SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(4), 1733.
https://doi.org/10.62617/mcb1733

Article

Mechanism-data hybrid modeling method for the drafting section of a sliver machine

Xiutian Zhang, Ming Li*

School of Intelligent Manufacturing Modern Industry, Xinjiang University, Urumqi 830017, China
*** Corresponding author:** Ming Li, xj_liming@163.com

CITATION

Zhang X, Li M. Mechanism-data hybrid modeling method for the drafting section of a sliver machine. Molecular & Cellular Biomechanics. 2025; 22(4): 1733. https://doi.org/10.62617/mcb1733

ARTICLE INFO

Received: 1 March 2025
Accepted: 13 March 2025
Available online: 19 March 2025

Abstract: The rapid development of information technology has enhanced the accuracy and depth of intelligent algorithms. However, significant challenges remain in expanding the application scope of intelligent systems and achieving synergistic coupling between mechanical and digital technologies. Inspired by multiscale dynamic feedback mechanisms in cellular biomechanics, such as cytoskeletal remodeling and mechanotransduction in extracellular matrix fiber networks, this study proposes a bio-inspired hybrid modeling framework that analogizes the integration of mechanical and digital systems to the coordinated molecule-cell-tissue multiscale mechanical responses in biological systems. A high-precision mechanistic model is constructed using Recurdyn multi-body dynamics simulation software to capture the physical characteristics of the drafting section through a strategy analogous to multiscale mechanical modeling of ECM (extracellular matrix) fibrous networks. Simultaneously, a GRU (Gated Recurrent Unit) neural network-based data-driven model is developed to emulate the adaptability of biological neural systems, particularly the feedback regulation of neuronal networks under dynamic mechanical stimuli. By calculating residuals between the mechanistic model, data-driven model, and experimental measurements, a dual-channel SKNet (Selective Kernel Networks) architecture is introduced to mimic the multiscale signal extraction properties of mechanosensitive ion channels in cellular biomechanics. Convolutional kernels of different scales extract residual features, and an LSTM (Long Short-Term Memory) residual model is constructed for compensation. Experimental validation demonstrates that the hybrid model significantly improves prediction accuracy and robustness, with its residual compensation mechanism functionally resembling the dynamic repair processes of cells under mechanical stress. This study provides an efficient solution for drafting section modeling and offers methodological insights for interdisciplinary applications in biomaterial fabrication and tissue engineering bio-inspired design.

Keywords: drawing frame; mechanistic model; biomechanics; data-driven; hybrid modeling; neural network; residual compensation; multiscale modeling

1. Introduction

With the deepening of global economic integration and technological revolution, the textile industry is facing unprecedented changes. The traditional textile production model is gradually transforming to become intelligent and digital in order to adapt to the rapidly changing market demand and enhance the flexibility of the industrial chain [1]. In this context, multiscale mechanical modeling serves as a pivotal methodology in biomedical engineering for deciphering dynamic coupling mechanisms spanning molecular, cellular, and tissue hierarchies. A representative example lies in extracellular matrix (ECM) fibrous networks, where mechanical properties govern cell migration and differentiation—a system demanding integration of molecular

dynamics simulations with macroscale experimental data. Similar challenges arise in textile machinery drafting component modeling: Both domains require balancing mechanistic interpretability against data-driven adaptability [2]. In the spinning system, after the fiber material has been carded, it needs to be drafted to improve the fiber condition. This process allows hook-crimped fibers to be further straightened and parallelized, thereby improving the quality of the sliver [3]. The drawing frame, a key piece of equipment used in the textile industry, is responsible for merging, drafting, mixing, and making uniform slivers from multiple fiber strips. The drawing frame process is a complex fiber movement process, influenced by various factors such as process parameter settings, fiber properties, and fiber speed point distribution. Among these factors, the distribution of variable speed points is closely related to the quality of the slivers. The quality of the slivers is better when the distribution of the variable speed points is more concentrated and closer to the front roller jaws. Therefore, it is of great significance to study the distribution theory of the variable speed points in order to optimize production process parameters and strengthen fiber control.

Drafting is a classic topic, focusing on drafting theory, many studies conducted by Chinese and foreign scholars and technicians, and the research content mainly focuses on the drawing frame drafting process, drafting structure, tension drafting, drafting roller drafting microtraction, uneven formation of lines in the traction process, and other research aspects. In many research fields, in order to effectively improve the correlation between the operating status of institutions and related models, many scholars have carried out in-depth research and achieved a series of valuable results [4]. Most of the existing studies focus on the optimization of the configuration and process analysis of the drafting mechanism, the optimal design of the draft part of the pressure bar, and the analysis of the influence of the drafting force on fiber fracture during the drafting process [5]. However, compared with the local optimization of the drafting mechanism, the discussion on the transmission mechanism of the front and rear gearbox and the dynamic behavior and efficiency analysis of the entire transmission system is still insufficient, and the digital twin technology can realize the accurate modeling of the key components of the transmission mechanism such as the front and rear gearboxes, and dynamically update the model through real-time data feedback, so as to simulate the actual operation state in the virtual environment, and the prediction and optimization of system performance with the help of real-time and global monitoring and analysis technology become more refined and efficient. Compared to traditional modeling methods, it offers significant advantages in improving system reliability, reducing failure rates, and optimizing energy efficiency. Lee et al. [6] took the feed system of CNC(Computer Numerical Control) machine tools as the research object, constructed a digital twin model driven by a mechanism and data mix, and proposed an adaptive model update method, which provides new ideas for research in this field. Du et al. [7] successfully established a hybrid model of constant hydraulic balance analysis of heat networks by using the method of data-driven correction mechanism model, which opened up a new path for the research in the field of heat networks. Li et al. [8] focused on the driving force of robotics in mechanical equipment, and the research results showed that this kind of microrobot has the ability of multi-mode motion adapted to the environment and can realize the function of transmission to the direction.

Focusing on the field of battery electrode manufacturing, Vijay et al. [9] proposed a hybrid modeling method, which organically combines physics-based simulation with deep learning, and successfully applied it to battery electrode manufacturing simulation. With its powerful data processing and pattern recognition capabilities, the deep learning part excavates complex nonlinear relationships from a large number of actual production data and captures the subtle features and coupling effects that are difficult to describe with traditional physical models, thus making up for the shortcomings of purely physics-based simulation. The application of hybrid models greatly improves the correlation between the model and the actual operating state in the battery electrode manufacturing process and provides technical support for the improvement of the battery electrode manufacturing process, the improvement of production efficiency, and the improvement of product quality. In addition, some researchers have conducted in-depth discussions on the optimization of the drawing frame's control system based on the mechanism of drawing frames, the application of data-driven and hybrid modeling in the traction and tensile processes of drawing frames, and the importance of neural networks in the design of predicted value compensation to mechanism models [10]. This research not only has a broad application prospect of multi-mode motion and directional transfer functions in the field of industrial manufacturing but also further promotes the in-depth integration and development of intelligent and automation technologies in the field of textile machinery, especially in the field of drawing frames. Roland et al. [11] used the Gated Recurrent Unit (GRU) neural network to update and predict deformation and found that deep learning far surpasses the Bayesian optimization update method in terms of effectiveness and efficiency. This greatly improves the timeliness and accuracy of deformation prediction and provides strong technical support for the safety assurance and risk control of perpetual motion machines. Leite et al. [12] combined the knowledge of the physical mechanism of the drafting process and constructed a model based on the mechanism and data fusion. They described the basic principles of fiber motion and friction in the drafting process through the mechanism model, built the basic framework of the model, and used the data processing ability of deep learning to mine the complex nonlinear relationships in the drafting process and the factors that are difficult to be directly described by the mechanism. This approach not only realizes the accurate prediction of key parameters such as drafting force but also provides strong support for the improvement of product quality and production efficiency and promotes the development and progress of drawing frame drafting modeling technology.

On the basis of previous research, this study studies the modeling method of the drafting part of the drawing frame based on the mechanism and data fusion method and uses the residuals of the mechanism model and the data model to train the residual model to compensate the mechanism model, which does not need to determine the parameters with greater influence than the tandem method, reduces the dependence on the prediction accuracy of the mechanism model compared with the parallel method, ensures the interpretability of the model, and improves the prediction accuracy of the mode [13,14]. At the same time, the hybrid model established by this method embodies one of the characteristics of the digital twin in the compensation mechanism model [15,16], but due to the lack of a real-time data interaction mechanism, it cannot

be regarded as a complete digital twin model. On the basis of the parallel algorithm, a multi-model residual learning method is proposed, which is used to establish a hybrid model of a mechanism model and a data-driven model, and the research algorithm focuses on the modeling process of the parallel hybrid model and improves on the basis of the existing method and elaborates the proposed multi-model hybrid modeling method in order to provide reference experience and reference for the drawing frame drawing part modeling method based on mechanism and data fusion.

2. Model principle and modeling process

This section elaborates on the principles and modeling processes for constructing both the mechanistic model and the data-driven model. The first part constructs the mechanistic model of the drawing section of the FA320A drawing machine, while the second part builds a data-driven model based on a GRU neural network.

Regarding the FA320A drawing machine, the focus of this study, its drive system design, stands as a core component of the equipment. It plays a crucial role in the stable operation and high-quality production of the machine. The drive system design is depicted in **Figure 1** (FA320A Drive System Diagram). Comprising several key components, such as 1) drive pulley, 2) transmission sprocket, 3) chain, 4) coaxial sprocket, 5) front gear box, 6) rear gear box, 7) main motor, etc., the drive system is integral. The front and rear gearboxes are of vital importance within it, coordinating and matching parameters like the tension and speed of different yarns. This ensures that multiple yarns can be arranged neatly under suitable conditions to ultimately form a uniform sliver.

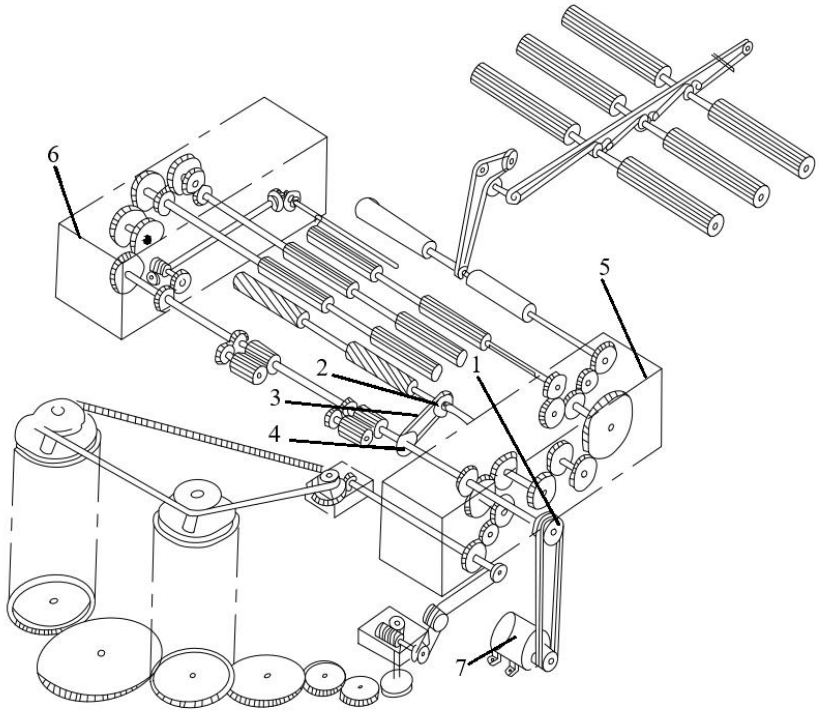

Figure 1. FA320A drive system diagram.

1: Drive pulley, 2: transmission sprocket, 3: chain, 4: coaxial sprocket, 5: front gearbox, 6: rear gearbox, 7: main motor.

As for the machine's configuration, the FA320A drawing machine adopts a 4-up4-down curve drafting pressure configuration with guide rollers. The upper rollers and the pressure springs are not directly connected to the drive section but are instead linked to the frame via bearings and supports. The model structure is presented in **Figure 2**. This design enables the upper rollers to flexibly adapt to the motion state of the yarn during operation. The pressure spring design also provides stable pressure, ensuring that the yarn tension remains stable during the drafting process.

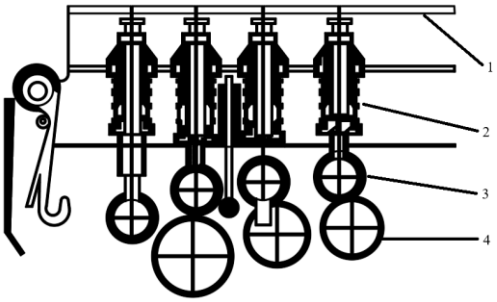

Figure 2. Tensioning section pressure apparatus.
1: Frame assembly, 2: compression spring, 3: upper roller, 4: lower roller.

2.1. Principle of the FA320A drawing machine transmission system

The FA320A drawing machine employs a pressure bar stretching format with four upper and four lower guide rollers. The transmission system, a core component of the machine, operates through gearboxes at both the front and rear ends, primarily coordinating the tension and speed of different yarns to align them side by side, ultimately forming a uniform sliver. The upper roller and the pressure springs are not directly connected to the transmission system but are instead mounted on the machine frame through bearings and support structures. The model structure is illustrated in **Figure 2**.

2.2. Constructing the mechanistic model using Recurdyn simulation software

A mechanistic model is a mathematical representation constructed based on a deep understanding of the internal working mechanisms of a system, with physical, chemical, and biological principles serving as the foundation. Its core feature lies in the ability to explain and depict the interactions between various components of the system at a fundamental level, as well as how these interactions collectively shape and determine the overall behavior of the system [17]. In many research and application scenarios in engineering, constructing an accurate and effective mechanistic model is of indispensable importance for analyzing the operating principles of the system, optimizing system performance, and conducting reliable predictive analyses [18,19].

When faced with the task of modeling and analyzing complex mechanical systems, selecting the appropriate tool is crucial. Recurdyn is a software specifically designed for multibody dynamics analysis, and it demonstrates outstanding advantages in handling complex mechanical transmission systems, ensuring high analytical precision while maintaining computational efficiency [20,21]. Particularly in large-scale multibody system simulations, Recurdyn's relatively fast computational

speed and powerful ability to handle a large number of components and complex constraints allow it to significantly reduce simulation time and greatly improve work efficiency [22].

In this study, the key components of the drawing section, such as shafts, gears, and rollers, were first modeled using SolidWorks and then imported into the multibody dynamics simulation software Recurdyn in the Parasolid file format. As the upper rollers and pressure springs are connected to the frame via bearings and supports, not directly linked to the transmission section, for simplicity and to reduce computational burden, this part was modeled as a frictional force and represented as a load in the simulation. The friction force equation is used as follows:

$$f = \mu \times F_N \tag{1}$$

In the section addressing the unilateral pressure F_N, the pressure exerted by the compression springs on the upper roller is sequentially 294 N, 294 N, 392 N, and 392 N. The coefficient of friction μ is taken as 0.1 in this context. The values for pressure correspond to the unilateral pressure, and when calculating frictional force, this value is multiplied by 2. Consequently, the loads on the roller are 58.8 N, 58.8 N, 78.4 N, and 78.4 N, respectively.

Parameter settings for the imported model in the Recurdyn software, including constraints, drives, and loads, are outlined in **Table 1**. **Figure 2** displays the model's state after adjusting constraints and contact parameters. A dynamic simulation is conducted on the adjusted model to establish the mechanistic model of the stretching section.

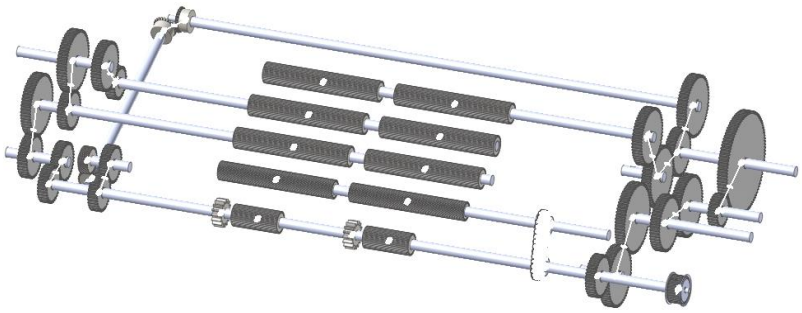

Figure 3. Drafting section model.

Figure 3 illustrates the primary transmission components within the transmission system: The right side depicts the front gearbox transmission section, the left side shows the rear gearbox transmission, the middle section contains four lower rollers, and the lower right corner features a pulley connected to the main motor, which is a 5.5 kW, 4-pole variable frequency motor.

Table 1 shows the constraints and contact types involved in the modeling process. In the modeling of the FA320A drawing machine's drawing section, based on the motion characteristics and interaction relationships of each component, appropriate fixed and rotational pairs are set for the gears, rollers, shafts, and chains. Surface contact is set between the gears, and solid contact is established between the chain and sprocket. A constant rotational speed of 1450 n/min is applied to the drive wheel to provide power input to the entire transmission system, ensuring that the

simulation accurately reflects the actual operating conditions.

Table 1. Constraints and contact statistics between components.

Type	Location	Add mode	Quantity
Joint	Gear, roller	Translational	40
	Shaft	Revolute	5
	Chain	Revolute	32
Contact	Between gears	Surface to surface	14
	Chain and sprocket	Solid to solid	64

In multibody dynamics simulations, the mechanistic model of the drafting machine is established through dynamic analysis [23]. This model reconstructs geometric features and component spatial relationships with high fidelity [24,25], akin to coarse-grained representations of cytoskeletal networks in cellular biomechanics that preserve macroscopic mechanical behavior while abstracting molecular-scale details. However, structural complexity and parameter sensitivity necessitate simplifications during modeling—a compromise reminiscent of tissue-level biomechanical studies where heterogeneous cell-ECM interactions are homogenized to tractable constitutive laws.

The deterministic nature of mechanistic frameworks further limits their capacity to emulate transient environmental perturbations, mirroring the challenge of predicting dynamic ECM remodeling driven by fluctuating cellular traction forces. Such gaps between idealized simulations and operational variability echo the fidelity-adaptability trade-off inherent in biological systems.

To bridge this gap, a data-driven model is introduced, inspired by biological systems that decode environmental noise through adaptive signaling. Just as mechanosensitive ion channels convert mechanical stimuli into electrochemical signals for real-time cellular feedback, the hybrid framework leverages neural networks to dynamically compensate for unmodeled perturbations, synergizing interpretable mechanics with data-driven adaptability [26,27].

2.3. LSTM and GRU neural networks

As illustrated in **Figure 4A**, the LSTM neural network's basic structure replaces the hidden state's activation function in traditional RNNs (Recurrent Neural Network) with memory units. It controls the flow of information through input gates, forget gates, and output gates, thus preventing issues of gradient vanishing and explosion. The specific implementation process of LSTM is as follows:

(1) Calculation of the forget gate f_t:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \tag{2}$$

The forget gate determines which information from the previous cell state C_{t-1} should be omitted. In this formulation, h_{t-1} and x_t serve as inputs, with W_f and b_f representing the weight matrix and bias term, respectively. The sigmoid function σ is employed to facilitate this determination.

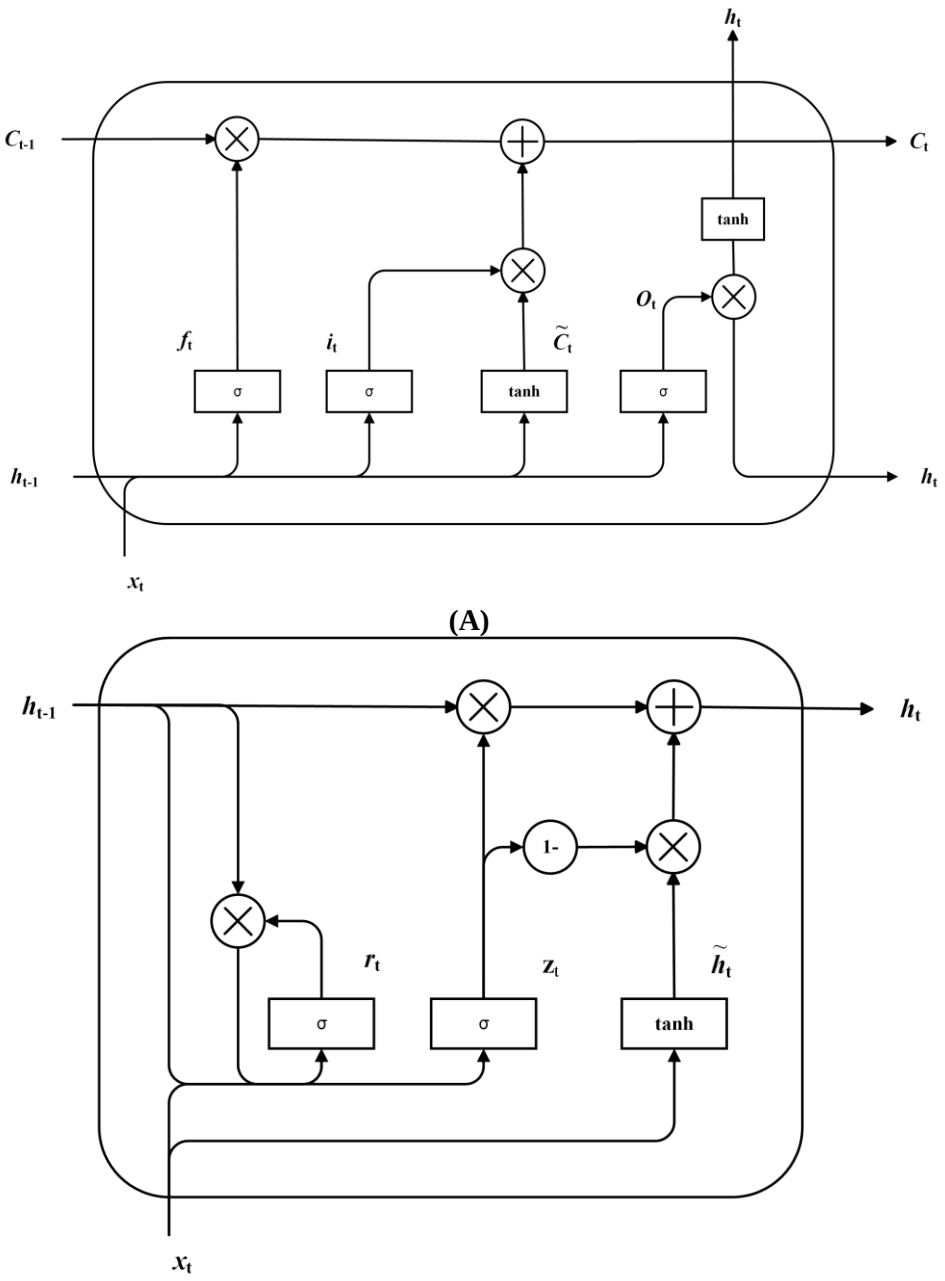

Figure 4. The basic structure of **(A)** LSTM; **(B)** GRU neural networks.

(2) Calculation of the input gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{3}$$

$$\widetilde{C_t} = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{4}$$

The input gate decides which information from the current input x_t and the previous hidden state h_{t-1} will be incorporated into the cell state. Here, W_i and b_i denote the weight matrix and bias term, respectively, while $\tanh$ serves as the activation function.

(3) Update of the cell state:

$$C_t = f_t \times C_{t-1} + i_t \times \widetilde{C_t} \tag{5}$$

The update of cell state C_t is facilitated through a conjunction of the forget gate and the input gate. The forget gate dictates the extent of old information from the cell state to be discarded, while the input gate determines the amount of current candidate information to be incorporated. In this context, C_{t-1} represents the cell state from the previous moment.

(4) Output gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{6}$$

$$h_t = o_t \times \tanh(C_t) \tag{7}$$

The output gate determines which portions of the current cell state C_t will become part of the final output as the hidden state h_t. Here, o_t denotes the value of the output gate, with W_o and b_o representing the weight matrix and bias term, respectively.

As shown in **Figure 4B**, the GRU is a simplified version of the LSTM model, featuring two main gating units: The update gate and the reset gate. Unlike LSTM, GRU does not separate the storage unit from the hidden state but controls the flow of information within the unit through these two gates. Specifically, the update gate determines the extent to which the previous time step's hidden state influences the current hidden state, and it also controls how much information is updated to the current state. A higher value of the update gate indicates that more information will be carried over to the next state unit, thus enhancing memory updating. On the other hand, the reset gate helps control the degree of forgetting, determining how the current input influences the updated state [28].

Specific computational processes of the GRU neural network are as follows:

(1) Updating the update gate z_t:

$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t] + b_x) \tag{8}$$

The update gate moderates the balance between the current and previous hidden states, influencing the extent of new state updates. In this formula, σ represents the sigmoid function, while W_z and b_x denote the weight matrix and bias terms, respectively. The vector $[h_{t-1}, x_t]$ constitutes the input.

(2) Calculating the reset gate:

$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t] + b_r) \tag{9}$$

The reset gate employs the sigmoid activation function to apply a linear transformation to the input x_t and the previous hidden state h_{t-1}, resulting in a value between 0 and 1. Outputs closer to 0 induce the network to forget information from earlier states; conversely, values near 1 prompt the retention of such information. In this context, W_r and b_r serve as the weight matrix and bias terms, respectively, determining which pieces of information are to be discarded.

(3) Computation of the candidate hidden state $\widetilde{h_t}$:

$$\widetilde{h_t} = \tanh(W \cdot [r_t \times h_{t-1}, x_t] + b \tag{10}$$

The candidate hidden state $\widetilde{h_t}$ represents a potential hidden state at the current moment, determined through the reset gate's modulation of the current input and the prior hidden state. The parameters W and b are the weight matrix and bias, respectively, while the $\tan h$ function acts as the activation mechanism. The term $r_t \times h_{t-1}$ denotes the element-wise multiplication of the reset gate with the previous hidden state, indicating the extent of memory retention.

(4) Updating the final hidden state h_t:

$$h_t = (1 - z_t) \times h_{t-1} + z_t \times \widetilde{h_t} \tag{11}$$

This step is pivotal in updating the current hidden state, with the update gate z_t balancing the retention of the old state against the adoption of the new candidate state.

The data-driven model relies on massive datasets collected via sensors, mobile internet, or relevant software tools. These models undergo training and fitting processes supported by extensive data to become automated decision-making systems. This paper develops a data-driven model for the drafting section of a ring-spinning frame based on a GRU recurrent neural network. The model includes three hidden layers with 128 neurons each, a dropout rate of 0.2, 500 training epochs, a batch size of 10 per iteration, and a learning rate of 0.001. It utilizes the Adam optimizer for parameter updates and incorporates the speeds of four rollers in the ring-spinning frame as input-output for the GRU network.

3. Improvement of hybrid model methods

In the modeling and analysis of complex industrial systems today, a single model often fails to fully and accurately meet practical requirements. To overcome the limitations of traditional single models and improve the performance and applicability of models, research and improvement of hybrid model methods become particularly important [28,29]. This section will provide a detailed explanation of the construction process of parallel hybrid models and, based on this, propose an innovative multi-model residual learning method. This will enable the establishment of more efficient and accurate hybrid models, providing new breakthroughs for the modeling of the drawing section of the spinning drawing machine.

3.1. Construction of parallel hybrid model

Previous work established individual mechanistic models and data-driven models for the drafting section of the spinning frame, although the single mechanistic model exhibits strong interpretability. However, its capacity for complex system modeling is limited, necessitating extensive domain knowledge and assumptions. Moreover, environmental factors in the production process are complex and cannot be represented in mechanistic models. Data-driven models, which often rely on the quality and quantity of data, may produce overfitting or bias and are typically considered "black box" models, making it challenging to understand their internal workings. To enhance model prediction accuracy and interpretability, this study integrates the strengths of mechanistic and data-driven models to establish a mechanistic-data hybrid model.

The parallel method establishes the hybrid model, which uses residual data from

the mechanistic model to train a neural network. The predictions from the neural network compensate for the mechanistic model, thus establishing a mechanistic-data hybrid model. The modeling process is illustrated in **Figure 5**.

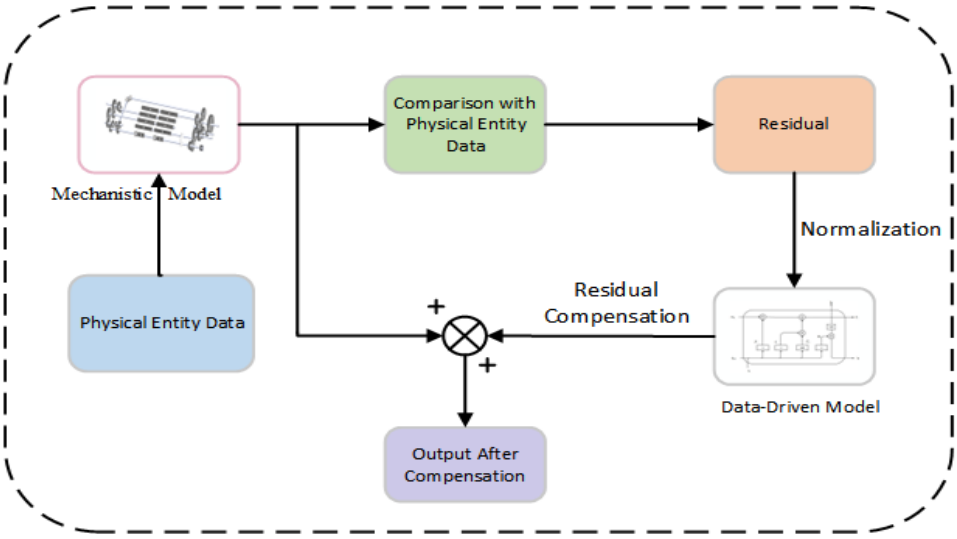

Figure 5. Process for establishing a hybrid model via parallel methodology.

Figure 5 shows the process flow for establishing the hybrid model using the parallel method, where the data-driven model is trained using residual data from the mechanistic model, and the predicted residuals are used to compensate and correct the mechanistic model. This method relies to some extent on the accuracy of the mechanistic model. Considering the numerous components and involved parameters in the drafting section of the spinning frame, simplifications are made during the construction of the mechanistic model, which may not ensure model accuracy. Therefore, on the basis of parallel hybrid modeling, a multi-model residual learning method is proposed to establish the hybrid model.

3.2. Construction of hybrid model using multi-model residual learning method

The multi-model residual learning method improves on the parallel method by reducing the dependency on the predictive accuracy of the mechanistic model. The hybrid modeling process using this method is depicted in **Figure 6**.

Figure 6 illustrates the improvements in the multi-model residual learning method, including constructing an additional data-driven model based on a GRU neural network. This model computes the residual data of the predicted values, which are then processed using SKNet to extract features from the residual data of both the mechanistic and data-driven models. These features are used to train a residual model based on the LSTM neural network, which uses the predicted residual values to compensate and correct the prediction data of the mechanistic model, thereby enhancing the overall prediction accuracy. The specific implementation process is as follows:

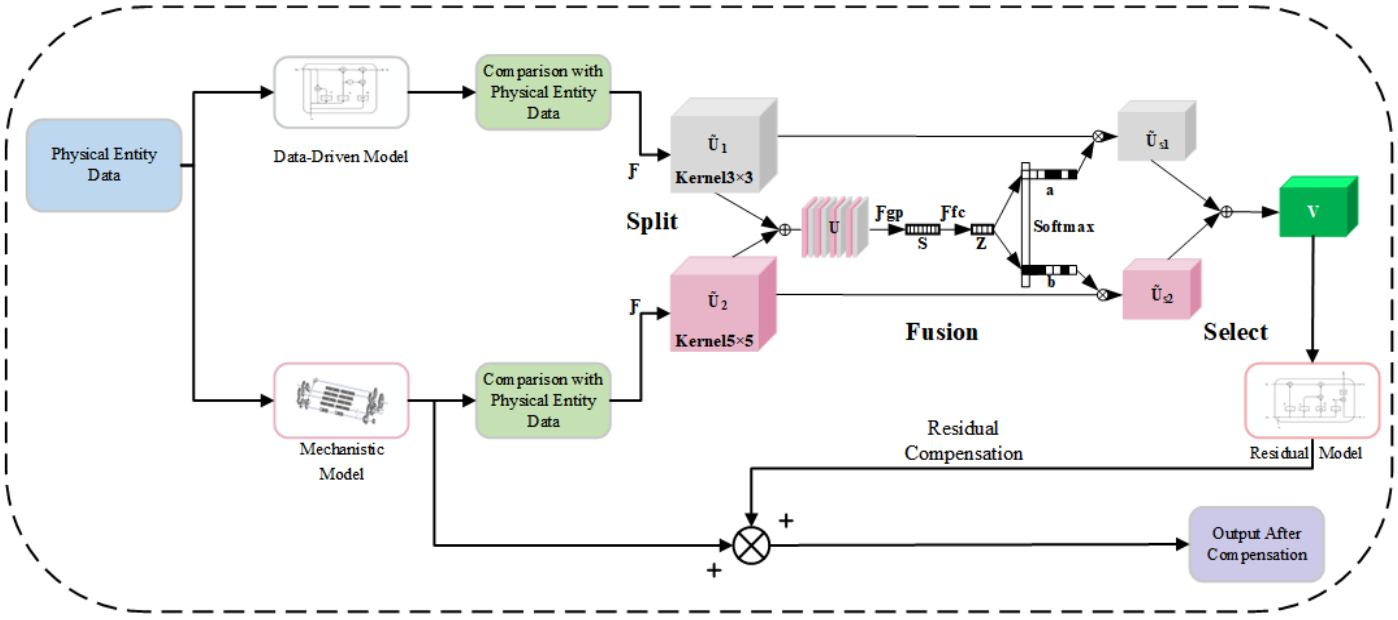

Figure 6. Process of establishing hybrid models using the multi-model residual learning method.

(1) Preprocessing of physical entity data.

Handling of missing data: Sensor malfunctions, such as equipment aging and network fluctuations, often result in data omissions. To ensure the precision of model predictions, missing data are rectified using linear interpolation.

$$x_{1(t+i)} = x_{1(t)} + \frac{i \cdot (x_{1(t+j)} - x_{1(t)})}{j}, 0 < i < j \tag{12}$$

In the equation, $x_{1(t+j)}$ and $x_{1(t)}$ denote the known data values at times $t+j$ and t respectively, while $x_{1(t+i)}$ represents the missing value at time $t+i$. Should the interval of missing values be substantial (indicated by a more significant j), interpolation using adjacent values is employed for data imputation.

Outlier Management: Under stable operating conditions, dramatic fluctuations in data are generally not expected in the spinning and doubling machinery. When a data set collected under unchanged conditions exhibits a variation exceeding ±10% at any given moment, this is considered an anomaly. Such data requires smoothing to mitigate this issue. The mean smoothing method is employed to address these outliers.

Should conditions $|x_{1(t)} - x_{1(t-1)}| > \epsilon_1$ or $|x_{1(t)} - x_{1(t+1)}| > \epsilon_2$ be met, then $x_{1(t)} = \frac{x_{1(t+1)} + x_{1(t-1)}}{2}$, ϵ_1 and ϵ_2 are designated as the respective error thresholds for adjacent data points.

(2) Construction of mechanistic and data-driven models.

This research uses Recurdyn software to establish the mechanistic model and a GRU neural network to establish the data-driven model.

(3) Residual sequences from both models are obtained using the following formula.

$$e_i = y_i - \hat{y}_i \tag{13}$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value by the model, and e_i is the residual.

(4) SKNet processes the two types of residual data for feature extraction.

SKNet, comprising the Split, Fuse, and Select modules, utilizes two differently sized perception fields (convolutional kernels), allowing incoming data to choose the most appropriate perception field during feature extraction. SKNet implements an attention mechanism for convolutional kernels, enabling the network to adaptively select the proper kernel based on different stimuli, thus enhancing feature extraction effectiveness. Furthermore, by adjusting the group convolution parameter to 3, SKNet not only reduces the number of parameters to one-third of a standard convolution but also enhances the training efficiency of the network model, achieving a lightweight design. Traditional network models typically use a single convolutional kernel from the same layer, neglecting data scale differences, which can lead to suboptimal results. Therefore, this study incorporates SKNet in the processing of data residuals, as shown in **Figure 6**, where two different-sized convolutional kernels impact differently scaled targets through adaptive adjustment of parameter weights, achieving feature extraction and adaptive output for differently scaled targets, further optimizing the extraction of residual data features.

(5) Normalization of the integrated residual data.

$$\overline{x_4} = \frac{x_3 - x_{3\min}}{x_{3\max} - x_{3\min}} \tag{14}$$

In this context, $\overline{x_4}$ represents the data normalized through the process of normalization, where x_3 denotes the original data, and $x_{3\min}$ and $x_{3\max}$ are respectively the minimum and maximum values within the dataset.

(6) A residual model was trained utilizing processed residual data.

This model is based on a Long Short-Term Memory (LSTM) neural network comprising three hidden layers, each with 256 neurons. The configuration includes a dropout rate of 0.2, with the training process spanning 500 epochs and a batch size of 10 for each iteration. The learning rate was set at 0.001, and parameter updates were facilitated using the Adam optimizer. The integrated residual sequence data were used to train the LSTM neural network, culminating in the development of the residual model.

(7) Compensation mechanism for predicted values from the residual model.

The compensation mechanism for the predicted values of the residual model was implemented using Equation (15) to enhance the accuracy of the compensation mechanism model.

$$y' = y + e \tag{15}$$

In the formula, y' represents the output from the hybrid model established by the multi-model residual learning method, y denotes the output predicted by the mechanistic model, and e is the residual value predicted by the residual model.

Building upon the parallel method framework, the multi-model residual learning method additionally incorporates a data-driven model based on GRU (Gated Recurrent Unit). This approach involves arranging residual data in columns, thereby increasing the dimensionality of residual data. Subsequently, the residual data are used to train an LSTM (Long Short-Term Memory) to develop a residual model. Finally, this model compensates for the predictive output of the mechanistic model, thereby reducing

dependency on the accuracy of the mechanistic model and establishing a robust hybrid model.

4. Experimental results and analysis

The data used in the experiment was sourced from the FA320A drafting machine in the spinning preparation workshop of Xinjiang Huashu Textile Technology Co., Ltd. At the time of data collection, the drafting machine was operating under the following conditions: The draft ratio was 6.57, the output speed was 256.61 m per second, the roller spacing was 2.52 mm, the main motor speed was 1450 revolutions per minute, and the feed rate was 150.62 m per second.

4.1. Grid search hyperparameter tuning

The hyperparameters of neural networks directly affect the model's learning efficiency, expressiveness, and generalization ability. Proper hyperparameter configuration can improve the model's performance and accuracy. This study used grid search to optimize key hyperparameters: learning rate, hidden layer size, and number of layers. The parameter space for the learning rate was set as [0.001, 0.005, 0.01]. The learning rate controls the step size for each parameter update; an appropriate learning rate helps accelerate convergence and avoids oscillations during the model training process. The parameter space for hidden size was set as [64, 128, 256]. The hidden layer size determines the network's capacity; larger hidden layers can capture more features but may also lead to overfitting. The parameter space for the number of layers was set as [2,3]. The number of layers determines the depth of the network; increasing the number of layers can improve the model's expressiveness, but it may also lead to issues such as vanishing gradients or overfitting. The loss value was chosen as the evaluation metric to evaluate the impact of different hyperparameter combinations on model performance. A grid search was then performed, traversing the hyperparameter space to find the optimal combination. In the experiment, 70% of the data was used as the training set and 30% as the test set. Finally, the optimal parameter combination was used to train the model. **Figure 7** shows the loss values of the LSTM and GRU networks with different parameter combinations.

Figure 7a,b display the loss values for the LSTM neural network, while **Figure 7c,d** show the loss values for the GRU neural network. In these figures, the model performance under different hyperparameter combinations is observed. In **Figure 7b**, when the learning rate, hidden layer size, and number of layers are set to 0.001, 256, and 3, respectively, the minimum loss value is 0.0120, indicating that the LSTM model performs best under this configuration. In **Figure 7d**, when the learning rate, hidden layer size, and number of layers are set to 0.001, 128, and 3, respectively, the minimum loss value is 0.0135, indicating that this configuration yields the best performance for the GRU model. Therefore, after optimization, the optimal hyperparameter values for the LSTM and GRU neural networks are 0.001, 256, 3, and 0.001, 128, and 3, respectively. These parameter combinations provide the best learning performance for both models.

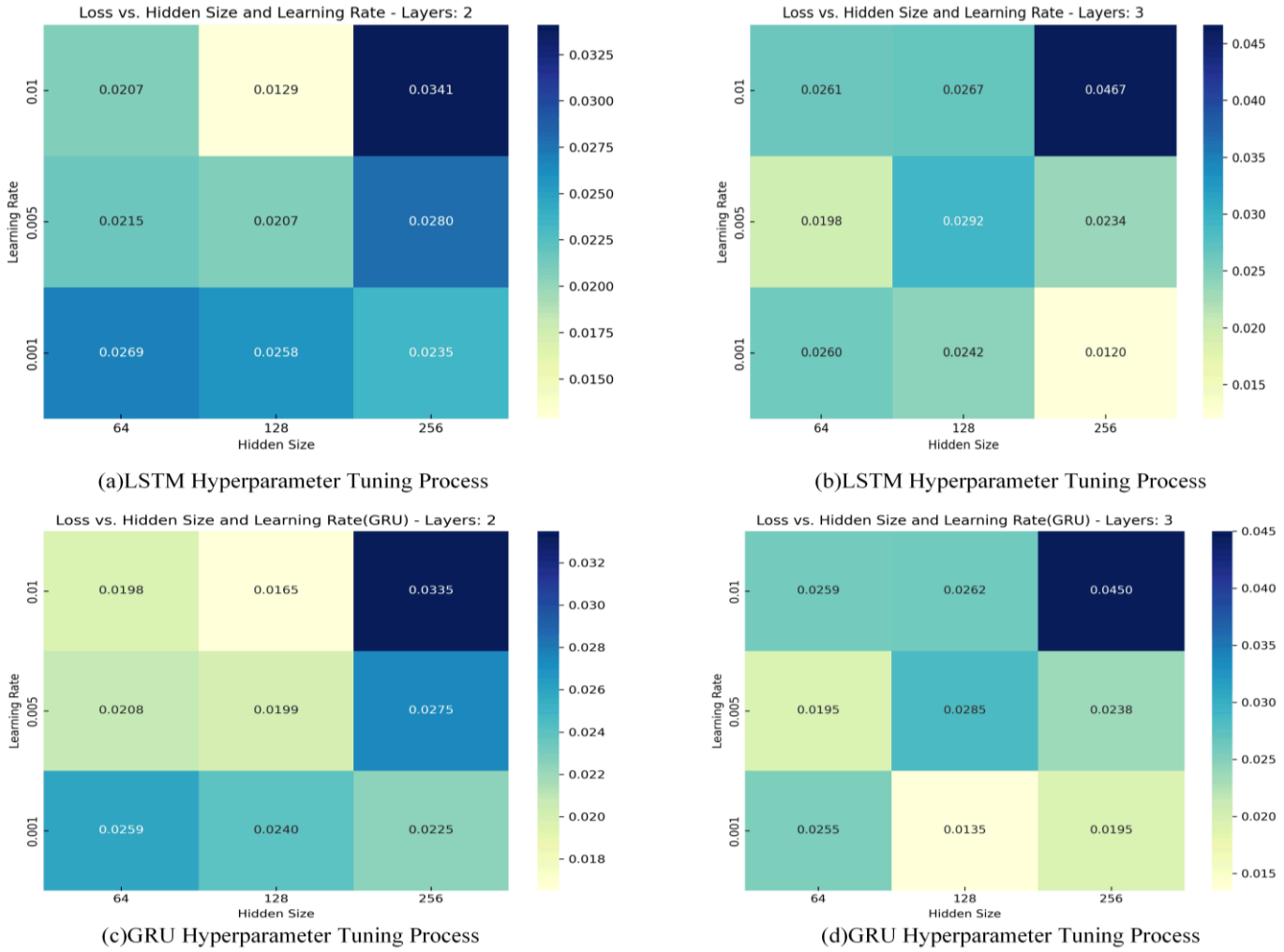

Figure 7. Hyperparameter selection process for LSTM and GRU.

4.2. Parallel comparative experiment

The measured roller speed data from the FA320A spinning drafting machine was compared with the predicted roller speed values from different models to evaluate the accuracy of the predictions made by each model.

(1) Performance of the Single Model

The single data-driven model demonstrates high accuracy in predicting the rotational speed of the rollers. Its predictions provide timely feedback during actual production, helping to monitor the dynamic behavior of the equipment in real-time [30]. However, this model has a notable limitation: It is a "black box" model with poor interpretability. When prediction deviations or equipment anomalies occur, it is difficult to quickly and accurately pinpoint the root cause of the issue, limiting its widespread application and in-depth analysis.

In contrast, the single mechanistic model has strong interpretability, helping researchers gain a deeper understanding of the model's working principles and internal mechanisms [31]. As seen in **Figure 8**, the loads on the three faster rollers exhibit significant changes, and their rotational speeds fluctuate greatly. In **Figure 8**, it is apparent that the rear roller, with a lower rotational speed, experiences relatively stable external loads, and its rotational speed does not exhibit significant fluctuations.

However, constructing an accurate mechanistic model for the many components in the drawing system presents significant challenges. To reduce modeling difficulty during simulation, some simplifications had to be made (such as simplifying the upper roller section in Chapter 2), which sacrifices a degree of model accuracy and leads to suboptimal overall prediction performance.

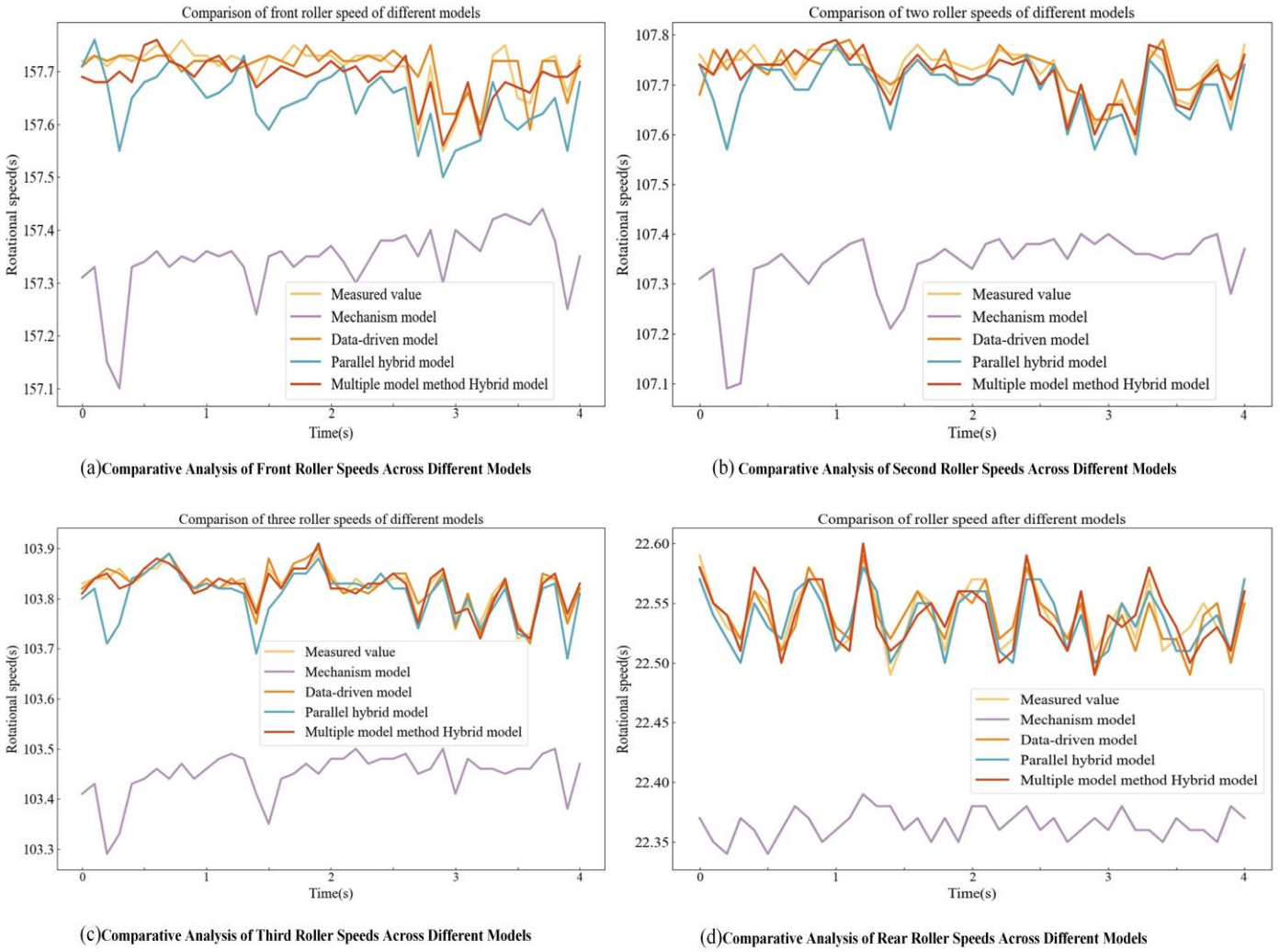

Figure 8. Group-comparison charts of speeds of each roller under different models.

(2) Performance of the parallel hybrid model

The parallel hybrid model combines the advantages of both the mechanistic model and the data-driven model. While ensuring interpretability, it improves the model's predictive accuracy. The parallel method only uses the residuals of the mechanistic model's predictions, which somewhat depends on the accuracy of the mechanistic model [32,33].

As seen in **Figure 8**, when the mechanistic model exhibits large fluctuations, the residuals of the mechanistic model at that moment are also large, reducing the effectiveness of residual compensation, and the parallel hybrid model shows larger errors at that time. In **Figure 8**, when the rear roller's rotational speed is relatively stable, the parallel method's predictions maintain high accuracy.

(3) Performance of the multi-model residual learning hybrid model

The multi-model residual learning method proposed in this study is an enhancement of the parallel hybrid model. This method uses SKNet to extract features from the residual data of both the data-driven and mechanistic models, effectively reducing the dependence on the accuracy of the mechanistic model [34]. As shown in **Figure 8**, the prediction accuracy of the rotational speeds for all four rollers significantly improved and was not affected by fluctuations in the mechanistic model's speed at certain moments. This indicates that the multi-model residual learning hybrid model successfully enhances predictive accuracy while maintaining interpretability, demonstrating superior performance and stability.

4.3. Comparison of model prediction errors

The table below presents the maximum and average relative errors in predicting the speeds of four rollers for four different models, as illustrated in **Table 2**. The speed of the rear roller, subject to a relatively stable external load, did not exhibit significant fluctuations in the mechanistic model predictions. The errors between the hybrid model developed using the parallel method and the hybrid model established through multi-model residual learning are minor; hence, the error analysis does not consider the rear roller.

Table 2. Prediction errors across different models.

Modeling method	Maximum relative error				Mean relative error			
	Front roller	Two roller	Three roller	Rear roller	Front roller	Two roller	Three roller	Rear roller
Mechanism model	0.390%	0.610%	0.530%	0.970%	0.230%	0.360%	0.360%	0.780%
Data-driven model	0.076%	0.074%	0.039%	0.170%	0.013%	0.025%	0.011%	0.047%
Parallel hybrid model	0.110%	0.160%	0.120%	0.089%	0.039%	0.033%	0.024%	0.050%
Multi-model method mixed model	0.051%	0.056%	0.029%	0.088%	0.017%	0.016%	0.011%	0.049%

The mechanistic model displayed a maximum relative error ranging from 0.39% to 0.61%, with predictions showing considerable fluctuations; the average relative error varied from 0.23% to 0.36%. Simplifications made during the modeling process led to suboptimal overall predictive outcomes.

The data-driven model had a maximum relative error between 0.039% and 0.076% and an average relative error ranging from 0.011% to 0.025%. Overall, the errors were minor, and the predictions remained stable, making this model suitable for scenarios where interpretability is not a concern, as it effectively fulfills the prediction requirements.

The parallel hybrid model exhibited a maximum relative error from 0.11% to 0.16%, affected by significant fluctuations in the mechanistic model's predictions, which compromised the stability of the outcomes; the average relative error was between 0.024% and 0.039%. The model compensates through the residuals, maintaining a relatively low overall predictive error.

The hybrid model developed using the multi-model residual learning method showed a maximum relative error from 0.029% to 0.056%, smaller than that of the parallel hybrid model and less affected by fluctuations in the mechanistic model's

predictions, resulting in more stable predictions; the average relative error ranged from 0.011% to 0.017%. This model benefits from handling residuals from both single-model types, achieving a relatively lower overall prediction error.

For the parallel and multi-model residual learning hybrid models, the latter mitigates the impact of significant fluctuations from the mechanistic model's predictions and completes the prediction tasks for all four roller speeds.

4.4. Ablation experiments

In order to gain a deeper understanding of each component's contribution to the model's overall performance, a series of ablation experiments were designed in this study. Ablation experiments evaluate the impact of each part on the overall performance by progressively removing or modifying key components of the model. This approach allows for identifying which modules or features play a decisive role in the model and which contribute little or have no impact on performance. This is crucial for further optimizing the model architecture, improving performance, and reducing computational resource consumption.

In the experiments, different settings for removing specific modules were defined: MD-S-L, MD-L, and MD-S. Specifically:

(1) MD-S-L refers to the hybrid model established by the multi-model residual learning method.

(2) MD-L refers to the model where SKNet is removed, and the residual data of the mechanistic model and the data-driven model are processed in parallel. These parallel data are then used for LSTM training, and the LSTM-predicted residual data are used to compensate for the mechanistic model, forming a hybrid model.

(3) MD-S refers to the model where LSTM is removed, and the residual data of the mechanistic model and data-driven model are directly processed by SKNet for feature recognition, compensating for the mechanistic model to form the hybrid model.

Table 3 shows the prediction accuracy of the multi-model residual learning method with different modules removed. By comparing the prediction accuracy of different configurations, the following observations can be made:

(1) Removal of LSTM: After removing LSTM, the model's ability to capture temporal dependencies significantly declines. Since there are temporal dependencies in the roller speed data of the drafting machine, the model fails to accurately capture the dynamic changes in the data, resulting in a decrease in prediction accuracy.

(2) Removal of SKNet: After removing SKNet, the model's ability to process the residual data of the mechanistic and data-driven models weakens. This is particularly evident when the residuals of the data are large. Simply arranging the residual data in parallel, without SKNet's data processing, leads to a decrease in prediction accuracy.

(3) Advantages of the MD-S-L Configuration: The complete MD-S-L configuration performs the best in terms of performance, demonstrating the complementary roles of SKNet and LSTM in residual processing and temporal modeling.

Therefore, the combination of LSTM and SKNet significantly improves the model's prediction accuracy when dealing with data that has temporal dependencies and complex residuals. The removal of either of these modules results in a decline in

model performance.

Table 3. Ablation experiments (Unit: %).

Experimental	Prediction accuracy											
	Front roller			Two roller			Three roller			Rear roller		
	MD-S-L	MD-L	MD-S	M-D-S-L	MD-L	MD-S	MD-S-L	MD-L	MD-S	MD-S-L	MD-L	MD-S
test1	99.64	98.45	97.55	99.65	98.55	97.25	99.72	98.45	97.63	99.57	98.65	97.22
test2	99.66	98.64	97.54	99.64	98.54	97.24	99.74	98.42	97.62	99.55	98.64	97.25
test3	99.61	98.12	97.52	99.61	98.52	97.24	99.75	98.43	97.61	99.52	98.62	97.23
test4	99.64	98.43	97.43	99.62	98.43	97.27	99.71	98.43	97.61	99.54	98.63	97.23
test5	99.64	98.45	97.55	99.64	98.45	97.26	99.78	98.47	97.66	99.54	98.65	97.24
test6	99.59	98.44	97.54	99.60	98.54	97.25	99.72	98.41	97.62	99.52	98.64	97.21
test7	99.65	98.54	97.54	99.61	98.54	97.24	99.74	98.42	97.64	99.54	98.64	97.24
test8	99.63	98.32	97.52	99.64	98.52	97.23	99.74	98.38	97.62	99.57	98.62	97.25
test9	99.67	98.76	97.56	99.67	98.56	97.22	99.76	98.42	97.63	99.52	98.66	97.24
test10	99.68	98.81	97.51	99.62	98.51	97.25	99.74	98.43	97.64	99.51	98.61	97.23

In summary, after experimental evaluation, the hybrid model established through the multi-model residual learning method offers greater flexibility in data processing, greater relevance in mechanistic interpretation, and effectively balances accuracy and interpretability. Overall, the hybrid model demonstrates significant advantages.

5. Conclusions

Despite its higher accuracy, the multi-model residual learning method introduces additional computational costs. Compared to hybrid models established through parallel methods, training an additional data-driven model and handling the residuals from both models is required, which increases complexity. Future research could explore optimizing this method to reduce computational overhead or investigate the development of more accurate mechanistic models for hybrid modeling using parallel methods.

Through experiments, the performance of various models in predicting roller speed data for the FA320A drafting machine was evaluated. The main findings are as follows:

(1) Limitations of a single model: The single mechanistic model offers strong interpretability, but due to the system's complexity, simplifications are necessary during modeling, which affects accuracy. The data-driven model achieves high prediction accuracy, meeting the goal of providing predicted values, but lacks interpretability, making it difficult to analyze and solve specific issues in practical applications.

(2) Advantages and trade-offs of hybrid models: The parallel hybrid model improves prediction accuracy while maintaining interpretability, but its performance is somewhat dependent on the accuracy of the mechanistic model. When the mechanistic model experiences significant errors, the prediction stability of the parallel hybrid model is affected. The multi-model residual learning method reduces

the reliance on mechanistic model predictions but increases computational costs and model complexity.

In conclusion, the hybrid model established through the multi-model residual learning method, after experimental evaluation, achieved more than 95% accuracy in predicting the roller speeds of different rollers in the FA320A spinning drafting machine and demonstrated better robustness compared to the hybrid model established using the parallel method.

Author contributions: Conceptualization, XZ and ML; methodology, XZ; software, XZ; validation, XZ and ML; formal analysis, XZ; investigation, XZ; resources, ML; data curation, ML; writing—original draft preparation, XZ; writing—review and editing, ML; visualization, XZ; supervision, XZ; project administration, ML; funding acquisition, ML. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Catanese C, Ayassi R, Pincemin E, et al. A Fully Connected Neural Network Approach to Mitigate Fiber Nonlinear Effects in 200G DP-16-QAM Transmission System. In: Proceedings of the 2020 22nd International Conference on Transparent Optical Networks (ICTON). pp. 1-4. doi: 10.1109/icton51198.2020.9203197
2. Meng J, Xu X, Jiang C, et al. Tensile force field plays a crucial role in local invasion of tumor cells through a mechano-chemical coupling mechanism. Soft Matter. 2024; 20(30): 6002-6015. doi: 10.1039/d4sm00335g.
3. Ye X, Yang Y, Jiao P, et al. Underwater minirobots actuated by hybrid driving method. Journal of Zhejiang University-SCIENCE A. 2023; 24(7): 596-611. doi: 10.1631/jzus.a2300056
4. Siddiqui Q, Abro Z, Yu C. Study of drafting force variability and sliver irregularity at the break draft zone of a draw frame. Textile Research Journal. 2015; 85(14): 1465-1473. doi: 10.1177/0040517514563724
5. Zhang Z, Yu C. Study on drafting force and sliver irregularity on drawing frame. Journal of The Textile Institute. 2012; 103(3): 341-347.
6. Lee J, Wu F, Zhao W, et al. Prognostics and health management design for rotary machinery systems—Reviews, methodology and applications. Mechanical Systems and Signal Processing. 2014; 42(1-2): 314-334. doi: 10.1016/j.ymssp.2013.06.004
7. Du C, Gong F, Zhou Y, et al. A Mechanistic-Based Data-Driven Modelling Framework for Predicting Production of Electric Submersible Pump Wells in Offshore Oilfield. Published online 2024. doi: 10.2139/ssrn.4946667
8. Li Q, Niu F, Yang H, et al. Magnetically Actuated Soft Microrobot with Environmental Adaptative Multimodal Locomotion Towards Targeted Delivery. Advanced Science. 2024; 11(43). doi: 10.1002/advs.202406600
9. Vijay U, Galvez-Aranda DE, Zanotto FM, et al. A hybrid modelling approach coupling physics-based simulation and deep learning for battery electrode manufacturing simulations. Energy Storage Materials. 2025; 75: 103883. doi: 10.1016/j.ensm.2024.103883
10. Michalopoulou F, Papathanasiou MM. An approach to hybrid modelling in chromatographic separation processes. Digital Chemical Engineering. 2025; 14: 100215. doi: 10.1016/j.dche.2024.100215
11. Roland W, Marschik C, Kommenda M, et al. Predicting the Non-Linear Conveying Behavior in Single-Screw Extrusion: A Comparison of Various Data-Based Modeling Approaches used with CFD Simulations. International Polymer Processing. 2021; 36(5): 529-544. doi: 10.1515/ipp-2020-4094
12. Leite WL, Jing Z, Kuang H, et al. Multilevel Mixture Modeling with Propensity Score Weights for Quasi-Experimental Evaluation of Virtual Learning Environments. Structural Equation Modeling: A Multidisciplinary Journal. 2021; 28(6): 964-982. doi: 10.1080/10705511.2021.1919895

13. Linthicum DS. Connecting Fog and Cloud Computing. IEEE Cloud Computing. 2017; 4(2): 18-20. doi: 10.1109/mcc.2017.37
14. Gusev M, Dustdar S. Going Back to the Roots—The Evolution of Edge Computing, An IoT Perspective. IEEE Internet Computing. 2018; 22(2): 5-15. doi: 10.1109/mic.2018.022021657
15. Mitrofanov S, Semenkin E. An Approach to Training Decision Trees with the Relearning of Nodes. In: Proceedings of the 2021 International Conference on Information Technologies (InfoTech). pp. 1-5. doi: 10.1109/infotech52438.2021.9548520
16. Xu W, Peng H, Zeng X, et al. A Hybrid Modeling Method Based on Linear AR and Nonlinear DBN-AR Model for Time Series Forecasting. Neural Processing Letters. 2021; 54(1): 1-20. doi: 10.1007/s11063-021-10651-2
17. Hu X, Shi L, Lin G, et al. Comparison of physical-based, data-driven and hybrid modeling approaches for evapotranspiration estimation. Journal of Hydrology. 2021; 601: 126592. doi: 10.1016/j.jhydrol.2021.126592
18. Panchikkil S, Manikandan VM, Zhang YD. A convolutional neural network model based reversible data hiding scheme in encrypted images with block-wise Arnold transform. Optik. 2022; 250: 168137. doi: 10.1016/j.ijleo.2021.168137
19. Pally RJ, Samadi S. Application of image processing and convolutional neural networks for flood image classification and semantic segmentation. Environmental Modelling & Software. 2021; 105-285.
20. Chen Z. Machine remaining useful life prediction method based on global-local attention compensation network. Reliability Engineering & System Safety. 2025; 255: 110652. doi: 10.1016/j.ress.2024.110652
21. Peng G, Qiu H, Li Y, et al. A Sub-Channel Spatial Homogeneity-Based Channel Estimation Method for Underwater Optical Densely Arrayed MIMO Systems. Journal of Marine Science and Engineering. 2024; 12(11): 2030. doi: 10.3390/jmse12112030
22. Cao Y, Ren X, Guo Q, et al. Model Reference Adaptive System of Permanent Magnet Synchronous Motor Based on Current Residual Compensation Without Position Measurement. Actuators. 2024; 13(11): 446. doi: 10.3390/act13110446
23. Li Y, Chen X. Attributed graph subspace clustering with residual compensation guided by adaptive dual manifold regularization. Expert Systems with Applications. 2024; 255: 124699. doi: 10.1016/j.eswa.2024.124699
24. Zhang X, Huang Z, Li Q, et al. Legged robot-aided 3D tunnel mapping via residual compensation and anomaly detection. ISPRS Journal of Photogrammetry and Remote Sensing. 2024; 214: 33-47. doi: 10.1016/j.isprsjprs.2024.05.025
25. Liu J, Yang X. Remote sensing image super-resolution based on cross residual compensation mechanism and attention mechanism. Multimedia Tools and Applications. 2023; 83(18): 55641-55657. doi: 10.1007/s11042-023-17645-3
26. Yang F, Chen X, Chai L. Optimal graph edge weights driven nlms with multi-layer residual compensation. EURASIP Journal on Advances in Signal Processing. 2021; 2021(1). doi: 10.1186/s13634-021-00800-z
27. Sustainability Research-Sustainable Development; Study Findings on Sustainable Development Detailed by Researchers at Memorial University of Newfoundland (Goal-Driven or Data-Driven Inventory of Sustainability Indicator Initiatives in Rural Canada). Information Technology Newsweekly. 2020; 994-998. Available: https://xueshu.baidu.com/usercenter/paper/show?paperid=130r0gx0ky120ed0f32v009078272700&site=xueshu_se (accessed on 10 March 2025)
28. Nekoonam A, Montazeri-Gh M. Noise-robust gas path fault detection and isolation for a power generation gas turbine based on deep residual compensation extreme learning machine. Energy Science & Engineering. 2023; 11(11): 4001-4018. doi: 10.1002/ese3.1576
29. Chen H, Hu L, Wang H, et al. Comparison of Active Full Compensation Control Algorithms for Single-phase Grounding Residual Current in Distribution Networks. Journal of Physics: Conference Series. 2023; 2530(1): 012011. doi: 10.1088/1742-6596/2530/1/012011
30. Information Technology; New Data from Russian Academy of Sciences Illuminate Findings in Information Technology (Data-driven Modeling of Stomatal, Mesophyll, and Biochemical Regulation of Scots Pine Photosynthesis). Information Technology Newsweekly. 2020; 350-360. Available: https://xueshu.baidu.com/usercenter/paper/show?paperid=1f0a0v2053160mg0ym180tn0rn477618&site=xueshu_se&hitarticle=1 (accessed on 10 March 2025)
31. Mo L, Xiao X, Chen Y, et al. A Rapid Decompression Failure Mechanism Model Coupling Diffusion-Deformation of O-Ring Sealing Rubber Exposed to High-Pressure Carbon Dioxide. Published online 2024. doi: 10.2139/ssrn.5021481
32. Guo F, Zhang H, Song C, et al. Identification and Precise Optimization of Key Assembly Error Links for Complex Aviation Components Driven by Mechanism and Data Fusion Model. Published online 2024. doi: 10.2139/ssrn.4944504

33. Xie N, Xu W. Analysis of the Evaluation Data of College Classroom Teaching Quality based on Fuzzy Mathematics Model. Frontiers in Computing and Intelligent Systems. 2024; 10(3): 86-94. doi: 10.54097/h7fd6821
34. Tuly SS, Joardder MUH, Welsh ZG, et al. A novel mechanistic model for predicting shrinkage kinetics in plant-based foods by integrating solid matrix mobility and viscoelasticity. Journal of Food Engineering. 2025; 387: 112346. doi: 10.1016/j.jfoodeng.2024.112346