LS and SVR

Input: measurement dataset $\{\Phi_0, \Phi_1, \ldots, \Phi_k\}$, measurement output set$\{y_0, y_1, \ldots, y_k\}$ and the input feature set $\{x_0, x_1, \ldots, x_k\}$
for $i = 1, 2, \ldots, k$ **do**
  Solve $\hat{\theta}_k$ using Algorithm 1
  Calculate the residual according to Eq (12)
  Use SVR training to solve the residual model (13)
  Correct the measurement output
  Update $\hat{\theta}_k$ using Algorithm 1
Output: The parameter estimation $\hat{\theta}$

5. Model bias correction

Although SVR has partially corrected the first-order approximation error, the nonlinear nature of the exponential term may still cause the high-order residuals to not be effectively fitted. In order to further improve the final prediction accuracy, this paper uses the XGBoost algorithm in the second stage to fit and correct the residuals between the predicted values and the actual values in the first stage.

Assume that the output of the first stage is $\hat{y}_k = \Phi_k^T \hat{\theta}$, the actual value is y_k, the residual is $e_k = y_k - \hat{y}_k$, and the regression tree model is introduced:

$$e_k = g(x_k) = \sum_{j=1}^{n} g_j(x_k),\ g_j \in \mathcal{G},$$

where $g(x_k)$ is the residual fitting function output by the XGBoost model, $g_j(x_k)$ represents the jth regression tree, n is the total number of trees, and $\mathcal{G}$ represents the function space of all optional tree models. XGBoost is trained by minimizing the following objective function:

$$\text{loss}_j = \sum_{i=1}^{k} \left(y_i - \hat{y}_i - g_j(x_i)\right)^2 + \Omega(g_j),$$

where $\Omega(g_j) = \gamma n + \frac{1}{2}\lambda \sum_{j=1}^{n} \varpi_j^2$ represents the complexity regularization term of the jth tree, which is used to prevent overfitting. γ is the penalty factor for controlling the complexity of the model, λ is the L2 regularization parameter, which is used to limit the size of the leaf node score, and ϖ_j is the weight of the jth leaf node.

By iteratively training the residual tree ensemble model $g(x_k)$, the first stage output is finally corrected for deviations to obtain a complete prediction model:

$$\tilde{y}_k = \hat{y}_(k) + g(x_k).$$

That is,

$$\tilde{y}_k = \Phi_k^T \hat{\theta} + g(x_k).$$

According to the above steps, this study adopts a staged hybrid estimation strategy, combines LS and SVR to estimate parameters, and then uses the XGBoost residual correction mechanism to gradually improve the accuracy and robustness of parameter estimation. The specific algorithm is implemented as follows.

6. The simulation

Under the set working conditions, for Eqs (1)–(4), this section uses MATLAB software to write a simulation program to explore the accuracy of the strip rough rolling width prediction model of mechanism and data fusion. First, based on engineering experience, the known parameter values are given and the unknown parameter values are set, as shown in Tables 1 and 2.

Algorithm 3: The parameter estimation of strip rough rolling width prediction model based on mechanism and data fusion

Input: rolling process data (temperature, pressure, size, etc.)
Output: width prediction value
Parameter identification stage:
Use LS to estimate unknown parameters
Use SVR to compensate for linearization errors
Use SVR output to correct measured values and re-estimate parameters
Residual correction phase:
Calculate the residual between the predicted value of the mechanical model and the actual value
Use the XGBoost algorithm to learn complex nonlinear patterns in the residual
Fusion prediction:
Final prediction = Mechanism model output + XGBoost residual correction

Table 1. Known parameter values.

System parameters	values
w_a	1094.67 mm
H_e	255 mm
H_o	202 mm
d_s	0.35
W_e	1073 mm
L	168.42 mm
q	4.2928
R_h	1307.2 mm
W_0	1073 mm
H	255 mm
R_e	1159.95 mm

Table 2. Unknown parameter value.

Unknown parameter	value
z_1	0.9944
z_2	0.0392
z_3	2.7048
z_4	3.5573
z_5	2.2847
d_1	0.9944
d_2	1.2090
d_3	-0.2268
d_4	8.4774

According to the above model parameter settings, the corresponding measurement output ΔW_f is obtained. We collect a 100 s data set with a sampling step of 0.01 s/time, and set the noise to follow a Gaussian distribution with an expectation of 0 and a variance of 0.01. The Gaussian distribution noise introduced simulates the typical level of sensor measurement error in actual industrial sites. This setting helps to test the robustness of the parameter estimation algorithm in a noisy environment. There are 10,000 sets of output data with measurement noise. Next, the algorithm in this paper is used to estimate the unknown parameters $z_1, z_2, z_3, z_4, z_5, d_1, d_2, d_3, and\ d_4$, and the corresponding estimated values $\hat{z}_1, \hat{z}_2, \hat{z}_3, \hat{z}_4, \hat{z}_5, \hat{d}_1, \hat{d}_2, \hat{d}_3, and\ \hat{d}_4$ are obtained. And, by adjusting the regularization terms in XGBoost (such as λ, γ) and combining cross-validation for parameter selection to control the risk of overfitting. The results are shown in Figures 1 and 2. As can be seen from the figure, the estimation algorithm in this paper ensures that the estimated values of the parameters converge to the true values. In addition, we find that features such as z_1 and d_1 contribute the most to residual correction.

In order to verify the parameter estimation performance of the proposed fusion method, this paper compares and analyzes the method with the traditional scheme that only uses LS for identification. In order to quantify the convergence accuracy of each estimated parameter, the following error index is defined:

$$\varsigma = \sum_{k} |\theta - \hat{\theta}_k|,$$

where ς reflects the cumulative deviation between the estimated value and the true value. The smaller the value, the higher the parameter estimation accuracy. Table 3 gives the error comparison results of the two methods under nine parameters to be estimated. It can be seen that in all parameter dimensions, the fusion method proposed in this paper is significantly better than the traditional LS. Especially in parameters such as z_2, d_2, and d_4, the error of the proposed method decreases by more than 30%, showing stronger convergence ability and robustness. This shows that the introduction of SVR and XGBoost to model the residual effectively compensates for the systematic deviation caused by the first-order linear approximation, thereby improving the overall estimation performance.

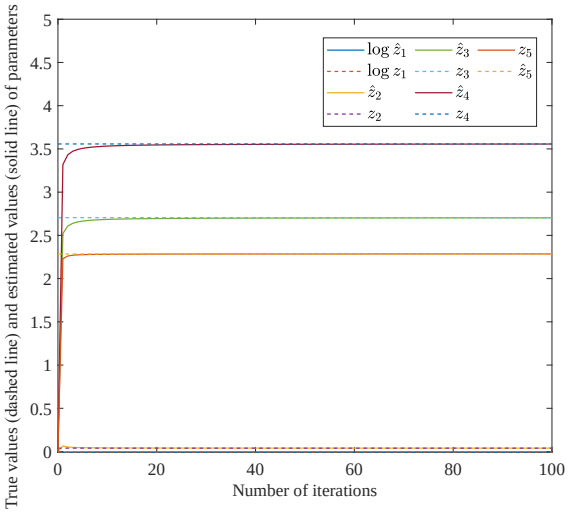

Figure 1. The true value and estimated value of $\log z_1, z_2, z_3, z_4, z_5$.

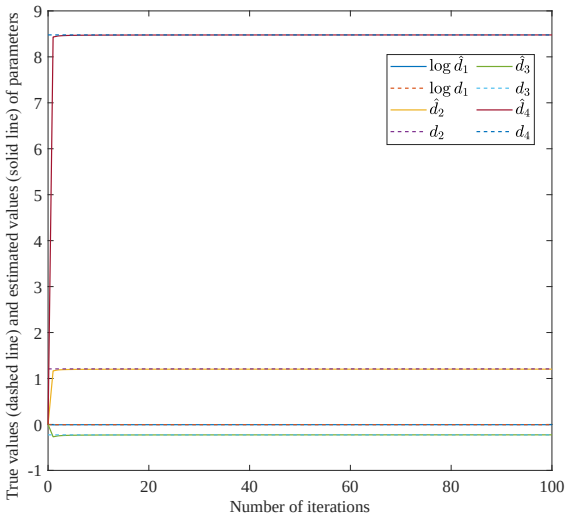

Figure 2. The true value and estimated value of $\log d_1, d_2, d_3, d_4$.

Table 3. The estimation effect comparison.

The error	$\log z_1$	z_2	z_3	z_4	z_5	$\log d_1$	d_2	d_3	d_4
LS	0.0001	0.0049	7.5126	12.9970	5.2426	0.0001	1.4784	0.0624	71.8802
This algorithm	0	0.0025	7.3704	12.7515	5.2240	0	1.4688	0.0542	71.8705

Remark 2. *The computational complexity of the LS part is linear ($O(n)$), and SVR and XGBoost are batch update methods, not real-time updates per sampling point. In a general PC environment, each round of iteration takes no more than 100 ms on average. If deployed in industry, near real-time applications can be achieved through code optimization and hardware acceleration.*

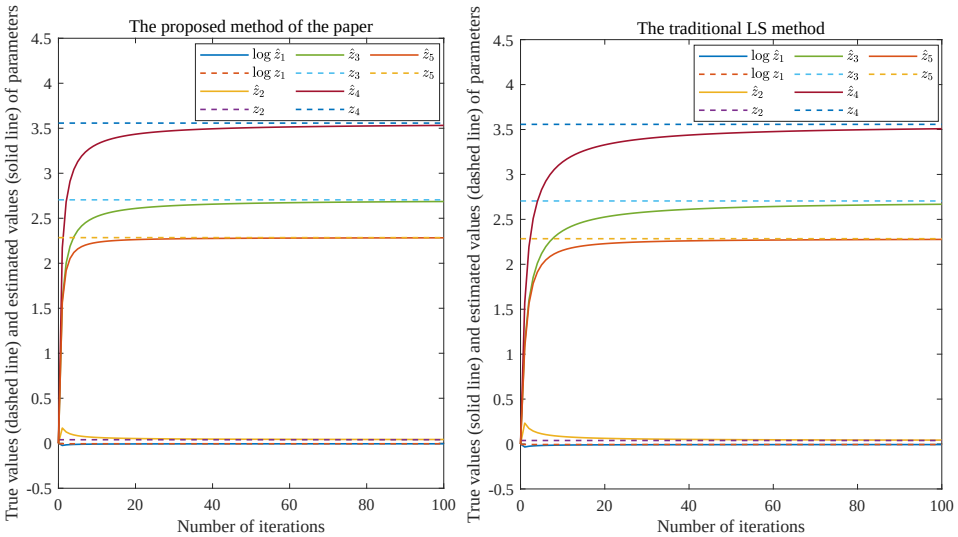

Figure 3. Comparison of the estimation effect of $\log z_1, z_2, z_3, z_4, z_5$ between the proposed method and the traditional LS method under small sample conditions.

In order to verify the sample efficiency advantage of this method in the case of small samples, we added a small sample experiment (the number of samples was reduced from 10000 to 1000) and compared the estimation performance of this method with the traditional LS method, as shown in Figures 3 and 4. The results show that this method still maintains a good parameter estimation effect in the small sample scenario, while the performance of the traditional LS method is significantly reduced.

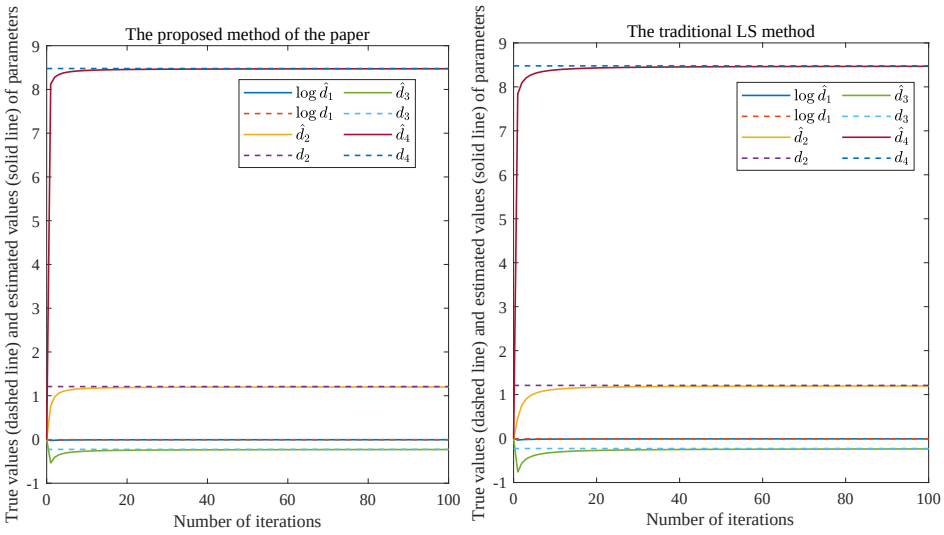

Figure 4. Comparison of the estimation effect of $\log d_1, d_2, d_3, d_4$ between the proposed method and the traditional LS method under small sample conditions.

7. Conclusions

This paper proposes a high-precision modeling framework that integrates mechanism modeling and data-driven methods to predict the width expansion during the rough rolling process of strip steel. This method is based on the Shibahara physical model, and linearization is achieved through Taylor expansion. The LS combined with SVR algorithm is used to estimate the key mechanism parameters online. Subsequently, the XGBoost algorithm is introduced to perform nonlinear correction on the mechanism modeling error, achieving the unity of physical interpretability and data adaptability. The simulation results show that the proposed method is superior to the traditional mechanism method in parameter identification accuracy, and is suitable for industrial field applications. In the future, we will further study the width modeling strategy based on deep learning, such as considering the use of a recurrent neural network or transformer structure based on an attention mechanism to better capture the timing characteristics of the rolling process. In addition, feature enhancement methods will be introduced to improve the model's ability to express complex nonlinear patterns, realize online adaptive updates of the model, and introduce parameter regularization, sensitivity analysis, and variable reconstruction to improve identifiability.

Author contributions

Shengyue Zong: Conceptualization, theoretical analysis, methodology design, supervision, project administration, and writing – review and editing; Xiaolong Li: Software development, data curation, experimental validation, visualization, and writing – original draft. All authors have read and approved the final version of the manuscript for publication.

Use of Generative-AI tools declaration

The authors declare that they have not used Artificial Intelligence (AI) tools in the creation of this article.

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. W. Wu, W. Peng, Y. Liu, J. Liu, X. Li, D. Zhang, et al., Predictive modeling of strip width based on incremental learning and adaptive-weight fusion during the hot rolling process, *J. Manuf. Process.*, **142** (2025), 157–176. https://doi.org/10.1016/j.jmapro.2025.03.091
2. P. Ren, Q. Liu, J. Song, B. Yan, M. Liang, S. Dai, et al., Heterogeneous microstructure and texture evolution of K-doped tungsten thick plate during hot-rolling processing, *Nucl. Mater. Energy*, **43** (2025), 101931. https://doi.org/10.1016/j.nme.2025.101931
3. R. Jia, T. Wang, W. Xue, J. Guo, Y. Zhao, Multi-time scale consensus algorithm of multi-agent systems with binary-valued data under tampering attacks, *IEEE Trans. Ind. Inf.*, **12** (2025), 9377–9388. https://doi.org/10.1109/TII.2025.3598451
4. J. Guo, X. Wang, W. Xue, Y. Zhao, System identification with binary-valued observations under data tampering attacks, *IEEE Trans. Autom. Control*, **66** (2021), 3825–3832. https://doi.org/10.1109/TAC.2020.3029325
5. X. Li, J. Wang, J. Wang, J. Wang, J. Chen, X. Yu, Research on CNC machine tool spindle fault diagnosis method based on DRSN–GCE model, *Algorithms*, **18** (2025), 304. https://doi.org/10.3390/a18060304
6. P. Yu, Y. Hu, Y. Wang, R. Jia, J. Guo, Optimal consensus control strategy for multi-agent systems under cyber attacks via a stackelberg game approach, *IEEE Trans. Autom. Sci. Eng.*, **22** (2025), 18875–18888. https://doi.org/10.1109/TASE.2025.3591858
7. M. Nilsson, M. Olsson, Microstructural, mechanical and tribological characterisation of roll materials for the finishing stands of the hot strip mill for steel rolling, *Wear*, **307** (2013), 209–217. https://doi.org/10.1016/j.wear.2013.09.002
8. K. Wang, R. Doerner, M. Baldwin, C. Parish, Flux and fluence dependent helium plasma-materials interaction in hot-rolled and recrystallized tungsten, *J. Nucl. Mater.*, **510** (2018), 80–92. https://doi.org/10.1016/j.jnucmat.2018.07.04

9. T. Shibahara, K. Misaka, T. Kono, M. Koriki, Y. Takemoto, Edger set-up model at roughing train in hot strip mill, *Tetsu-to-Hagane*, **67** (1981), 2509–2515. https://doi.org/10.2355/tetsutohagane1955.67.15_2509
10. J. Guo, Q. Zhang, Y. Zhao, Identification of FIR systems with binary-valued observations under replay attacks, *Automatica*, **172** (2025), 112001. https://doi.org/10.1016/j.automatica.2024.112001
11. F. Jing, F. Li, Y. Song, J. Li, Z. Feng, J. Guo, Root cause tracing using equipment process accuracy evaluation for looper in hot rolling, *Algorithms*, **17** (2024), 102. https://doi.org/10.3390/a17030102
12. S. Zhang, Z. Liu, T. An, X. Cui, X. Zeng, N. Shi, et al., Hot rolled prognostic approach based on hybrid Bayesian progressive layered extraction multi-task learning, *Expert Syst. Appl.*, **249** (2024), 123763. https://doi.org/10.1016/j.eswa.2024.123763
13. L. Meng, J. Ding, X. Li, G. Cao, Y. Li, D. Zhang, Novel shape control system of hot-rolled strip based on machine learning fused mechanism model, *Expert Syst. Appl.*, **255** (2024), 124789. https://doi.org/10.1016/j.eswa.2024.124789
14. T. Hussain, J. Hong, J. Seok, A hybrid deep learning and machine learning-based approach to classify defects in hot rolled steel strips for smart manufacturing, *Comput., Mater. Continua*, **80** (2024), 2099–2119. https://doi.org/10.32604/cmc.2024.050884
15. J. Lu, P. Wang, H. Huang, L. Hao, X. Li, Q. Peng, et al., Novel online prediction model for thermal convexity of work rolls during hot steel rolling based on machine learning algorithms, *Expert Syst. Appl.*, **254** (2024), 124384. https://doi.org/10.1016/j.eswa.2024.124384
16. C. Ding, J. Sun, X. Li, W. Peng, D. Zhang, A high-precision and transparent step-wise diagnostic framework for hot-rolled strip crown, *J. Manuf. Syst.*, **71** (2023), 144–157. https://doi.org/10.1016/j.jmsy.2023.09.007
17. Z. Wang, Y. Liu, T. Wang, J. Wang, Y. Liu, Q. Huang, Intelligent prediction model of mechanical properties of ultrathin niobium strips based on XGBoost ensemble learning algorithm, *Comput. Mater. Sci.*, **231** (2024), 112579. https://doi.org/10.1016/j.commatsci.2023.112579
18. M. Li, G. Li, X. Li, L. Lu, P. Mitrouchev, Prediction of Mechanical Properties of Hot-Rolled Strip Steel Based on XGBoost and Metallurgical Mechanism, *Phys. Metals Metallogr.*, **125** (2024), 1728–1738. https://doi.org/10.1134/S0031918X23600653
19. I. Trofimov, P. Korotaev, I. Ivanov, A. Malginov, A. Tokhtamyshev, A. Yanilkin, et al., Machine learning assisted design of reactor steels with high long-term strength and toughness, *Mater. Design*, **254** (2025), 114014. https://doi.org/10.1016/j.matdes.2025.114014
20. B. Duan, Y. Zhu, D. Cao, X. Cai, X. Mao, F. Liu, Machine-learning-based performance prediction for partially encased steel-concrete composite columns after fire exposure, *Structures*, **77** (2025), 109082. https://doi.org/10.1016/j.istruc.2025.109082
21. H. Emami, M. Ghazani, An interpretable stacking machine-learning model to predict the hot torsion flow characteristics of a micro-alloyed steel, *Results Eng.*, **26** (2025), 105326. https://doi.org/10.1016/j.rineng.2025.105326
22. F. Liang, Z. Gao, F. Chai, H. Su, X. Luo, Investigation on predicting flow stress of a Cu-containing steel based on machine learning algorithms, *Mater. Today Commun.*, **46** (2025), 112728. https://doi.org/10.1016/j.mtcomm.2025.112728

23. C. Wang, X. Wei, S. Zwaag, Y. Tan, From creep-life prediction to ultra-creep-resistant steel design: An uncertainty-informed machine learning approach, *Acta Mater.*, **292** (2025), 121073. https://doi.org/10.1016/j.actamat.2025.121073

24. L. Guo, *Time-varying stochastic systems–stability, estimation and control*, Changchun: Jilin Science and Technology Press, 1992.

AIMS Press