ISSN 1392-1207. MECHANIKA. Vol. 32, No. 3, 2026: 251–256

Experimental Research on Thermal Deformation Error of CNC Machine Tools Based on Neural Network

Bo YU*, Xiao-peng CHANG**
*School of Mechanical and Electrical Engineering, Changchun Institute of Technology, Changchun 130012, Jilin, China
**School of Mechanical and Electrical Engineering, Changchun Institute of Technology, Changchun 130012, Jilin, China, E-mail: 1202201011@stu.ccut.edu.cn (Corresponding Author)

https://doi.org/10.5755/j02.mech.44179

1. Introduction

Driven by high-end manufacturing, the precision requirements for CNC machine tool machining are continuously increasing, while thermal error accounts for approximately 40%-70% of the total machine tool error, becoming a key factor restricting precision [1, 2]. Traditional thermal error modeling methods (such as multiple linear regression and static neural networks) mostly rely on instantaneous temperature-error mapping, ignoring the dynamic temporal characteristics of the thermal process, which leads to "compensation lag" caused by time delays in data acquisition, calculation, and compensation execution during actual machining, severely affecting the compensation effect [3]. Although existing time-series models based on Recurrent Neural Networks (RNN) and their variants have made some progress, two major limitations still exist: first, most models only establish a direct mapping between temperature and error, failing to explicitly decouple the two stages of temperature prediction and error mapping; second, the prediction lead time is usually short or not clearly defined, making it difficult to meet the response time requirements of actual compensation systems [4, 5].

To address the aforementioned problem, this paper proposes a two-stage predictive method based on Long Short-Term Memory (LSTM) networks. Its innovativeness is in the following aspects: first, the prediction task is explicitly decomposed into two serial stages-temperature time-series advance prediction and temperature-thermal error mapping-thereby introducing physical interpret; second, by analyzing the thermal characteristics of the machine tool and the response time of the compensation system, a 10-minute advance prediction step is determined to reserve a sufficient window compensation execution; finally, the entire process is validated using an aspherical CNC grinder as the experimental platform. Experimental results demonstrate that this method significantly outperforms the multiple linear regression modeling in both prediction accuracy and stability, providing an effective technical path for solving the compensation lag problem [6].

2. Theoretical Analysis of Ahead-of-Time Prediction

The Recurrent Neural Network (RNN), proposed in 1997, captures temporal dependencies in sequences through recurrent connections but suffers from the long-term dependency problem, where training on long sequences is prone to gradient vanishing or explosion [7, 8]. To address this, the Long Short-Term Memory (LSTM) network introduces memory cells and three gating mechanisms: the forget gate, input gate, and output gate building upon the RNN framework. The gate outputs a continuous value in the (0,1) interval via the Sigmoid function, representing the proportion of information passing through, where 0 indicates complete and 1 indicates complete passage.

In any RNN algorithm, let i, j, and k represent the number of nodes in the input layer, hidden layer, and output layer, respectively. Let $X = \{X_1, X_2, X_3, \cdots, X_i\}$ and $Y = \{Y_1, Y_2, Y_3, \cdots, Y_k\}$ denote the network input samples and network output samples, respectively. Let U_{ij} represent the weight from input layer node i to hidden layer node j, and V_{jk} represent the weight from hidden layer node j to output layer node k.

When the RNN is unfolded along the time dimension, its topological structure is illustrated in Fig. 1.

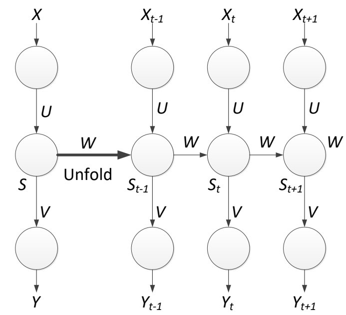

Fig. 1 RNN unfolding structure diagram

In Fig. 1, $X = \{X_1, X_2, X_3, \cdots, X_t, \cdots, X_i\}$ and $Y = \{Y_1, Y_2, Y_3, \cdots, Y_t, \cdots, Y_k\}$ represent the network input samples and network output samples, respectively $S = \{S_1, S_2, S_3, \cdots, S_t, \cdots, S_i\}$ denotes the hidden layer outputs. U, V, and W are the connection weights from the input layer to the hidden layer, from the hidden layer to the output layer, and from the hidden layer to the hidden layer, respectively. Thus, the hidden layer output S_t at any time step t is calculated based on the previous hidden layer output S_{t-1} and the current input X_t, as shown in Eq. (1)

$$S_t = f\left(UX_t + WS_{t-1}\right) + b_t . \quad (1)$$

In Eq. (1), the function $f(UX_t + WS_{t-1})$ represents a nonlinear activation function, and b_t denotes the threshold of the current hidden layer.

The LSTM algorithm achieves long-term retention of historical information by monitoring the state of memory cells. The basic structure of an LSTM memory cell is illustrated in Fig. 2.

In Fig. 2, X_t is the input to the memory cell, while

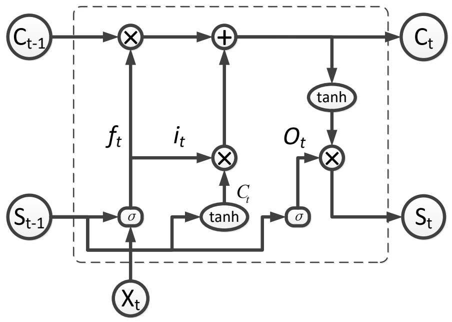

Fig. 2 LSTM memory cell

S_t and S_{t-1} represent the outputs of the memory cells at adjacent time steps. C_t and C_{t-1} denote the states of the memory cells at adjacent time steps. The symbol σ represents the sigmoid function, and ⊗ denotes the element-wise multiplication operation. The gating units of the LSTM algorithm, from right to left, are the output gate, the input gate, and the forget gate. The output of the σ function is limited to either 0 or 1, where 0 signifies complete blockage and 1 signifies full passage.

The forget gate controls the retention ratio of the previous state C_{t-1} based on the previous output S_{t-1} and the current input X_t. Assuming b_f and W_f represent the threshold (bias) and weights in the forget gate operation, the output f_t can be calculated as shown in Eq. (2)

$$f_t = \sigma\left[W_f\left(S_{t-1}, X_t\right) + b_f\right]. \tag{2}$$

The input gate is used to store new information into the memory cell, while the candidate memory information $\bar{C}_t$ is computed via the *tanh* function. Assuming b_c and W_c represent the threshold (bias) and weights in the input gate operation, $\bar{C}_t$ can be derived from Eq. (3)

$$\bar{C}_t = tanh\left[W_c\left(S_{t-1}, X_t\right) + b_c\right]. \tag{3}$$

Assuming b_i and W_i represent the threshold (bias) and weights for the input gate operation, the input gate value i_t can be derived from Eq. (4)

$$i_t = \sigma\left[W_i\left(S_{t-1}, X_t\right) + b_i\right]. \tag{4}$$

Based on the results of the input gate and forget gate calculated from Eqs. (1) to (4), the memory cell state C_t can be derived using Eq. (5)

$$C_t = f_t \otimes C_{t-1} + i_t \otimes \bar{C}_t, \tag{5}$$

$$\begin{cases} O_t = \sigma\left[W_o\left(S_{t-1}, X_t\right) + b_o\right] \\ S_t = O_t \otimes tanh\left(C_t\right) \end{cases}. \tag{6}$$

In Eq. (6), b_o represents the threshold (bias) of the sigmoid function, W_o denotes its weight, and O_t is the output of the sigmoid function.

The LSTM algorithm utilizes the operational formulas of the output gate, input gate, and forget gate from Eqs. (1) to (5), along with Equation Set (6), to achieve forward propagation of key information.

3. Ahead-of-Time Prediction Analysis

To achieve effective ahead-of-time prediction of thermal errors in CNC machine tools, this study proposes a two-stage serial modeling strategy. The objective of the first stage is to establish a time-series ahead-of-time prediction model for key temperature measurement points, i.e., using historical temperature data to predict temperature values for multiple future time steps. The second stage involves constructing a mapping model between temperature and thermal error, where the predicted future temperatures from the first stage serve as inputs, thereby indirectly enabling the prediction of future thermal errors. The core of the overall scheme lies in the LSTM time-series prediction model of the first stage.

3.1. Experimental data collection and preprocessing

The experimental platform is an aspheric CNC grinding and polishing machine independently developed by the laboratory. To obtain the dataset required for model training and testing, simulated machining cycle experiments were designed. Each experiment lasted 6 hours, simulating the heating and cooling processes of the machine tool during actual machining. The specific experimental conditions are as follows: PT10 platinum resistance thermometers were used as temperature sensors, the laser displacement sensor model was Keyence LK-H020, and the synchronous sampling period of the data acquisition card was set to minute. The ambient temperature was maintained within 20±1°C by air conditioning, the spindle running speed was constant at 3000 r/min, the heating consisted of continuous spindle operation, and the cooling phase was natural heat dissipation after shutdown. Multiple temperature sensors were arranged on the machine tool to monitor the temperature changes of key parts such as the, ball screw nut pair, column, and bed. Among them, the specific installation positions of the three key measurement points T1, T6, and T12, screened through clustering and grey relational analysis, are shown in Table 1.

Table 1

Locations of key temperature measurement points

Sensor ID	Installation position description
T1	Outer wall of the front bearing housing of the spindle
T6	Bed near the center of the worktable
T12	Middle surface of the rear side of the column

Meanwhile, a high-precision laser displacement sensor is used to measure the axial thermal deformation error of the grinding wheel spindle. The data acquisition system synchronously records the readings of temperature measurement points and the thermal error value of the spindle once per minute. Therefore, a single experiment generates a data sequence containing 360 time steps. A total of 180 sets of time-synchronized temperature-thermal error data pairs are obtained from 5 experiments, forming the data basis of this study.

Fig. 3 presents the temperature variation curves of three representative key temperature measurement points (T1, T6, T12) optimized through fuzzy clustering and grey relational analysis across five experiments. Fig. 4 displays the corresponding axial thermal error variation curves of the

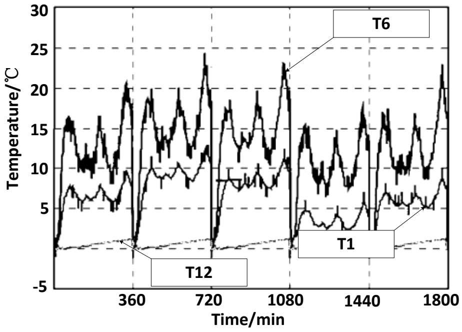

Fig. 3 Temperature change curve

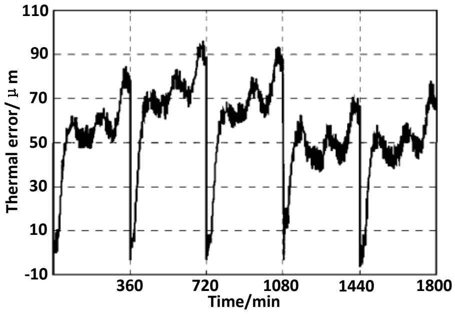

Fig. 4 Thermal error curve

grinding wheel spindle. It can be observed that both temperature and thermal error exhibit typical time-varying characteristics: a rapid increase during the startup phase, followed by a relatively stable yet fluctuating state, and a gradual decline after shutdown. While a clear correlation exists between the temperature and thermal error curves, they do not demonstrate a simple immediate correspondence. Instead, a certain phase difference and dynamic response relationship are evident, highlighting the necessity of constructing a time-series prediction model.

3.2. Construction of the LSTM ahead-of-time prediction model

Model training adopts a supervised learning approach. To strictly avoid data leakage, the data from 5 complete experiments are partitioned chronologically: the first3 experiments (1080 samples) serve as the training set, the 4th experiment (360 samples) as the validation set, and the 5th experiment (30 samples) as the test set, without using random shuffling or cross-experiment sliding window partitioning. Input features and labels are Z-score normalized, with the mean and standard deviation calculated from training set and then applied to the validation and test sets. The training batch size is set to 32, the optimizer is Adam, the loss function is mean squared error, and the early criterion terminates training when the validation set loss does not decrease for 50 consecutive epochs. The model is implemented based on the PyTorch 2.0 framework, and a fixed random seed of 2 is used for all experiments to ensure reproducibility. Training data are generated using a sliding window method: from a sequence with a total length of 1800, segments of length60 are sequentially extracted as input samples X, and the temperature value at the 10th time point after the starting point of the segment is taken as the label Y for that sample. window slides forward by 1 time step each time, generating a large number of overlapping training samples. The goal of model training is to minimize the mean squared error between the predicted temperature value and actual temperature value. The optimizer selected is the adaptive moment estimation optimizer, and the parameter settings such as the learning rate are shown in Table 2.

Table 2

Parameters of the LSTM Algorithm

No.	Name	Value
1	Time horizon	10 minutes
2	Time step	60 minutes
3	Number of memory cells	90
4	Learning rate	0.000001
5	Number of epochs	1000
6	Weight adjustment ratio per iteration	0.0001

Regarding parameter selection: the look-back window (60 minutes), the number of memory units (90), and the learning rate (10^{-6}) were initially determined through grid search, with search ranges of window [30, 60, 90], units [32, 64, 90, 128], and learning rate [10^{-5}, 10^{-6}, 10^{-7}], respectively, selecting the final values based on the criterion of minimizing validation set loss; the number of iterations (1000 epochs) is used in conjunction with an early stopping mechanism, with the actual termination epoch determined by the validation set loss; weight decay coefficient (0.0001) was obtained by fine-tuning based on common settings for LSTM time-series forecasting. The key temperature measurement points T1, T6 and T12, optimized using fuzzy clustering and grey relational algorithms, are used as training samples for the temperature data to experimentally verify the LSTM algorithm.

Let the input matrix be $T_M \in R^{60\times3}$ (60 time steps, 3 temperature measurement points), the LSTM network mapping function be $f_{LSTM}()$, then the temperature prediction output 0 minutes in advance is $T_P \in R^3$, $T_P \in R^3$, given by the following formula, where W and b are the trainable weights and biases of the network, respectively

$$T_P = f_{LSTM}\left(T_M, W, b\right). \quad (7)$$

4. Experimental Analysis

4.1. Ahead-of-time temperature prediction results

To objectively evaluate the performance of the ahead-of-time temperature prediction model, data from key temperature measurement points T_1, T_6, and T_{12} during the period from 1101 to 1111 minutes were selected from the test dataset for validation. This period includes multiple states such as temperature rise, relative stability, and slight fluctuations, making it highly representative. The model's prediction results were compared with the actual measured values, as shown in Figs. 5-7.

From the Figs. 5-7, it can be intuitively observed that for all three key measurement points, the LSTM model's 10-minute ahead prediction curves align very closely with the actual temperature curves. Whether in terms of trend changes or subtle fluctuations, the model demonstrates highly accurate tracking and prediction capabilities. This provides preliminary evidence that the model effectively captures the temporal dynamics governing temperature variations.

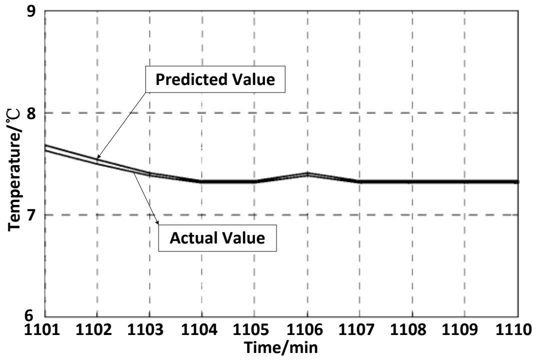

Fig. 5 Prediction result of T1

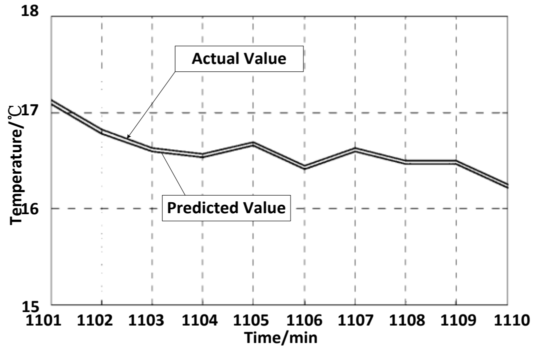

Fig. 6 Leading prediction result of T6

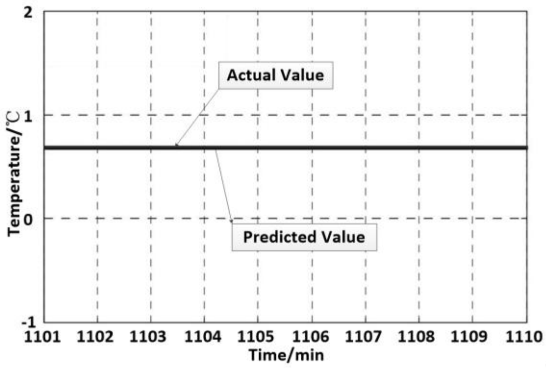

Fig.7 Leading prediction result of T12

To quantitatively assess prediction accuracy, the residuals between the predicted values and the actual values were calculated, and their maximum, mean, and root mean square error (RMSE) were summarized in Table 3. The residual mean Δe and residual standard deviation RMSE are defined as shown in Eq. (8), and the corresponding error values in Tables 3-5 of this paper are calculated according to the definition in Eq. (8)

$$\begin{cases} \Delta e = \dfrac{1}{n}\sum_{i=1}^{n} e_i \\ RMSE = \sqrt{\dfrac{1}{n}\sum_{i=1}^{n} e_i^2} \end{cases} . \tag{8}$$

Analysis of the data in Table 3 reveals the following:

1. **Mean Residual Error.** The mean residual error for all three measurement points is significantly below 0.05°C, indicating that the model exhibits minimal systematic bias, and the overall prediction results are unbiased.

Table 3

Model test data

No.	Name	Key Measurement Point /°C		
		T1	T6	T12
1	Δe	0.025	0.033	0.021
2	Max R	0.048	0.048	0.023
3	RMSE	6.39e-004	1.57e-003	3.61e-004

2. **Maximum Residual Error.** The maximum deviation is also controlled within 0.05°C. In the context of temperature field monitoring for precision machine tools, temperature fluctuations of 0.05°C generally fall within the range of sensor accuracy and noise levels. Therefore, this maximum prediction deviation can be considered fully acceptable in practical engineering applications, even regarded as a "negligible magnitude."

3. **Root Mean Square Error (RMSE).** The RMSE values are very small (on the order of 10^{-3} to 10^{-4}), reflecting minimal fluctuations of the predicted values around the actual values. This demonstrates excellent stability and consistency in the predictions.

4. **Comprehensive Analysis of Charts and Data.** Based on the combined analysis of charts and data, the LSTM-based ahead-of-time temperature prediction model demonstrates outstanding performance. It successfully achieves high-precision estimation of temperatures 10 minutes into the future, laying a solid foundation for accurate ahead-of-time prediction of thermal errors in subsequent steps.

4.2. Ahead-of-time thermal error prediction results

The trained thermal error mapping model was connected in series with the ahead-of-time temperature prediction model to forecast the spindle axial thermal error 10 minutes ahead for the period from 1201 to 1261 minutes. A comparison between the predicted results and the actual measured values is shown in Fig. 8.

From Fig. 8, it can be observed that the predicted curve of thermal error closely aligns with the actual curve in terms of variation trends throughout the entire time period. In regions where thermal error changes gradually, the two curves are nearly identical. In areas with fluctuations or turning points, the predicted curve promptly follows the actual trend, exhibiting only slight deviations in amplitude or

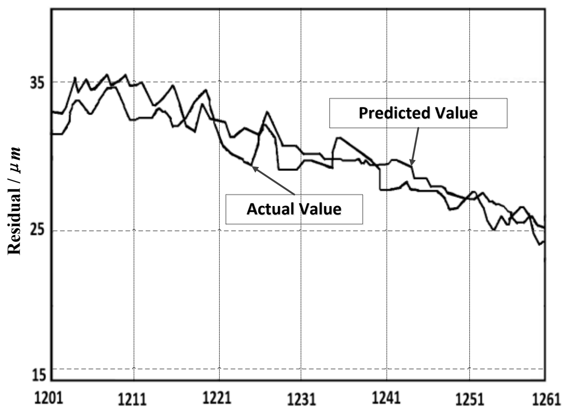

Fig. 8 Thermal error forward prediction results

Table 4

Thermal error test data

No.	Name	Value, μm
1	Δe	1.631
2	Max R	4.912
3	RMSE	0.945

phase. This clearly demonstrates that the cascaded prediction system successfully transfers the ahead-of-time prediction capability from temperature to thermal error.

A quantitative statistical analysis of the prediction errors was conducted, and the results are summarized in Table 4.

The data in Table 4 show that during the 60-minute ahead-of-time prediction test, the mean residual of thermal error prediction is 1.631 μm, the maximum residual is 4.912 μm, and the root mean square error (RMSE) is 0.945 μm. Considering that the machining accuracy requirement of this aspheric grinding machine is at the micro meter level, and the thermal error itself can reach tens or even dozens of micro meters, this prediction accuracy holds significant engineering application value. The small RMSE indicates low dispersion of prediction errors, demonstrating the stability and reliability of the model output. These data strongly support the conclusion that the two-stage ahead-of-time prediction model based on LSTM can effectively achieve accurate ahead-of-time estimation of thermal deformation errors in CNC machine tools, providing a feasible technical solution to address the issue of compensation lag.

4.3. Comparative analysis

To highlight the advantages of the LSTM time-series prediction method proposed in this paper, a comparative experiment was conducted with the widely used traditional static thermal error modeling method based on multiple linear regression. The multiple linear regression model assumes an immediate linear relationship between thermal error and multiple temperature measurement points. The same training data were used to fit the multiple linear regression model.

The multiple linear regression model adopts the exact same data partitioning as the LSTM model (training on the first 3 experiments, validation on the 4th, and on the 5th), and performs predictions under the same 10-minute lead time (i.e., inputting data from time t-59 to t to predict error at time t 10) to ensure a fair comparison.

Within the same test period (1204 to 1260 minutes), sampling was performed at 4-minute intervals. The trained multiple linear regression model and the LSTM ahead-of-time prediction model proposed in this paper were used to predict thermal errors, and the prediction residuals of both methods were calculated. Detailed comparative data are presented in Table 5.

Based on the comparative analysis of the data in Table 5, the following conclusions can be drawn:

1. Overall Accuracy: The mean residual of the LSTM ahead-of-time prediction model (2.526 μm) is significantly lower than that of the multiple linear regression model (3.463 μm), representing a reduction of approximately 27%. This indicates that, overall, the predictions of the LSTM model are closer to the true values.

2. Worst-Case Performance: The maximum residual of the LSTM model (3.764 μm) is much lower than that of the multiple linear regression model (8.137 μm). This means that at the most challenging prediction points, the LSTM model controls prediction errors within a smaller range, demonstrating stronger robustness.

Table 5

Residual data comparison

Sampling time, min	Multiple linear regression modeling residual, μm	
	Multiple linear regression modeling	LSTM algorithm modeling
1204	7.642	2.931
1208	3.460	0.847
1212	5.297	1.320
1216	3.601	3.764
1220	0.194	2.142
1224	4.426	3.019
1228	8.137	3.235
1232	2.664	2.896
1236	2.539	1.237
1240	3.118	0.358
1244	3.979	3.566
1248	3.157	2.671
1252	0.297	3.764
1256	2.551	3.145
1260	0.896	3.009
Δe	3.463	2.526
Max R	8.137	3.764
RMSE	2.2169	1.0524

3. Prediction Stability: The root mean square error (RMSE) of the LSTM model (1.0524) is less than half of that of the multiple linear regression model (2.2169). A smaller RMSE indicates lower fluctuations in the LSTM model's predictions, higher consistency, and more stable and reliable performance.

4. The comparative experiments robustly validate that the ahead-of-time prediction method based on LSTM neural networks not only theoretically addresses the issue of compensation lag but also demonstrates superior prediction accuracy and stability over traditional static model methods in practice. This highlights its significant technological advancement and application potential.

5. Conclusions

1. Through literature review, this study summarizes the critical issue currently faced in the implementation of thermal error compensation for CNC machine tools: "the time delay between thermal error prediction and actual compensation leads to reduced compensation accuracy." It proposes the application of the LSTM algorithm to address this problem and discusses the implementation steps of the LSTM-based approach.

2. An LSTM-based ahead-of-time thermal error prediction model was established. Experiments were conducted using temperature data from key measurement points T1, T6, and T12, as well as axial thermal error data of the grinding wheel spindle. The results demonstrate that the LSTM-based ahead-of-time thermal error prediction achieves high accuracy.

3. A comparative analysis was performed between traditional thermal error prediction using multiple linear regression and LSTM-based ahead-of-time prediction. The latter not only ensures timely thermal error prediction but also further enhances prediction accuracy. In terms of

maximum residual error, mean residual error, and root mean square error, the LSTM-based ahead-of-time prediction exhibits superior precision, further validating the effectiveness of the proposed methodology.

Acknowledgments

This work is supported by the Key Research and Development Projects of Jilin Province Science and Technology Agency, No.20240302050GX.

References

1. **Peng, J.; Yang, K.; Zhang, Z.; Hu, X.; Yin, M.; Mao, Y.; Zhao, Z.; Li, L.; Yuan, C.** 2026. Modeling of the radial thermal error in precision CNC machine tools induced by ambient temperature rise, Advanced Engineering Informatics 74: 104780. https://doi.org/10.1016/j.aei.2026.104780.
2. **Mayr, J.; Jedrzejewski, J.; Uhlmann, E.; Donmez, M.; Knapp, W.; Härtig, F.; Wendt, K.; Moriwaki, T.; Shore, P.; Schmitt, R.; Brecher, C.; Würz, T.; Wegener, K.** Thermal issues in machine tools, CIRP Annals 61(2): 771-791. https://doi.org/10.1016/j.cirp.2012.05.008.
3. **Li, Y.; Zhao, W.; Lan, S.; Ni, J.; Wu, W.; Lu, B.** 2015. A review on spindle thermal error compensation in machine tools. International Journal of Machine Tools and Manufacture, 95: 20-38. https://doi.org/10.1016/j.ijmachtools.2015.04.008.
4. **Yin, L.; Li, C.; Peng, Y.; Tang, H.; Wang, N.; Chen, H.** 2026. A Novel Hybrid Neural Network with Optimized Feature Selection for Spindle Thermal Error Prediction, Applied System Innovation 9(2): 40. https://doi.org/10.3390/asi9020040.
5. **Liu, Y.; Wang, X.; Zhu, X.; Zhai, Y.** 2021. Thermal error prediction of motorized spindle for five-axis machining center based on analytical modeling and BP neural network, Journal of Mechanical Science and Technology 35(1): 281-292. https://doi.org/10.1007/s12206-020-1228-7.
6. **Wang, W.; Zhang, Y., Fan, K., Yang, J.** 2013. A Fourier Series-Neural Network Based Real-Time Compensation Approach for Geometric and Thermal Errors of CNC Milling Machines, Advances in Mechanical Engineering 5: 357920. https://doi.org/10.1155/2013/357920.
7. **Kim, K.; Sasahara, H.; Imura, I. J.** 2025. Adaptive false data injection attack detection in load frequency control using recurrent neural networks, SICE Journal of Control, Measurement, and System Integration 18(1): 2597566. https://doi.org/10.1080/18824889.2025.2597566.
8. **Zeng, F.; Jiang, C.; Liu, P.** 2026. An artificial neural network-based cyclic p-y curve model for predicting the lateral cyclic response of piles in sand, Soil Dynamics and Earthquake Engineering 207: 110352. https://doi.org/10.1016/j.soildyn.2026.110352.

B. Yu, X. Chang

EXPERIMENTAL RESEARCH ON THERMAL DEFORMATION ERROR OF CNC MACHINE TOOLS BASED ON NEURAL NETWORK

S u m m a r y

Thermal deformation error of CNC machine tools accounts for 40% to 70% of the total machining error, acting as a critical bottleneck precision. Existing thermal error prediction methods are mostly based on static instantaneous mapping, ignoring the temporal evolution characteristics of thermal errors, and suffer from compensation lag due to delays in data acquisition, model computation, and compensation execution, degrades machining quality. To address this issue, this paper proposes a two-stage proactive thermal error prediction method based on Long Short-Term Memory (LSTM) neural networks: first, a lead prediction model is established using historical time-series temperature data, and then a temperature-thermal error mapping model is constructed to achieve advance estimation of the thermal error for the next 0 minutes. Experimental validation was completed by collecting time-series temperature and thermal error data on a self-developed aspherical CNC grinder. The results show that the maximum value, mean value and root mean square error of the LSTM model's prediction residuals are all at a low level, and the prediction accuracy is significantly superior to that of the traditional multiple linear regression model, verifying the effectiveness advanced nature of this method in solving the compensation lag problem and improving the machining precision of machine tools.

Keywords: CNC machine tools, thermal deformation error, compensation lag, ahead-of-time prediction, Long Short-Term Memory (LSTM) neural network, time-series modeling.

Received January 28, 2026
Accepted June 19, 2026