

Received May 9, 2022, accepted May 18, 2022, date of publication May 23, 2022, date of current version May 27, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3176900

A Model-Data-Fusion Pole Piece Thickness Prediction Method With Multisensor Fusion for Lithium Battery Rolling Machine

YANJUN XIAO^{1,2,3}, SHUHAN DENG³, FURONG HAN^{2,3}, XIAOLIANG WANG³,
ZONGHUA ZHANG³, AND KAI PENG³

¹Tianjin Key Laboratory of Power Transmission and Safety Technology for New Energy Vehicles, School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China

²Career Leader Intelligent Control Automation Company, Suqian, Jiangsu 223800, China

³School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China

Corresponding authors: Yanjun Xiao (xyj@hebut.edu.cn), Zonghua Zhang (hebutzzh123@163.com), and Kai Peng (h313930@163.com)

This work was supported in part by the Jiangsu Province Training Fund under Grant BRA2020244.

ABSTRACT Trend prediction based on sensor data is an important topic in the thickness control system of lithium battery electrode mills. As the number of sensors increases, we can measure and store more and more data. The characteristics of nonlinearity, uncertainty, and time-variability in the lithium battery electrode thickness control system. The increase of control system complexity and data volume does not effectively improve the prediction performance. This paper proposes a physical-data fusion modeling prediction method based on a multiform coupling model and Bayesian LSTM (Bayesian Long Short-Term Memory) to achieve dynamic prediction of lithium battery electrode thickness, overcome data irrelevance and sensor noise, ensure the consistency of electrode thickness, and improve the operational efficiency of battery electrode production: Firstly, we establish the underlying physical model of the roll to further obtain the specific parameters affecting the thickness control and overcome the data irrelevance and sensor noise; secondly, we use Bayesian method to obtain the characteristics of the weight distribution of the sub-prediction network and construct the Bayesian LSTM predictor. An MLP (Multilayer Perceptron) is used as the fusion layer to fuse the results of different sub-predictions to improve the robustness of the nonlinear control system prediction model and solve the problems of slow approximation speed and ease to fall into local minimization of traditional neural networks. Finally, the advantages of the deep learning model are analyzed in terms of data feature self-extraction and model generalization generalizability. Compared with other neural network models, Bayesian LSTM has better generalizability for small sample data. The results show that the predictor can effectively model the large measurement data of the thickness control system of lithium battery electrode mills and improve the prediction performance.

INDEX TERMS Multiform coupling, time series forecasting, Bayesian LSTM, deep fusion predictor.

I. INTRODUCTION

The lithium-ion battery electrode coating process is a key process in the manufacture of lithium-ion batteries, and the uniformity of the processed electrode thickness directly affects the safety, consistency, and other key indicators of lithium-ion batteries. Therefore, the real-time monitoring and control of the thickness of the electrode piece are crucial. [1]. The process flow of the lithium battery wafer is shown in Fig.1. In the production process of battery wafer, the most

important process is coated drying and electrode compaction. The coating is the uniform coating of a stable, viscous, and liquid slurry on the positive and negative collectors. In the electrode coating process ensure that the parameters before and after the electrode are consistent, to effectively avoid problems such as differences in battery capacity and large differences in cycle life. The electrode compaction process relies on the formation of friction between the roll and the electrode, which continuously drags the electrode between the upper and lower rolls, and then under the pressure provided by the rolling pressure system, the electrode is plastically deformed. As shown in Fig. 1, the rolling force

The associate editor coordinating the review of this manuscript and approving it for publication was Yiqi Liu.

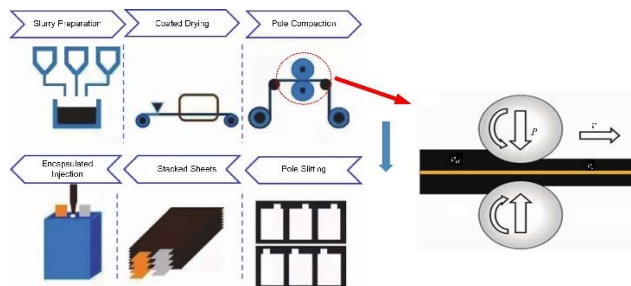


FIGURE 1. Process flow diagram of lithium battery electrodes.

set by the upper and lower rolls is P , and the compaction running speed of the electrode piece with V . The increase of the relative density of the lithium-ion particles in the electrode sheet is mainly manifested in the displacement of the particles, which are displaced between the pores by the roller rolling, filling the particles in the base material. At the same time, a small part of the deformation of the particles occurs, when the rolling force is continuously increased, the void of lithium-ion slurry is filled after a larger deformation [2]. The substrate foil strip is denser ($\rho_{c0} > \rho_c$) with the active material attached to the surface. The electrode's compacted density is maintained in a certain range of space and the consistency of the electrode thickness is maintained. After that, the electrodes are slit and stacked, and finally encapsulated and filled with liquid.

In the electrode rolling process, the thickness of the electrode sheet is not uniform after rolling due to the uneven thickness of the electrode sheet in the previous coating process. When the thickness of the electrode sheet is not uniform, the rate of lithium-ion and electron transport and conduction in the electrode sheet will be different. It is easy to cause the precipitation of lithium dendrites, which is unfavorable to the performance of the electric core. In addition, the contact resistance between the active material and the collector fluid is also different when the thickness of the electrode is different, and the thicker the electrode is, the greater the internal resistance, and the more serious the battery polarization is, which affects the capacity of the electrode. Therefore, it is a key part of the whole processing process to ensure the consistency of the thickness of the electrode.

To solve the problem of inconsistent electrode thickness, many scholars have used real-time thickness measurement to achieve continuous high-quality production. The laser thickness measurement method is characterized by high sensitivity and fast sampling speed, using dual laser displacement sensors up and down differential thickness measurement and adding calibration devices to keep the thickness consistency [2]–[4]. However, the optical measurement method is easily disturbed by impurities in the external air and is prone to errors.

And with the rapid development of artificial intelligence technology, machine learning algorithms are gradually applied to machinery production status prediction.

Fu *et al.* [5] established a multi-intelligence collaborative control system to achieve dynamic and synchronous control of the rolling mill. However, due to the complexity of the rolling process and the uncertainty of the influence of various unknown factors, it is difficult to obtain accurate mathematical models of thickness and plasticity coefficients. Therefore, he established the thickness AGC prediction intelligence and used the fuzzy neural network approach to encapsulate the neural network-based thickness and plasticity coefficient predictions into two bits of intelligence respectively, which replace the ordinary simple mathematical models and interact with other intelligence as auxiliary intelligence, thus further improving the system accuracy. In addition, machine learning algorithms have been applied to the field of power system prediction [6], [7], photovoltaic power generation prediction [8], [9] and Agri-environmental prediction [10], [11], among others. In the field of machining, many scholars have found that the factors affecting the thickness of rolled parts in steel rolling mills are temperature [12], [13], irregular vibration of the machine base [14]–[16], and rolling force [17]–[19]. It was found that machine learning models can be used instead of mathematical models to eliminate irregular vibrations of the machine [14]–[16]. Starting from the characteristics of hot-rolled strip steel production data, Li *et al.* [16] used systematic clustering to determine the number of clusters and then used the K-means algorithm to divide the production data into K clusters. Each data cluster is used to build its own BP neural network prediction model, PSO (Particle Swarm Optimization) is used to optimize the network parameters to avoid the neural network from falling into them. While the stability of rolling force is one of the key parameters affecting the quality of rolled parts, Wang *et al.* [18] proposed two rolling force prediction methods combining improved PSO and BP (Back Propagation) neural networks in order to improve the prediction accuracy of rolling force during dynamic rolling in reversible cold rolling mills, using a large amount of actual data as the neural network training input and fully considering the effect of influence between the input parameters.

During the production of electrode electrodes, the roll rolling force stability and the thickness consistency of the electrodes are the keys to improving the production efficiency [20]–[22]. Wang *et al.* [21] studied a mill rolling force prediction model based on an improved support vector machine. A least-squares support vector machine based on RBF kernel function and polynomial kernel function was established, and the parameters of the hybrid function were optimized using a cooperative quantum particle swarm algorithm to improve the prediction performance of the prediction model. The electrode thickness control system of lithium batteries has the characteristics of nonlinearity, uncertainty, and time variation, Xu *et al.* [22] used a genetic algorithm and backpropagation method to optimize the neural network and established the thickness control prediction model, which eliminated the nonlinearity and uncertainty

problems due to mechanical vibration, and the RMSE was within 6.72, which effectively improved the accuracy and precision of prediction.

In contrast to machine learning algorithms that require manual input of features, deep learning algorithms can automatically select features that affect the prediction results of the model [23]–[25]. Deep learning can easily input problem-specific constraints into the model, which can easily reduce model bias [25]. RNN (Recurrent Neural Network) is a special type of neural network model that has the ability of short-term memory, which preserves the association between data by “remembering” relationships [25]. Liang *et al.* [26] used LSTM and Encoder-Decoder LSTM models to predict the thickness of coal seams. When the hyperparameters were not optimized, the error of the model’s coal seam thickness prediction results was both large. The introduction of an expert knowledge base and optimization of hyperparameters such as the number of neurons, training period, and input-output sequence step size can effectively fit the real coal seam thickness true values. Relying only on data-driven models to predict the operating state tends to ignore the characteristics of the mechanical equipment itself. Zhang *et al.* [27] used a theoretically simplified rolling force model and a bounce model based on the theory to determine the prediction model input parameters and established a DBN-LSSVM hybrid model to predict the strip thickness, and the prediction accuracy was further improved compared with the conventional deep confidence network model.

The essence of model-based methods is to predict the system output by constructing an observer and to determine the fault based on the residuals between the predicted and observed values [28]. This type of approach is more intuitive in terms of the predicted response of the system, but with the increasing complexity of the structure and working mechanism of mechanical equipment and the increase in random disturbances, the complexity of modeling increases significantly. Therefore, model-based methods have many constraints in complex engineering application scenarios and can no longer adapt to the current trend of mechanical equipment [29]. The data-driven approach analyzes the implied laws from historical data and trains the model through continuous iterations to make its “input-output” fit the distribution characteristics of historical data [30]. Given this feature, the data-driven dynamic prediction and diagnosis technology is largely free from the reliance on a priori knowledge, without the need to establish complex physical models, only need to select the appropriate parameters from the sensor monitoring data center and supporting the corresponding learning algorithm to achieve accurate early warning. However, there is a contradiction between computational speed and accuracy in the practical application of online prediction methods that rely solely on physical or data models for frequency dynamics. Hao *et al.* [31] established a mathematical model based on the influence function method to accurately obtain the thickness distribution during the angle rolling process. After that,

a real-time dynamic prediction model was established relying on experimental data. The model can be applied to the actual production process to develop effective plate shape control strategies and flexible rolling protocols to meet the various thickness requirements for custom production. Wang *et al.* [32] proposed a method for online prediction of frequency dynamics based on the idea of physical-data fusion modeling: the transient frequency influencing factors were divided into critical and non-critical factors, the system frequency response model was applied to the critical factors to preserve the causal link between electrical information, and the error correction model based on the limit learning machine was applied to the non-critical factors to characterize the correlation. Therefore, we can introduce the nonlinear deformation formulas (including boundary and initial conditions) in battery electrode production into the data-driven model as training constraints, which can achieve the goal with fewer data samples.

In summary, the electrode thickness, as an important indicator to assess the operating status of the electrode mill, needs to be monitored in real-time. In order to solve the problems of wasted industrial resources and increased costs due to inconsistent sheet thickness, a more advanced prediction method is needed to predict the machine operating status for some time in the future. At present, the traditional method of predicting the thickness of lithium battery electrodes can no longer meet the demand for high precision and still has the following problems:

1. Many current models can often only analyze data purely based on data-based machine learning, without actually considering the underlying physical mechanics more [32];
2. The processing of Li-ion battery electrodes is a very complex non-linear system, and even two identical electrodes may react differently when subjected to the same environment;
3. We can predict the behavior of nonlinear systems based on deep learning algorithms, but it is not easy to train and validate this system, often requiring large amounts of data.

In response to the above problems, we adopt a physical-data fusion modeling prediction method based on the coupled multimorph model and Bayesian LSTM to achieve dynamic prediction of the thickness of lithium battery electrodes, ensure the consistency of the electrode thickness, and improve the operating efficiency of electrode production. Firstly, we establish the underlying physical model of the roll to further obtain the specific parameters affecting thickness control and reduce the redundancy and complexity of the data; secondly, we use LSTM and Bayesian LSTM models to predict the electrode piece thickness under variable working conditions, improve the robustness of the prediction model of the nonlinear control system, and solve the problems of slow approximation speed and easy to fall into local minimization of the traditional neural network. The details of the research are as follows:

Firstly, the two upper and lower rolls of the mill will give pressure to the electrode piece and produce a plastic deformation state, which makes the thickness of the electrode piece change, while the rolls are subjected to the reaction force of the electrode piece and the bouncing phenomenon occurs. In addition, the calculation accuracy of the rolling force directly affects the quality and thickness of the lithium battery electrode piece. Therefore, we establish a multi-deformation coupled physical model based on the roll stiffness deformation model, elastic bending model, and rolling force model.

Second, the ordinary neural network predictive control model has the following problems:

- (1) Uncertainty about the initial connection and the size of the threshold;
- (2) Slower convergence of errors;
- (3) Easy to fall into the problem of local miniaturization;
- (4) Variable choice of network structure;
- (5) The contradiction between prediction ability and training ability.

Therefore, we use LSTM deep learning model to solve the problem of slow convergence and easy to fall into local minimization. However, the number of hyperparameters of LSTM is large, and the prediction accuracy of the model is poor under the influence of noisy data. Further, we use the Bayesian algorithm to optimize the LSTM model and suppress the noise effect of the data. We use the variable parameters of the deformation coupling model as the input variables of the prediction model, and the predictions based on the selected input data are fused by a nonlinear fusion network, which effectively balances the predictive and training capabilities of the model. In addition, we use MSE, RMSE, MAE, R, and R^2 metrics to evaluate the convergence ability and fitting accuracy of the neural network.

Finally, we compare other deep network models with the method proposed in this paper. The prediction accuracy between the GA-BP model, DBN-LSSVM model, LSSVM model, GRNN model, and the prediction model proposed in this paper is mainly compared and analyzed. The advantages of deep learning models in data feature self-extraction and model generalization generalizability are analyzed. Compared with other neural network models, Bayesian LSTM has better generalizability for small sample data.

Our proposed multi-deformation coupled model and data-driven method for battery electrode thickness prediction control method have the following innovations:

1. In order to improve the prediction accuracy of lithium battery electrode thickness, maintain the thickness consistency and improve the mill operation efficiency. In this paper, we innovatively propose a distributed prediction model with multi-deformation coupled model and Bayesian LSTM, in which the variable parameters of the deformation coupled model are used as the input variables of the prediction model, and the predictions based on selected input data are fused by a nonlinear fusion network, which effectively balances the prediction ability and training ability of the model.

The requirement of large data volume is reduced based on improving convergence ability and fitting accuracy.

2. To solve the contradiction between computational speed and accuracy of traditional data-driven online prediction methods in practical applications. In this paper, we propose a multi-deformation coupling model of rolls and a rolling force model, which introduces the nonlinear deformation formula (including boundary and initial conditions) in the production of battery electrodes into the data-driven model as training constraints, which can achieve the target with fewer data samples.

3. Under variable working conditions, the uncertainty of LSTM hyperparameters leads to poor prediction results. To solve this problem. In this paper, we propose a Bayesian statistical approach to optimize the LSTM hyperparameters, which reduces the influence of noisy data on the prediction results. We use MSE, RMSE, MAE, R, and R^2 metrics to evaluate the convergence ability and fitting accuracy of the neural network.

4. To further analyze the applicability of the Bayesian LSTM model, we compare other deep network models with the method proposed in this paper. The results show that the Bayesian LSTM has better generalization in small sample data compared with other neural network models. It is also more suitable for future prediction of the operating state of a double-roller lithium battery electrode mill.

This paper is organized as follows: In Section 2, a coupled model of the electrode mill roll deformation is developed, including a stiffness deformation model, an elastic bending model, and a rolling force model. In Section 3, the distributed LSTM predictor and the MLP-based Bayesian LSTM depth fusion predictor are proposed. In Section 4, we conduct experiments to verify the accuracy of our proposed prediction models.

II. DEFORMATION COUPLING MODEL

In the rolling process, there are two main reasons for changes in the thickness of lithium battery electrodes: one is the reason for the rigidity of the electrode mill itself; the other is the reason for the battery electrodes themselves. While the mill performance changes are often the main factor, the battery electrode strip is generally a secondary factor. Therefore, we study lithium battery electrode mill changes on the impact of the electrode thickness. The following specific introduction of mill stiffness deformation model, roll elastic deformation model, and rolling force model.

A. STIFFNESS DEFORMATION MODEL

When the electrode mill stiffness is small, the frequent fluctuation of the rolling force will affect the accuracy of the electrode thickness. At the same time, because the rolls work for a while along the roll surface wear distribution is not uniform, which also causes the change of the mill stiffness, so the electrode mill design generally uses a high stiffness design.

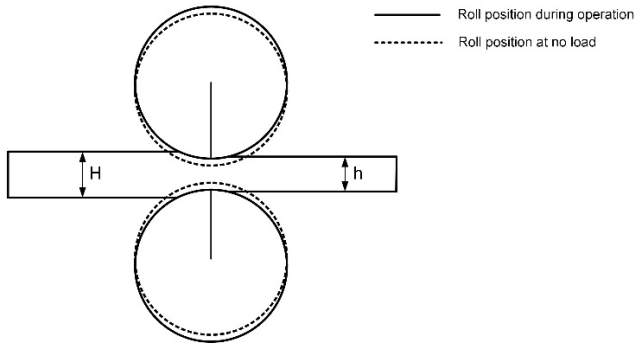


FIGURE 2. Seat elastic deformation.

The overall elastic deformation of the mill mainly includes the elastic deformation of the mill plate [33], the elastic deformation of the roll bearing seat, the elastic deflection, and flattening of the rolls, etc. The overall elastic deformation of the mill, which includes the deformation of various components, is collectively referred to as the mill bounce. Seat elastic deformation is shown in Fig.2.

During the rolling process of the lithium battery electrode, the two upper and lower rolls of the mill will give pressure to the electrode, producing a plastic deformation state, which makes the thickness of the electrode change, while the lithium battery electrode will act on the opposite force of the roll, producing an elastic deformation state, making the roll gap larger and the bouncing phenomenon occur [34]. Combined with the mill operating parameters to establish the following bouncing model:

$$S = S_0 + \frac{F_P - F'_P}{K} + S_r + S_w + S_e + \Delta_0 \quad (1)$$

where, S denotes the actual roll gap of the roll, mm . S_0 denotes the original roll gap of the roll, mm . F_P denotes the preset rolling force, KN . F'_P denotes the preset rolling force, KN . K denotes the mill stiffness coefficient. S_r denotes the compensation amount of the roll wear, mm . S_w denotes the compensation amount of the electrode width, mm . S_e denotes the compensation amount of the thermal expansion of the roll, mm .

B. ELASTIC BENDING MODEL

To simplify the analysis process, we simplify the roll to a solid stepped axis and use the reference system method to establish the deflection and angle synthesis relations of the beam model in multiple reference systems. The simplified model is shown in Fig.3.

As shown in Fig.3, respectively, take the roll left journal midpoint, the two end faces of the roll surface, the roll right journal midpoint, for four cross-sections: S_1, S_2, S_3, S_4 . The lengths between two sections are L_1, L_2, L_3 . The flexural stiffness of the three segments are: EI_1, EI_2, EI_3 . The support reaction force at P is R_P and the shear force at S_2 is Q_B . In the dynamic system I, Q_B and R_P can produce a relative deflection of the section S_4 in the beam model concerning

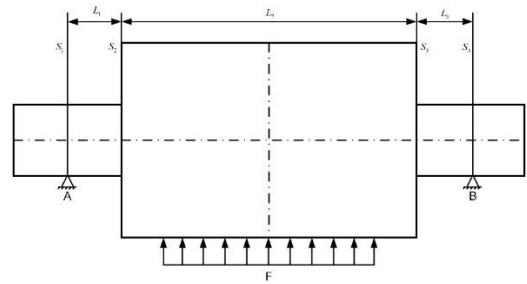


FIGURE 3. Simplified model of rolls.

section S_1 , which is set as $f_{r_1(Q)}^{S_4}, f_{r_1(F)}^{S_4}$, its calculation formula is as follows:

$$f_{r_1(Q)}^{S_4} = -\frac{Q_B L_1^3}{3EI_1} - \frac{Q_B (\frac{L_2}{2})^3}{3EI_2} - \frac{Q_B L_3^3}{3EI_3} - \frac{Q_B L_1}{EI_1} \times (L_1 + L_2 + L_3)(L_2 + L_3) - \frac{Q_B L_2}{EI_2} (L_2 + L_3)L_3 \quad (2)$$

$$f_{r_1(F)}^{S_4} = \frac{F L_1^3}{3EI_1} + \frac{F (\frac{L_2}{2})^3}{3EI_2} + \frac{F L_1}{EI_1} (L_1 + \frac{L_2}{2}) \frac{L_2}{2} + (\frac{F L_1^2}{2EI_1} + \frac{F (\frac{L_2}{2})^2}{2EI_2} + \frac{F L_1}{EI_1} \cdot \frac{L_2}{2}) \cdot (\frac{L_2}{2} + L_3) \quad (3)$$

Then, the total relative deflection of section S_4 in the beam model for the dynamic system I am:

$$f_{r_1}^{S_4} = f_{r_1(Q)}^{S_4} + f_{r_1(F)}^{S_4} \quad (4)$$

The cross-section B is at the support B, so its absolute deflection is $f_{a_1}^B = 0$. The static reference system is represented by the support B [35], Then the absolute deflection of section B is equal to the sum of the implicated deflection of the section by the dynamic system I and the relative deflection of section B to the dynamic system I, i.e. $f_{a_1}^B = f_{e_1}^B + f_{r_1}^B$, and the implicated deflection is $f_{e_1}^B = \theta_A \cdot (L_1 + L_2 + L_3)$. The shear force and bending moment at the section where the maximum deformation of the roll is taken, and the relative deflection of the section m with respect to the dynamic system I fixed to the beam section I is $f_{r_1(Q)}^m, f_{r_1(F)}^m, f_{r_1(M)}^m$:

$$f_{r_1(Q)}^m = -\frac{Q_m L_1^3}{3EI_1} - \frac{Q_m (\frac{L_2}{2})^3}{3EI_2} - \frac{Q_m L_1}{EI_1} (L_1 + \frac{L_2}{2}) \frac{L_2}{2} \quad (5)$$

$$f_{r_1(F)}^m = f_{r_1(F)}^B \quad (6)$$

$$f_{r_1(M)}^m = \frac{M_m L_1^2}{2EI_1} + \frac{M_m (\frac{L_2}{2})^2}{2EI_2} + \frac{M_m L_1}{EI_1} \cdot \frac{L_2}{2} \quad (7)$$

So:

$$f_{r_1}^m = f_{r_1(Q)}^m + f_{r_1(F)}^m + f_{r_1(M)}^m \quad (8)$$

And the implicated deflection of the dynamic system I to the section m is:

$$f_{e_1}^m = \theta_A \cdot (L_1 + \frac{L_2}{2}) \quad (9)$$

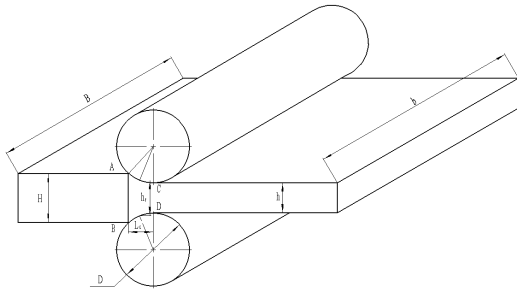


FIGURE 4. Rolling process plastic deformation area.

TABLE 1. Deformation amount classification.

Absolute deformation	Relative deformation
Δh	η_h
Δl	η_l
Δb	η_b

Then the deformation of the section m is:

$$f_{a1}^m = f_{e1}^m + f_{r1}^m \quad (10)$$

Therefore, in a certain dynamic system, the deformation of the roll deformation section depends on some force parameters, which are set as input parameters of the prediction model in the subsequent establishment of the prediction model using the equivalent transformation.

C. ROLLING FORCE MODEL

The rolling pressure model is used to calculate the rolling pressure in the production of the electrode, and it is also a very important variable in the mill control system, and its calculation accuracy directly affects the quality and thickness of the lithium battery electrode. We mainly study the area where the plastic deformation occurs between the roll and the lithium battery electrode piece in contact with the force, as shown in Fig.4, to form a more ideal rolling model.

1) PLASTIC DEFORMATION

In the plastic deformation region, we defined the absolute and relative deformations [36], as shown in Table 1.

Where, Δh refers to absolute depression, Δl refers to absolute extension, Δb refers to absolute spreading, η_h refers to depression rate, η_l refers to extension rate, η_b refers to spreading rate.

Δh is the result of subtracting the exit thickness from the entrance thickness of the lithium battery electrode, which is defined as:

$$\Delta h = H - h \quad (11)$$

Δl is the result of subtracting the exit height from the entrance height of the lithium battery electrode, which is

defined as:

$$\Delta l = L_H - l_h \quad (12)$$

Δb is the result of subtracting the exit width from the entrance width of the lithium battery electrode, which is defined as:

$$\Delta b = B_H - B_h \quad (13)$$

η_h is the percentage of the pressed down the amount of the lithium battery electrode to the entrance thickness, which is defined as:

$$\eta_h = \frac{\Delta h}{H} \times 100\% \quad (14)$$

η_l is the extension of the lithium battery electrode as a percentage of the entrance height, which is defined as:

$$\eta_l = \frac{\Delta l}{L_H} \times 100\% \quad (15)$$

η_b is the percentage of the widespread of the lithium battery electrode to the entrance width, which is defined as:

$$\eta_b = \frac{\Delta b}{B_H} \times 100\% \quad (16)$$

From expert experience, the volume of the lithium battery electrode does not change before and after rolling, so:

$$HL_H B_H = hl_h b_h \quad (17)$$

Deforming Equation (17), we can get:

$$\frac{HL_H B_H}{hl_h b_h} = 1 \quad (18)$$

where: H indicates the entrance thickness of the electrode; h indicates electrode exit thickness; L_H indicates the entrance height of electrode; l_h indicates the exit height of electrode; B_H indicates the entrance width of the electrode; b_h indicates the exit width of electrode.

2) FRONT-SLIP AND BACK-SLIP

During the rolling process, the thickness of the electrode decreases gradually, but its rolling speed is slowly increasing, resulting in a velocity difference between the electrode and the roll. When the velocity of the electrode mass located in the plastic deformation zone is larger than the velocity of the horizontal component of the roll speed, the front slip phenomenon is generated; when the velocity of the electrode mass is larger than the horizontal component of the roll speed, the back slip phenomenon is generated. The front-slip and back-slip zones allow the rolling process to continue in an equilibrium state. Assuming that the width of the electrode does not change and no deformation occurs in the rolling, the relationship between the thickness of the electrode before and after rolling and the speed is:

$$v_H H = v_h h \quad (19)$$

where, v_H denotes the roll speed, v_h denotes the velocity of the electrode.

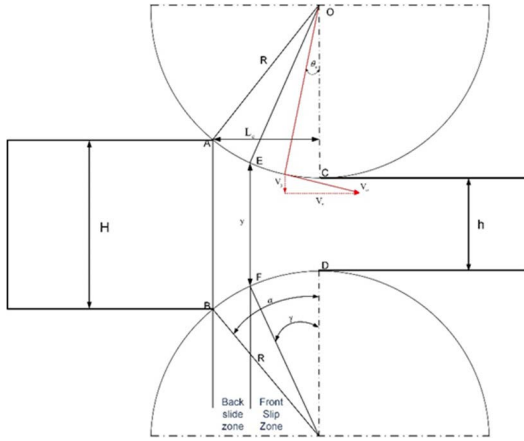


FIGURE 5. Schematic diagram of the front slip zone I and the back slip zone II.

As shown in Fig.5, we define the front-slip zone I and the back-slip zone II. A neutral surface exists in the middle of zone I and zone II, indicating that the horizontal speed of the roll corresponding to this surface is equal to the horizontal speed of the electrode ($v_H \cos \alpha = v_h$). The corresponding roll point on this side is the neutral point F , the angle of the line with the center of the roll is the neutral angle γ [37], [38].

f_h is the difference between the electrode speed and the electrode rolling speed and the percentage of the roll speed at the exit of the electrode roll, which is defined as:

$$f_h = \frac{v_h - v}{v} \times 100\% \quad (20)$$

f_H is the difference between the speed of the horizontal direction of the roll and the rolling speed and the percentage of the horizontal speed of the roll, which is defined as:

$$f_H = \frac{v \cos \alpha - v_h}{v \cos \alpha} \times 100\% \quad (21)$$

where: f_h denotes front-slip value; f_H denotes back-slip value; v denotes rolling speed;

Therefore, we need to consider the front-slip value when using continuous rolling with tension. The front-slip and back-slip values need to be considered when calculating the torque required to rotate the rolls and the tension between the stands. In addition, the front-slip value needs to be calculated when adjusting the mill, otherwise, the tension will be too high and the electrodes will be pulled off.

D. THICKNESS CONTROL INFLUENCING FACTORS

According to the first three sections, we can obtain the expression for the polymorphic coupling model for the thickness of the lithium battery pole piece:

$$Q = \sqrt{\frac{1 - \eta_h}{\eta_h}} \left[S + \frac{1}{2} \sqrt{\frac{B_H}{b_h}} \ln \frac{1}{1 - \eta_h} \sqrt{\frac{H}{h}} \ln \frac{L_H}{l_h} + \frac{\pi}{2} \arctan(\theta_A - \theta_B) + \frac{T_h - f_h}{K \cdot v_h} - \frac{\gamma}{K} \sqrt{\frac{L_2}{\Delta h}} \right] \quad (22)$$

From the above equation, it can be seen that the electrode thickness formula is complicated and the accuracy cannot be guaranteed, which is not conducive to practical application. Therefore, when the traditional pure mathematical model is used to predict the electrode thickness, the maximum deviation will reach more than 20% [32], this has not been able to meet the demand in actual production. We summarize the factors affecting thickness control according to the previous article as:

- The thickness of the electrode entrance and the thickness of the electrode exit, whose thickness variation ultimately affects the actual thickness of the electrode.
- The variation of the rolling speed of the battery electrode sheet. It directly affects the size of the front-slip and back-slip, the size of the rolling force in the plastic deformation zone at the exit of the electrode sheet, and the actual thickness of the electrode sheet.
- Variation in the tension of the lithium battery electrode sheet. In the winding and unwinding mechanism of a rolling mill, the rolling tension can change the stress state of the electrode and ultimately the deformation resistance of the electrode [18]. The unwinding tension and the winding tension can fine-tune the thickness of the electrode sheet.
- The variation of roll gap. The electrode will be subject to friction and plastic deformation between the roll and the electrode piece during the rolling process, resulting in a larger roll gap, which affects the actual thickness of the electrode.

Through specific analysis of the rolling process and control system, we can get that H (Thickness of the electrode entrance), h (Thickness of the electrode exit), B_H (Width of the electrode entrance), v (Rolling speed), F_P (Rolling force), T_h (Unwinding tension), T_H (Winding tension), S (Actual roll gap) are the main influencing factors that cause the fluctuation of the thickness of the lithium battery electrode. In subsequent sections, the control parameters affecting thickness determined by the deformation coupling model are used as input variables for the LSTM prediction model and the Bayesian LSTM prediction model, ultimately enabling real-time prediction of cell thickness. The deformation coupling model can effectively simplify the input parameters of the prediction model, reduce the complexity of the model, and improve the efficiency of model training.

III. LSTM MODEL AND ITS VARIANTS

A. LSTM MODEL

When the network is too deep, standard RNNs tend to suffer from long-time dependence and gradient disappearance [25], [39], [40]. In other words, when the time step is too large, the information carried by the preceding neurons is lost, because no structure in the standard recursive layer can control the flow of memory itself alone. To solve this problem, long and short-term memory networks have been proposed, which are improved recursive electrodeular structures whose differences from RNNs are shown in Fig.6. The network

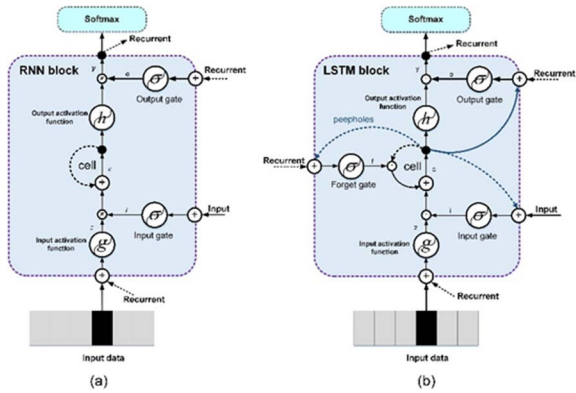


FIGURE 6. (a) RNN neural network model; (b) LSTM neural network model.

structure of LSTM is much more complex than that of RNN. Microscopically, the LSTM adds three gate structures and a unitary state structure to the RNN structure to achieve the function of remembering the information state values at any past moment [41].

In order to better characterize the coupling relationship between variables, we adopt a single-to-single structure LSTM model, and use the training error of each category of variables as the input bias of the next category of variables, and the model structure is shown in Fig.7. The gate structures are: the oblivion gate, input gate, and output gate, which are responsible for different functions in the LSTM. Among them, the forgetting gate is mainly responsible for information filtering, determining the trade-off of information, and this decision is made by the "sigmoid" function in the forgetting gate [40]; The input gate is the information selection storage unit in the LSTM, which is responsible for selecting the appropriate and useful information to be retained in the LSTM; the output gate mainly determines the output of the neural network.

The LSTM uses storage elements to transfer information from past outputs, rather than making the output of the RNN unit a nonlinear function of the weighted sum of the current input and the previous output [41], [42]. In other words, instead of using only the hidden state H , the LSTM uses the electrode state C to preserve long-term information. LSTMs mainly uses three gates (forgetting gate, input gate, and output gate) to control the electrode state C . The forgetting gate is used to control from previous electrode states C_{t-1} to the information of current electrode states C_t ; The input gate determines how much input should be held in the C_t ; The output gate determines the output H_t according to C_t .

Compute the LSTM output for step t using the following equation:

$$F_t = \sigma(W_f \cdot [H_{t-1}, X_t] + b_f) \quad (23)$$

$$I_t = \sigma(W_i \cdot [H_{t-1}, X_t] + b_i) \quad (24)$$

$$\tilde{C}_t = \tanh(W_c \cdot [H_{t-1}, X_t] + b_c) \quad (25)$$

$$C_t = F_t * C_{t-1} + I_t * \tilde{C}_t \quad (26)$$

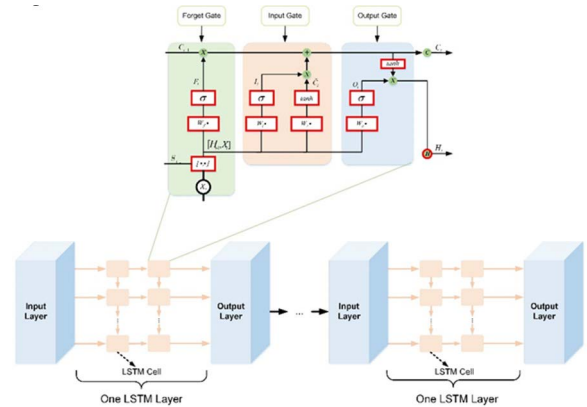


FIGURE 7. Block diagram of the one-to-one LSTM model.

$$O_t = \sigma(W_o \cdot [H_{t-1}, X_t] + b_o) \quad (27)$$

$$H_t = O_t * \tanh(C_t) \quad (28)$$

Among them, ' \cdot ' represents the inner product between variables, ' $*$ ' represents the outer product between variables, W and b are trainable weights and biases, respectively, and I , F and O represent input gates, forgetting gates and output gates, respectively. These three doors have the same shape, $[H_{t-1}, X_t]$ different, and need to be learned in the training process. The candidate state \tilde{C}_t cannot be used directly, it must pass through the input gate and then be used to calculate the internal storage C_t . And C_t is not only affected by the hidden state, but also by C_{t-1} , which is controlled by the forgetting gate. On top of C_t , a layer of \tanh functions is applied to the output information H_t , which is bounded by the output gate. The presence of gates allows the LSTM to fulfill long-term dependencies in sequences, and by learning the parameters of the gates, the network can find the appropriate internal storage [25].

B. BAYESIAN LSTM MODEL

The parameters in the LSTM model are divided into two main categories, one is the model parameters, which are updated by the gradient descent algorithm during training, and the other is the hyperparameters, which are generally fixed values or vary with predefined rules during training, such as batch size, learning rate, weight decay, and gamma in kernel functions, etc [43]. The goal of hyperparameter tuning is usually to minimize the generalization error, but other optimization goals can be customized for specific tasks. Bayesian conditioning uses a continuously updated probabilistic model to "focus" promising hyperparameters by extrapolating past results and is suitable for optimization on spaces smaller than 20 dimensions. Through Monte Carlo sampling, Bayesian deep learning networks train the network several times and take the average of all losses, which is then used for back-propagation to obtain the distribution of weights and biases.

In a normal LSTM network, the parameters including all weights and biases are trained constants. Bayesian LSTM

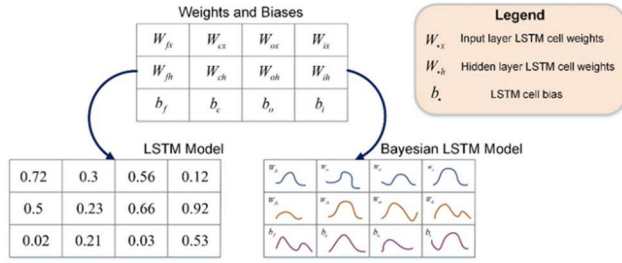


FIGURE 8. The difference between the normal LSTM network and the Bayesian LSTM network.

can treat the weights and deviations as a random distribution that is constantly updated with iterations. Each parameter obtained through Bayesian LSTM network training is based on the mean and variance of the weight and bias distribution [44], [45]. Fig.8 shows the differences between the ordinary LSTM network and the Bayesian LSTM network. The ordinary LSTM network obtains a certain weight and bias after training. In contrast, the Bayesian LSTM obtains weights and biases after training as a distribution with mean and variance.

Assuming training data D , Bayesian inference can be used to compute the posterior distribution of the weights $P(w|D)$ [44]. The predicted distribution of the input data x is given by $P(y|x) = E_{P(\theta|D)} [P(y|x, \theta)]$. Until now, it is still difficult to find a suitable collection function. Variational approximation of the Bayesian posterior distribution of the weights is a feasible approach. Variational learning finds the parameter (μ, σ) of a distribution overweights $q(\theta|\mu, \sigma)$ that minimizes the Relative Entropy (RE) [45]. The true Bayesian posterior probability is:

$$(\mu, \sigma) = \arg \min_{\mu, \sigma} RE[q(\theta|\mu, \sigma)||P(\theta|D)] \quad (29)$$

According to Bayesian theory:

$$P(\theta|D) = \frac{P(D|\theta)P(\theta)}{P(D)} \quad (30)$$

and RE, equation (29) can be transformed into:

$$(\mu, \sigma) = \arg \min_{\mu, \sigma} \int q(\theta|\mu, \sigma) \log \frac{q(\theta|\mu, \sigma)}{P(D|\theta)P(\theta)} d\theta \quad (31)$$

For the evaluation of the effect of modeling the feature dataset, we define the cost function:

$$L = \int q(\theta|\mu, \sigma) \log \frac{q(\theta|\mu, \sigma)}{P(D|\theta)P(\theta)} d\theta \quad (32)$$

For the variance to be expressed as a non-negative real number, we make the following definition:

$$\sigma = \log(1 + e^p) \quad (33)$$

$$\theta = \mu + \log(1 + e^p) \otimes \varphi, \varphi \sim N(0, 1) \quad (34)$$

After that we can get:

$$L = \log q(\theta|\mu, \sigma) - \log P(\theta) - P(D|\theta)P(\theta) \quad (35)$$

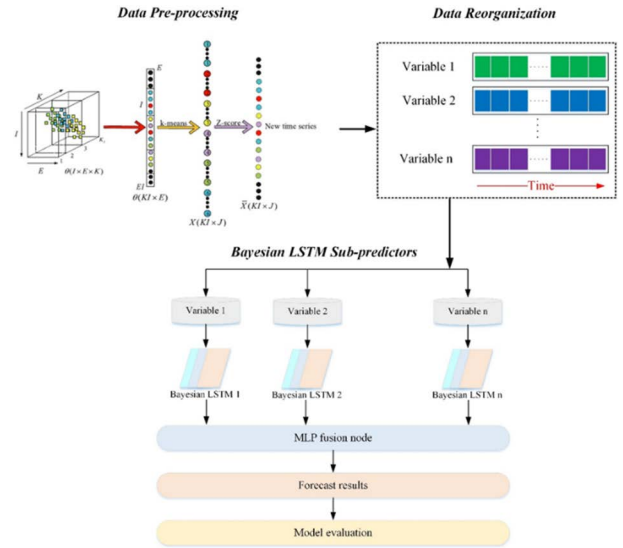


FIGURE 9. Model framework for predicting the thickness of lithium battery electrodes.

$$\frac{\partial}{\partial p} L = \frac{\partial L}{\partial \theta} \frac{\varphi}{1 + e^{-p}} + \frac{\partial L}{\partial p} \quad (36)$$

where $\frac{\partial L}{\partial \theta}$ is the gradient found by the backpropagation algorithm on an ordinary LSTM network. Thus, to understand the mean and standard deviation, we can calculate the gradient by backpropagation and then scale and transform it [46], [47].

C. THICKNESS PREDICTION MODEL FRAMEWORK

We propose a distributed prediction model combining a deformation coupling model and a deep learning network for the thickness prediction problem of lithium battery electrode mills. The model framework is shown in Fig.9, which consists of three main components: data pre-processing, data reorganization, and Bayesian LSTM sub-predictors. Data pre-processing using K-means and Z-score methods to reduce data dimensionality and eliminate redundant data. Data reorganization classifies the processed features with a total of data variables. For n data variables, n Bayesian LSTM sub-predictors are designed. Finally, we use fusion nodes to fuse the predictions of multiple sub-predictors. The fusion node uses an artificial neural network MLP. MLP is a fully connected combination of artificially designed neurons, which applies a nonlinear activation function to model the relationship between input and output. Our proposed thickness prediction model framework is able to mine the “historical information” of the data while fusing the correlations between different feature data.

IV. EXPERIMENTS

A. DATASETS

Our experiments use the data set of the lithium battery electrode thickness control system of a battery equipment

TABLE 2. Raw data of the sample section.

Time Steps	F_P (t)	H (mm)	h (m)	B_H (mm)	v (mm/s)	T_H (N)	T_h (N)	S (mm)	Q (mm)
10	49.9	0.1855	0.1305	500.75	698	149.4	120.1	0.043	0.1253
20	54.2	0.1882	0.1283	499.86	700	149.0	127.3	0.038	0.1255
30	45.4	0.1888	0.1306	500.27	698	143.8	122.9	0.044	0.1251
40	52.6	0.1899	0.1292	500.14	700	141.3	121.8	0.042	0.1268
50	46.7	0.1925	0.1302	500.25	702	140.5	120.8	0.045	0.1258
60	53.0	0.1860	0.1315	500.70	703	144.9	120.4	0.041	0.1253
70	52.9	0.1945	0.1251	500.44	699	143.6	126.4	0.043	0.1265
80	53.5	0.1927	0.1272	499.99	700	142.4	127.4	0.038	0.1264
90	52.4	0.1895	0.1268	499.62	701	149.4	123.6	0.042	0.1266
100	52.1	0.1878	0.1282	500.80	699	142.3	129.2	0.045	0.1246
10	61.7	0.1812	0.1197	500.75	698	149.2	130.7	0.038	0.1197
20	61.9	0.1791	0.1193	499.86	700	152.5	129.9	0.042	0.1229
30	59.3	0.1791	0.1200	500.27	698	151.9	130.1	0.045	0.1184
40	60.5	0.1814	0.1185	500.14	700	152.9	130.1	0.038	0.1193
50	56.4	0.1829	0.1199	500.25	702	148.7	131.4	0.035	0.1204
60	62.4	0.1774	0.1197	500.70	703	150.6	131.4	0.039	0.1202
70	63.9	0.1820	0.1198	500.44	699	152.9	133.3	0.036	0.1217
80	59.1	0.1811	0.1194	499.99	700	153.0	132.6	0.043	0.1201
90	63.4	0.1777	0.1195	499.62	701	151.0	131.5	0.038	0.1220
100	64.8	0.1828	0.1195	500.80	699	151.4	131.6	0.044	0.1186
10	72.1	0.1714	0.1173	499.99	700	154.5	134.2	0.043	0.1156
20	68.1	0.1724	0.1160	499.62	701	155.4	134.6	0.038	0.1144
30	67.6	0.1753	0.1172	500.80	699	155.3	133.7	0.044	0.1157
40	72.3	0.1716	0.1168	500.10	700	154.1	133.7	0.042	0.1151
50	69.5	0.1743	0.1167	499.53	702	154.2	134.2	0.045	0.1166
60	67.2	0.1729	0.1171	500.28	701	154.9	134.6	0.041	0.1161
70	69.2	0.1722	0.1171	500.70	699	155.0	133.1	0.041	0.1170
80	73.0	0.1755	0.1165	500.44	700	154.0	134.6	0.045	0.1166
90	69.8	0.1720	0.1169	499.99	700	155.6	134.0	0.041	0.1154
100	74.6	0.1706	0.1155	499.62	699	155.3	133.5	0.040	0.1159

company in Xingtai, Hebei Province, China. These data include system preset and actual thickness values. The lithium battery electrode thickness was selected for prediction to validate the proposed model. The dataset contains electrode thicknesses under the action of 3 different F_P (45t-55t, 55t-65t, and 65t-75t). The data under different working conditions are collected according to time steps, and there is a strong correlation between the time series data, that is, the data of the previous time step will have an impact on the thickness prediction results of the next time step. Datasets contains a total of 30,000 data samples, and Table 2 shows some of the data. We selected the first 90% of the data for training and the remaining 10% for testing. Between model training, we need to normalize the parameters

to map the original data into the $(-1, 1)$ interval, maintaining the correlation between the input and output data [10].

B. MODEL PARAMETER SETTINGS

We use LSTM deep neural networks and Bayesian LSTM deep neural networks for our experiments. We use Relu as the linear activation function for the Bayesian LSTM layer and the LSTM layer.

For the Bayesian LSTM layer, we set up an MLP layer with the size of each layer set to 24. Supervised training was performed using the Adam algorithm, and the model was trained using small batch sampling. The model hyperparameters, such as learning and batch size, were obtained from experiments, as shown in Tab.3.

TABLE 3. Hyperparameters for the experiments.

Layers	Design Details	Parameters Setting
LSTM	Number of layers: 3	Batch size: 20
	Number of neurons: 96	Epochs: 60
	Sampling number: 4	Learning rate: 0.005
		Learning times: 250
Bayesian LSTM	Number of layers: 6	Batch size: 32
	Number of neurons: 96	Epochs: 100
	Sampling number: 4	Learning rate: 0.001
		Learning times: 125
		Dropout value: 0.5
MLP	Number of layers: 1	
	Number of neurons: 24	

For the LSTM layer, we also use Adam’s algorithm for supervised training. The training-test dataset partitioning is kept consistent with the Bayesian LSTM layer. The hyperparameters of the model are set based on expert experience, as shown in Table 3.

The performance of the model is evaluated by the following five metrics. The mean square error (MSE) can reflect the value of the convergence loss function of the neural network and is defined as:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \tag{37}$$

where \hat{y}_i is the prediction, y_i is the ground truth, and n is the number of data.

The root-mean-squared error (RMSE) is a better way to describe the data and is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \tag{38}$$

The mean absolute error (MAE) and Pearson’s correlation coefficient (R) between the predicted and reference values were also explored in the experiments [48]. In addition, to describe the relationship between the predictor and response variables, we explored the R-squared coefficient (R^2):

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \tag{39}$$

$$R = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^n (y_i - \hat{y}_i)^2 \sum_{i=1}^n (\hat{y}_i - \bar{\hat{y}})^2}} \tag{40}$$

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \tag{41}$$

C. EXPERIMENT I

Under this experiment, the performance of the LSTM model was verified by predicting the electrode thickness

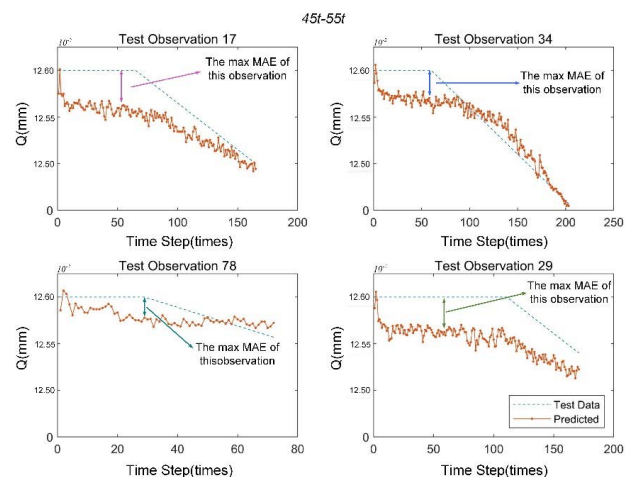


FIGURE 10. LSTM model training progress.

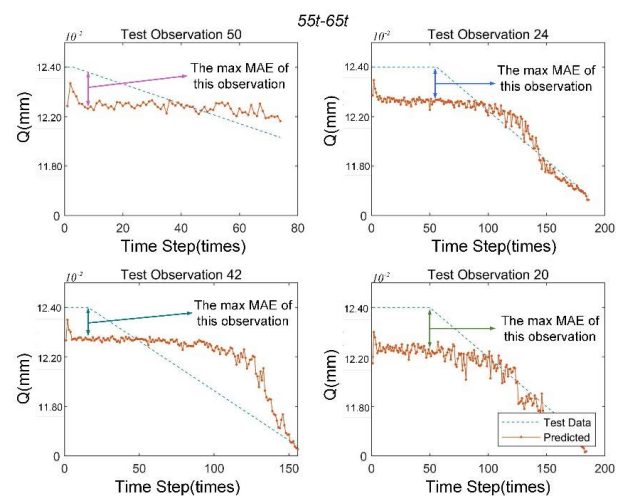
and evaluating the causality. We used 7 variables selected to influence the thickness control as input data for the distributed depth model. The tests are divided into three groups, and there are 45t-55t, 55t-65t, and 65t-75t. The training process of the LSTM model is to optimize the RMSE by continuously iterating. The training process of the LSTM model is shown in Fig.10. Using the 45t-55t condition as an example, we experimentally set the total number of iterations to 250, specified an initial learning rate of 0.005, and reduced the learning rate after 125 training rounds by multiplying by a factor of 0.2. As can be seen from the fig.10, the RMSE is close to stabilization near the 80th generation and is in the 20-23 interval. Therefore, the LSTM model is fast to train and has low approximation error.

We set the maximum time step in the time series to 250, i.e., including 250 observations. The blue and yellow lines indicate the basic facts of the thickness and the predictions of the model, respectively. Fig.11 shows a comparison of the measured data and the results of the 250-step prediction. The best fits were obtained for Observation 17 and Observation 34 at 45t-55t conditions. The poor prediction of Observation 78 is due to the small-time variation trend of the measured data, which leads to a poor fitting effect. The best fits were obtained for Observation 20 and Observation 24 at 55t-65t conditions. While the prediction results of Observation 42 are located on both sides of the measurement results, the reason is that as the number of data increases, the hyperparameters and the amount of data does not fit. Under the conditions of 65t-75t, Observation 84 has the best fitting effect. While Observation 4 and Observation 28 have the same trend as the measured data, but fluctuate more, because the data itself is too noisy, which affects the prediction accuracy. The maximum MAE of most observation points (Observation 20, Observation 24) appears at the data inflection point, which indicates that the data is noisy and the LSTM is not enough to suppress the noise. However, the predicted trend is close to the measured data, and most of the predicted values are within the confidence interval.

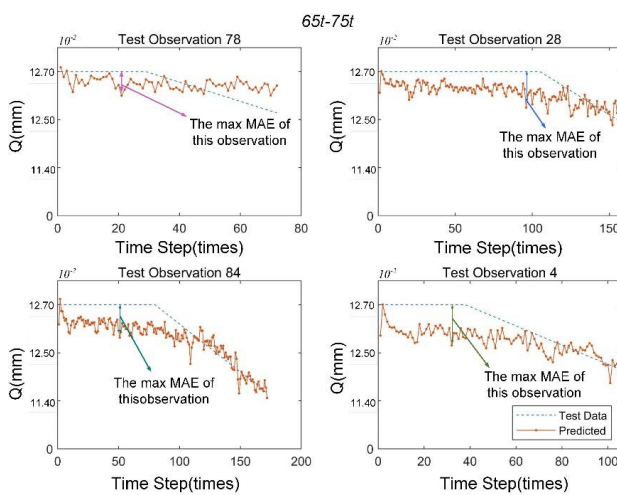
We counted the RMSE of each observation point and the results are shown in Fig.12. The frequency of RMSE=0 for the pole piece thickness samples is the highest in the 45t-55t and 55t-65t conditions, and the error curve has a centrosymmetric distribution. Under the conditions of 65t-75t, the frequency of RMSE of the pole piece thickness



(a)



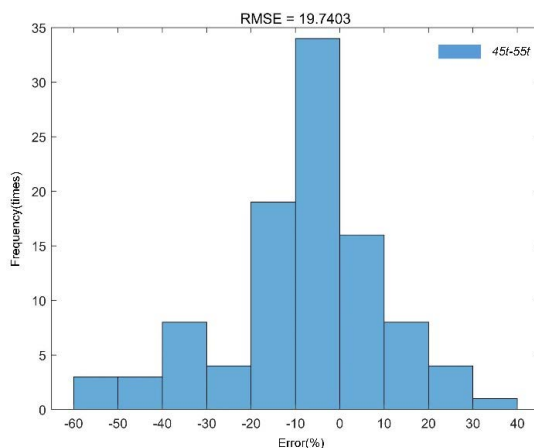
(b)



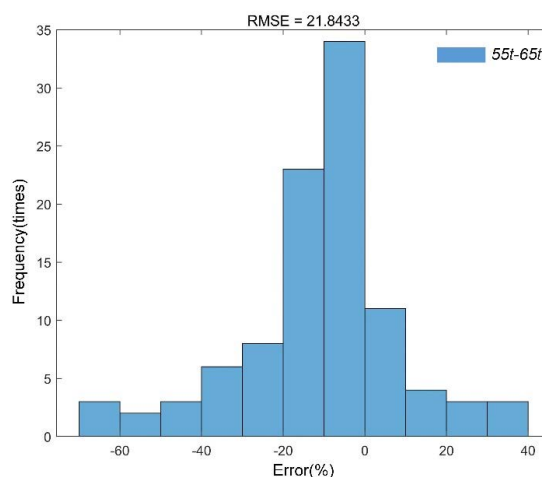
(c)

FIGURE 11. (a) Measured and predicted values of electrode thickness at 45t-55t; (b) Measured and predicted values of electrode thickness at 55t-65t; (c) Measured and predicted values of electrode thickness at 65t-75t.

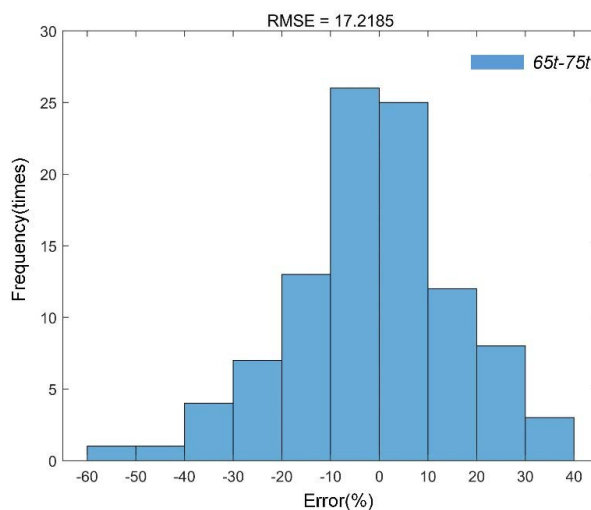
samples is the highest in the interval of $[-10, 10]$, which indicates that the data distribution is not uniform and the noise



(a)



(b)



(c)

FIGURE 12. (a) RMSE for electrode thickness prediction at 45t-55t; (b) RMSE for electrode thickness prediction at 55t-65t; (c) RMSE for electrode thickness prediction at 65t-75t.

is large. We counted additional metrics for the observation sites, as shown in Table 4. We can find that the absolute error

TABLE 4. Prediction performance with different observation.

Condition	Observation	MSE	RMSE	MAE	R	R ²
45t-55t	Observation: 17	29.452	5.427	6.023	0.673	0.532
	Observation: 29	85.252	9.281	10.281	0.682	0.672
	Observation: 34	8.753	2.753	5.753	0.792	0.751
	Observation: 78	145.67	14.495	5.495	0.802	0.766
55t-65t	Observation: 20	113.48	10.653	32.653	0.631	0.623
	Observation: 24	65.607	8.062	24.620	0.812	0.453
	Observation: 42	279.47	19.731	30.731	0.842	0.569
	Observation: 50	225.02	15.902	12.902	0.737	0.602
65t-75t	Observation: 4	52.560	7.673	8.673	0.645	0.752
	Observation: 28	80.073	10.043	15.043	0.743	0.628
	Observation: 78	24.771	5.771	21.771	0.872	0.525
	Observation: 84	214.62	14.602	14.602	0.805	0.481

TABLE 5. Prediction performance with MLP fusion predictor.

Condition	MSE	RMSE	MAE	R	R ²
45t-55t	94.57	9.72	2.708	0.875	0.861
55t-65t	23.72	4.87	2.305	0.860	0.591
65t-75t	5.11	2.25	1.852	0.800	0.322

between predicted and observed values is smaller and the data fits better in the 45t-55t condition; in the 55t-65t condition, although the absolute error is larger, the correlation between the data is stronger. However, the RMSE of the training model is too large to meet the accuracy of pole piece production. Therefore, an optimization algorithm is needed to adjust the LSTM model hyperparameters.

D. EXPERIMENT II

In experiment I, the accuracy of the LSTM prediction model could not meet the production requirements. Therefore, in experiment II, we use the Bayesian LSTM model to predict the pole piece thickness. Under this experiment, we used RMSE to evaluate the error of the training data and MAE, R, and other metrics to evaluate the model accuracy. Experiment II verified the performance of the Bayesian LSTM model and evaluated the causality. The test data is the same as Experiment I. The Bayesian LSTM model uses bidirectional logic parameter control, and the results of each Bayesian LSTM predictor output the final thickness prediction through the MLP network. The prediction results of each sub-predictor after MLP fusion are shown in Fig.13. The blue and black lines indicate the thickness measurement data and the model prediction respectively. The Bayesian LSTM model performs better than the LSTM model when there is a large amount of data noise in each working condition. In Fig. 13 (a), there is spike noise in the data, but the prediction results are not disturbed and the prediction results remain in a stable range; Fig.13 (b) and (c) show that the data from multiple sub-predictors contain a large amount of attribute noise, resulting in a large range of data fluctuations, but the prediction results remain in a stable range.

Similar to experiment I, we counted the MSE and RMSE of the Bayesian LSTM predictor under three working conditions, as shown in Fig.14. The MSE of the predictor under 45t-55t working condition is 94.57 and the RMSE is

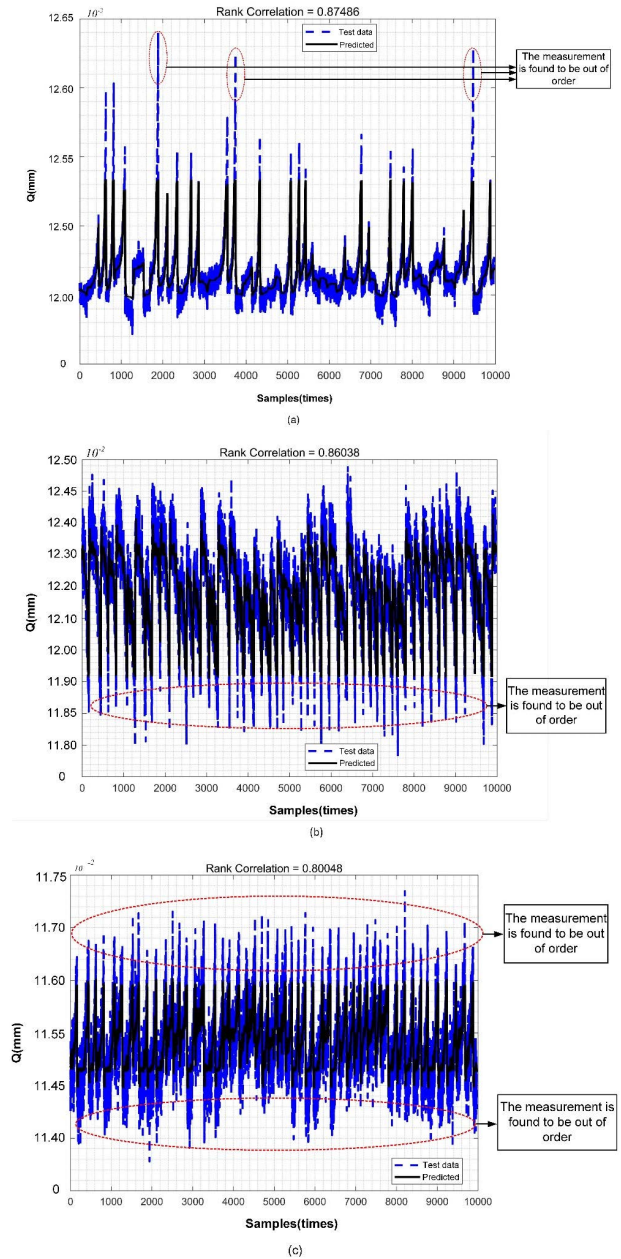


FIGURE 13. (a) Measured and predicted values of electrode thickness at 45t-55t; (b) Measured and predicted values of electrode thickness at 55t-65t; (c) Measured and predicted values of electrode thickness at 65t-75t.

9.72, which indicates that the convergence function error is larger; the MSE and RMSE of the predictor under 55t-65t and 65t-75t working conditions are lower, and the predictor performance is better. In addition, the maximum R and R² values of the proposed Bayesian LSTM model represent the best fit between the predicted and observed values. The results show that the fitted data are highly correlated and can effectively reduce the uncertainty of some of the large-scale parameters or parameter combinations. The specific predictor indexes are shown in Table 5. Compared with the LSTM predictor, the Bayesian LSTM predictor has a high fitting ability and strong data denoising ability, which can effectively

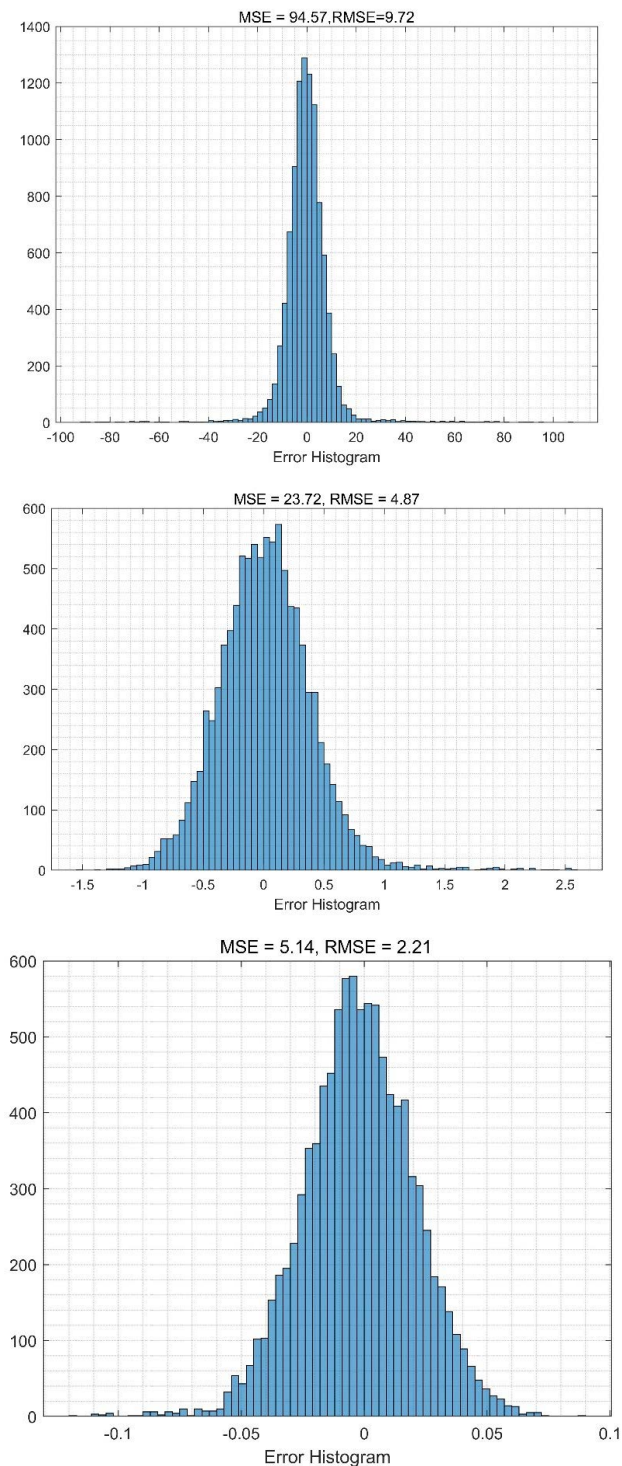


FIGURE 14. (a) RMSE and MSE for electrode thickness prediction at 45t-55t; (b) RMSE and MSE for electrode thickness prediction at 55t-65t; (c) RMSE and MSE for electrode thickness prediction at 65t-75t.

characterize the relationship between the predictor variables and the response variables.

E. EXPERIMENT II

In this experiment, we compared other deep network models with the approach proposed in this paper. As there are few

TABLE 6. Prediction performance with other models.

Model	MSE	RMSE	MAE	R
GA-BP[22]	45.158	6.723	2.708	0.845
DBN-LSSVM[27]	30.404	5.514	2.325	0.841
LSSVM[49]	38.937	6.245	5.282	0.772
GRNN[50]	44.649	6.682	2.539	0.880
The proposed LSTM	29.452	5.427	2.305	0.860
The proposed Bayesian LSTM	26.030	5.102	1.952	0.875

studies related to thickness prediction for lithium battery pole mills, the operating principle and deformation parameters of steel rolling mills are similar to those of lithium battery pole mills. Therefore, we have analyzed and compared the relevant results from the thickness prediction studies of steel rolling mills. Among them, no baseline model includes a feature selection process and uses all features as network inputs. In our previous study, we used a BP neural network prediction model, but the model tended to fall into local minimization. Later, a genetic algorithm was used to optimize the weights and thresholds of the BP neural network [22]. Other model prediction results are shown in Table 6.

We chose other network models including GA-BP [22], DBN-LSSVM [27], LSSVM [49] and GRNN [50]. In order to make the comparison results more accurate, we chose the mean value of the metrics for analysis. The RMSE of chosen models, the proposed LSTM and Bayesian LSTM in this paper were 6.723, 5.514, 6.245, 6.682, 5.427 and 5.102, respectively. The MAEs were 2.708, 2.325, 5.282, 2.539, 2.305 and 1.952. The RMSE of Bayesian LSTM decreased by 6.37% and 30.97%, and MAE decreased by 18.08% and 30.07% compared to LSTM and GRNN, respectively. The results show that the Bayesian LSTM model has the smallest RMSE of 5.102, decreased by 8.07% and 22.40% respectively, and the smallest MAE of 1.952, decreased by 19.11% and 170.59% respectively, compared with other hybrid models such as DBN-LSSVM and LSSVM. The above experimental data show that our proposed model outperforms other models in terms of prediction accuracy and result fitting, demonstrating the applicability of the model in the field of rolled piece thickness prediction in rolling mills.

V. CONCLUSION

This paper mainly introduces a physics-data-driven prediction model for the control of the thickness of lithium battery electrodes. This paper analyzes the current development status of the research on lithium battery electrode mills and electrode thickness prediction control and summarizes the development trend, research content, and shortcomings of electrode thickness prediction control methods. To effectively reduce the difficulty of data acquisition in industrial production and further improve the operation performance and productivity of the electrode sheet mill, this paper establishes a multi-deformation coupling model based on the elastic deformation of the roll and the rolling force. Meanwhile, to improve the robustness of the prediction model of nonlinear control system under variable working

conditions and to solve the problem of slow approximation speed of the traditional neural network, this paper proposes to establish the thickness prediction model by combining physical model and deep learning method and to optimize the LSTM neural network by using Bayesian theory.

In building the physical model, we analyze the influence of the elastic deformation of the roll on the thickness by starting from the key components affecting the thickness of the electrode and establishing the stiffness deformation model and elastic bending model. Meanwhile, the rolling force is also one of the factors affecting the thickness of the electrode. We define the plastic deformation region of the rolling process, determine the amount of plastic deformation between the roll and the electrode, and also analyze the effect of front slip and back slip on the consistency of the electrode. We obtained mathematical expressions to control the pole piece thickness production and determined the influencing factors affecting the electrode thickness control.

In building the physics-data-driven prediction model, we introduce the nonlinear deformation formula in battery electrode production into the data-driven model as a training constraint, considering the influence of the physical mechanics factors of the underlying pole mill on the prediction accuracy. Meanwhile, we use the LSTM predictor to learn the prediction for long time-series data. We find that the LSTM model does not have high prediction accuracy and does not suppress noise well under variable working conditions and insufficient data volume. Therefore, we use Bayesian theory to optimize the hyperparameters of the LSTM neural network, modeled by weighted sampling, and employ MLP to fuse multiple sub-predictors to obtain a more stable output. The Bayesian LSTM can automatically find the optimal hyperparameters under different operating conditions and can eliminate redundant data, which reduces the requirement of large data volume while improving the convergence capability and fitting accuracy. The RMSE, MSE, MAE, R, R^2 of the Bayesian LSTM model are 5.11, 2.25, 1.852, 0.800, and 0.322, respectively. It has significantly lower errors and better characterizes the correlation between data compared with other prediction models.

Although we have improved the thickness prediction accuracy of lithium battery pole pieces and evaluated the correlation between different sensor data, there are still some problems that we need to solve further:

1. The working environment of the lithium battery pole piece rolling mill is complex. The related parameters of the pole piece thickness should also include the original density, coating quality and environmental factors. The thickness prediction model should add more nonlinear factors to improve its robustness of the thickness prediction model.

2. The deformation coupling model proposed in this paper can reduce the network's input to a certain extent and realize feature dimension reduction. However, this method cannot remove the redundant noise of the sensor and cannot extract the high-dimensional hidden features of the data. We can seek to adopt some feature extraction algorithms suitable

for real-world industrial data and consider the coupling relationship between different features.

3. Although the Bayesian network can estimate the model uncertainty, it relies on prior knowledge and is computationally expensive. Other optimization algorithms can be sought to improve the uncertainty expression ability of the model.

With the demand for industrial production, the lithium battery pole piece rolling mill is gradually converted from the single-rolling mode to the double-rolling mode, and the process technology will be more complicated. Therefore, we will continue to study the coupling relationship between the double rolls and the battery pole pieces, propose a deformation coupling model that is more in line with the actual operating conditions, and continuously improve the thickness prediction model. The method proposed in this paper can be combined with other parameter estimation algorithms to study nonlinear parameter identification of different disturbances of lithium battery pole piece rolling mills and applied to other fields such as signal processing and process control systems.

Data Availability: The code and data used to support the findings of this study have been deposited in the battery electrode thickness using polymorphic variable coupling model and data-driven method repository and can be obtained from the corresponding author upon request.

REFERENCES

- [1] S. Shidian, "Lithium-ion battery pole piece laser thickness measurement technology and system development," Ph.D. dissertation, Huazhong Univ. Sci. Technol., Wuhan, China, 2018.
- [2] J. N. Sun et al., "Research progress of lithium battery electrode rolling technology," *China Metall.*, vol. 31, no. 5, pp. 12–18, 2021.
- [3] S. Dingyu and G. Lin, "Design of control system for power lithium-ion battery pole piece rolling mill based on STM32," *Instrum. Technol. Sensors*, vol. 1, no. 6, pp. 83–86 and 91, 2020.
- [4] L. Wei, W. Dingxiang, and T. Lijun, "Design of laser measurement system for thickness of lithium battery pole piece," *Electron. World*, vol. 12, no. 5, pp. 145–148, 2019.
- [5] F. Ziyang, "Multi-agent theory and its application in thickness control system," M.S. thesis, Northeastern Univ., Beijing, China, 2012.
- [6] X. Shenping and Z. Weiwei, "Power load forecasting based on improved genetic algorithm optimizing BP neural network," *Electron. Test.*, vol. 5, no. 22, pp. 42–44, 2020.
- [7] M. K. Wei, "A short-term electric load forecasting method based on self-organizing feature neural network and least squares support vector machine," *Mod. Electr.*, vol. 38, no. 1, pp. 17–23, 2021.
- [8] W. Yingli, T. Shuai, H. Xiaoxiao, and Q. Hong, "Short-term photovoltaic power generation prediction based on GA-BP neural network based on MIV analysis," *Chin. J. Solar Energy*, vol. 41, no. 8, pp. 236–242, 2020.
- [9] X. Zhu, "Short-term prediction of photovoltaic power generation based on improved NARX neural network algorithm," *J. Wuhan Univ. Sci. Ed.*, vol. 66, no. 5, pp. 505–511, 2020.
- [10] X. Jin, J. Zhang, J. Kong, T. Su, and Y. Bai, "A reversible automatic selection normalization (RASN) deep network for predicting in the smart agriculture system," *Agronomy*, vol. 12, no. 3, p. 591, Feb. 2022.
- [11] J. Kong, H. Wang, C. Yang, X. Jin, M. Zuo, and X. Zhang, "A spatial feature-enhanced attention neural network with high-order pooling representation for application in pest and disease recognition," *Agriculture*, vol. 12, no. 4, p. 500, Mar. 2022.
- [12] Y.-W. Wang, J.-G. Cao, C.-N. Song, L.-L. Wang, L. Sun, D. Xie, and Y.-L. Lu, "Research on high-precision transverse thickness difference control strategy based on data mining in 6-high tandem cold rolling mills," *Steel Res. Int.*, Jan. 2022, Art. no. 2100514.

- [13] H. Bo, Y. Zhang, L. Sihan, Z. Fei, Q. Guo, J. Zhang, and T. Yildirim, "Temperature prediction for finish entry of hot strip mill based on data-driven," in *Proc. 39th Chin. Control Conf. (CCC)*, Jul. 2020, pp. 2487–2493.
- [14] X. Wang and X. Yan, "Dynamic model of the hot strip rolling mill vibration resulting from entry thickness deviation and its dynamic characteristics," *Math. Problems Eng.*, vol. 2019, Feb. 2019, Art. no. 5868740.
- [15] D. Zhikui, L. Pengwei, Q. Chaoyue, S. Jianliang, Z. Jingyi, and L. Mingli, "Research on vibration prediction of hot-rolled high-strength steel sheet mill based on DBN algorithm," *Mining Metall. Eng.*, vol. 40, no. 4, pp. 135–141 and 144, 2020.
- [16] X. Li and Y. Dai, "Prediction model of hot rolled strip quality based on K-means clustering and neural network," in *Proc. 11th Int. Symp. Comput. Intell. Design (ISCID)*, Dec. 2018, pp. 150–153.
- [17] W. Qianfeng, "Prediction of rolling force of rolling mill based on improved support vector machine algorithm," *Forging Technol.*, vol. 44, no. 4, pp. 131–137, 2019.
- [18] Z.-H. Wang, D.-Y. Gong, G.-T. Li, and D.-H. Zhang, "Bending force prediction model in hot strip rolling based on artificial neural network optimize by genetic algorithm," *J. Northeastern Univ., Natural Sci.*, vol. 39, pp. 1717–1722, Dec. 2018.
- [19] G. Zheng, L.-H. Ge, Y.-Q. Shi, Y. Li, and Z. Yang, "Dynamic rolling force prediction of reversible cold rolling mill based on BP neural network with improved PSO," in *Proc. Chin. Automat. Congr. (CAC)*, 2018, pp. 2710–2714.
- [20] X. Wang, Y. Li, T. Chen, Q. Yan, and L. Ma, "Quantitative thickness prediction of tectonically deformed coal using extreme learning machine and principal component analysis: A case study," *Comput. Geosci.*, vol. 101, pp. 38–47, Apr. 2017.
- [21] L. Binbin, D. Xiaozhong, W. Rongjun, and Y. Shijian, "Research on rolling technology of electrodes for power lithium ion battery," *Mech. Sci. Technol.*, vol. 37, no. 4, pp. 592–598, 2018.
- [22] Y.-X. Xu, L.-C. Niu, H. Yang, Y.-C. Xiao, and Y.-J. Xiao, "Optimization of lithium battery pole piece thickness control system based on GA-BP neural network," *J. Nanoelectron. Optoelectron.*, vol. 14, no. 7, pp. 978–986, Jul. 2019.
- [23] L. Naipeng, C. Xiao, L. Yaguo, X. Pengcheng, W. Wenting, and W. Biao, "A digital-analog linkage mechanical residual life prediction method fused with multi-sensor data," *Chin. J. Mech. Eng.*, vol. 57, no. 20, pp. 29–37 and 46, 2021.
- [24] X. Wang, T. Wang, A. Ming, Q. Han, F. Chu, W. Zhang, and A. Li, "Deep spatiotemporal convolutional-neural-network-based remaining useful life estimation of bearings," *Chin. J. Mech. Eng.*, vol. 34, no. 3, pp. 128–142, 2021.
- [25] Y. Wang, Y. Zhao, and S. Addepalli, "Practical options for adopting recurrent neural network and its variants on remaining useful life prediction," *Chin. J. Mech. Eng.*, vol. 34, no. 1, pp. 45–64, Dec. 2021.
- [26] L. Shua, W. Shibo, X. Yang, and G. Shirong, "Research on dynamic prediction method of coal seam thickness based on LSTM," *Coal Sci. Technol.*, vol. 49, no. S1 pp. 150–157, 2021.
- [27] Z. Xiaoxiong, "Steel strip thickness prediction based on deep belief network," M.S. thesis, Metall. Automat. Res. Des. Inst., Beijing, China, 2020.
- [28] N. Li, N. Gebrael, Y. Lei, X. Fang, X. Cai, and T. Yan, "Remaining useful life prediction based on a multi-sensor data fusion model," *Rel. Eng. Syst. Saf.*, vol. 208, Apr. 2021, Art. no. 107249.
- [29] P. Hao, A. He, and W. Sun, "Predicting model of thickness distribution and rolling force in angular rolling process based on influence function method," *Mech. Ind.*, vol. 19, no. 3, p. 302, 2018.
- [30] W. Qi, L. Feng, T. Yi, and X. Yusheng, "Online prediction method of power grid transient frequency characteristics based on physics-data fusion model," *Automat. Electr. Power Syst.*, vol. 42, no. 19, pp. 1–9, 2018.
- [31] X. Li, Y. Fang, and L. Liu, "Decoupling predictive control of strip flatness and thickness of tandem cold rolling mills based on convolutional neural network," *IEEE Access*, vol. 8, pp. 3656–3667, 2020.
- [32] H. Li, C. Yao, J. Shao, A. He, Z. Zhou, and W. Li, "Prediction and control technology of stainless steel quarter buckle in hot rolling," *Metals*, vol. 10, no. 8, p. 1060, 2020.
- [33] Y. Conglin, "Modeling, simulation and experimental study of hydraulic control system of $\Phi 800$ battery pole mill," M.S. thesis, Yanshan Univ., Qinhuangdao, China, 2013.
- [34] Z. Xiuchao, "Development of hydraulic thickness automatic control system for heavy plate rolling mill," M.S. thesis, Harbin Inst. Technol., Harbin, China, 2015.
- [35] S. Dongfa, Z. Shenwang, and Z. Ning, "Calculation of deformation of variable section shaft," *J. Jiamusi Univ., Natural Sci. Ed.*, vol. 28, no. 6, pp. 853–857, 2010.
- [36] T. Meng, "Research on process control system and rolling force prediction of single-stand cold rolling mill," M.S. thesis, Dalian Univ. Technol., Dalian, China, 2012.
- [37] H. Dongxiao, L. Haoran, L. Jizhao, and L. Fei, "Modeling and simulation of dynamic rolling force under asymmetric lag deformation of rolling stock," *Chin. J. Mech. Eng.*, vol. 48, no. 14, pp. 38–43, 2012.
- [38] H. Hongqian, C. Furong, G. Renguo, Z. Zhanyong, and X. Zhenhuan, "Mathematical model and theoretical study on unit rolling pressure distribution in strong rheological rolling of semi-solid alloys," *J. Metals*, vol. 47, no. 3, pp. 291–297, 2011.
- [39] X. Bampoula, G. Siaterlis, N. Nikolakis, and K. Alexopoulos, "A deep learning model for predictive maintenance in cyber-physical production systems using LSTM autoencoders," *Sensors*, vol. 21, no. 3, p. 972, 2021.
- [40] L. Chen, "Research on LSTM-based state trend prediction method for mechanical equipment," M.S. thesis, China Univ. Petroleum, Beijing, China, 2019.
- [41] I. Anagiannis, N. Nikolakis, and K. Alexopoulos, "Energy-based prognosis of the remaining useful life of the coating segments in hot rolling mill," *Appl. Sci.*, vol. 10, no. 19, pp. 6827–6845, 2020.
- [42] S. Hong, Z. Zhou, E. Zio, and K. Hong, "Condition assessment for the performance degradation of bearing based on a combinatorial feature extraction method," *Digit. Signal Process.*, vol. 27, pp. 159–166, Apr. 2014.
- [43] H. Yang, J. C. K. Lam, V. O. K. Li, and D. Reiner, "A Bayesian LSTM model to evaluate the effects of air pollution control regulations in Beijing, China," *Environ. Sci. Policy*, vol. 115, pp. 26–34, Jan. 2021.
- [44] X.-B. Jin, X.-H. Yu, T.-L. Su, D.-N. Yang, Y.-T. Bai, J.-L. Kong, and L. Wang, "Distributed deep fusion predictor for a multi-sensor system based on causality entropy," *Entropy*, vol. 23, no. 2, p. 219, Feb. 2021.
- [45] L. Yang, M. Xu, Y. Guo, X. Deng, F. Gao, and Z. Guan, "Hierarchical Bayesian LSTM for head trajectory prediction on omnidirectional images," *IEEE Trans. Pattern Anal. Mach. Intell.*, early access, Oct. 1, 2022, doi: 10.1109/TPAMI.2021.3117019.
- [46] J. Chen, D. Pi, Z. Wu, X. Zhao, Y. Pan, and Q. Zhang, "Imbalanced satellite telemetry data anomaly detection model based on Bayesian LSTM," *Acta Astronaut.*, vol. 180, pp. 232–242, Mar. 2021.
- [47] C. Ma, H. Gui, and J. Liu, "Self learning-empowered thermal error control method of precision machine tools based on digital twin," *J. Intell. Manuf.*, pp. 1–2, Aug. 2021, doi: 10.1007/s10845-021-01821-z.
- [48] X.-B. Jin, W.-T. Gong, J.-L. Kong, Y.-T. Bai, and T.-L. Su, "PFVAE: A planar flow-based variational auto-encoder prediction model for time series data," *Mathematics*, vol. 10, no. 4, p. 610, Feb. 2022.
- [49] Y. A. W. Shardt, S. Mehrkanoo, K. Zhang, X. Yang, J. Suykens, S. X. Ding, and K. Peng, "Modelling the strip thickness in hot steel rolling mills using least-squares support vector machines," *Can. J. Chem. Eng.*, vol. 96, no. 1, pp. 171–178, Jan. 2018.
- [50] G. Bin, M. Lingqi, D. Yong, and M. Shengbiao, "Thickness prediction of medium thick plate mills based on GRNN neural network," *J. Central South Univ.*, vol. 42, no. 4, pp. 960–965, 2011.



YANJUN XIAO received the bachelor's degree in industrial automation and the master's degree in mechanical design and manufacturing and automation from the Hebei University of Technology, in 2000 and 2009, respectively. From 2001 to 2007, he worked with the Central Laboratory, School of Mechanical Engineering, Hebei University of Technology, where he has been a Professor with the School of Mechanical Engineering, since 2017. He is currently teaching at the School of Mechanical Engineering, Hebei University of Technology. He is also working at the Tianjin Key Laboratory of Power Transmission and Safety Technology for New Energy Vehicles, School of Mechanical Engineering, Hebei University of Technology. He is a Professional at Jiangsu Career Leader Company Ltd. His main research interests include waste heat recovery and industrial control. His awards and honors include the title of "Three Three Three" Three-Level Talent in Jiangsu Province in 2018, the 2017 and 2019 Hebei Science and Technology Invention Award, and the Honorary Title of "Hebei Science and Technology Talent" in 2017.



SHUHAN DENG graduated from the Harbin University of Science and Technology, in 2020. He is currently pursuing the master's degree with the Hebei University of Technology. He mainly engaged in intelligent industrial connectivity and embedded development.



FURONG HAN graduated from the Hebei University of Technology, in 2020. She is currently pursuing the master's degree with the Hebei University of Technology. Her research interests include intelligent control and health management.



XIAOLIANG WANG is currently pursuing the master's degree with the Hebei University of Technology. His major is electronic information. His research interests include embedded technology and intelligent control.



ZONGHUA ZHANG received the Ph.D. degree from the School of Precision Instruments and Optoelectronics Engineering, Tianjin University, in 2001. From September 2012 to March 2013, he was a Visiting Scholar at the University of Huddersfield, U.K. From August 2013 to August 2014, he was an Exchange Visitor at the Centre for Surface Precision Measurement, a National Research Institute for Advanced Manufacturing, University of Huddersfield. He is currently a Professor and a Ph.D. Supervisor with the School of Mechanical Engineering, Hebei University of Technology, a member of the Academic Committee of the University, the Chairperson of the Academic Committee of the School of Mechanical Engineering, and the Director of the Key Laboratory of Tianjin City Enterprises. He is now in charge of many national and provincial projects, such as the National Key Research and Development Program "Major Scientific Instruments and Equipment Development" Key Special Project, the National Natural Science Foundation of China, the Key Basic Research Project of Hebei Province Applied Basic Research Program, the Tianjin Science and Technology Program, and the Open Fund of the Key Laboratory of Optoelectronic Devices and Systems of the Ministry of Education.



KAI PENG received the Ph.D. degree in instrumentation engineering from Tianjin University. He is currently working as a Teacher with the Department of Measurement and Control, Hebei University of Technology, with the main research direction of geometric measurement technology, machine vision, and applications. He has published more than ten core articles, including three EI retrieved articles, more than ten authorized patents, and six copyrights. He has presided over one provincial and ministerial-level project, five enterprise projects, participated in or led three national projects and five provincial and ministerial-level projects. He has won the Second Prize of the Hebei Science and Technology Award for the Main Researcher and the Hebei Challenge Cup Excellent Guidance Teacher.

...