Thermal error modeling of electric spindle based on particle swarm optimization-SVM neural network

Zhaolong Li1,2 · Wenming Zhu1 · Bo Zhu1 · Baodong Wang1 · Qinghai Wang1

Received: 2 June 2022 / Accepted: 20 July 2022 / Published online: 23 July 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2022

Abstract
High-speed electric spindle is an important part of computer numerical control (CNC) machining equipment. The thermal displacement generated by the electric spindle during operation is the main reason that affects the machining stability and machining accuracy of the electric spindle. Compensating the thermal error of the high-speed electric spindle can effectively improve the CNC machining. Improve equipment processing performance. Therefore, it is particularly important to establish the accuracy of the thermal error prediction model. Taking the A02 high-speed electric spindle as the research object, ANSYS is used to analyze the thermal characteristics of the electric spindle, and the temperature and thermal displacement monitoring points of the electric spindle are arranged according to the simulation results, and the temperature and thermal displacement data of the monitoring points under different rotational speeds are collected; using K-means to classify temperature measurement points, uses the gray relation analysis degree to determine the correlation between the temperature measurement point and the thermal displacement data, and selects 4 temperature-sensitive points from 10 temperature measurement points. Finally, particle swarm optimization (PSO) is used to optimize the penalty factor and kernel function of support vector machine (SVM), and the PSO-SVM prediction model is established to compare with the neural network prediction model of SVM and genetic algorithm (GA) optimized SVM. The results show that PSO-SVM has better robustness, stability, and generalization ability.

Keywords Electric spindle · Thermal error · Temperature measurement point · Particle swarm optimization · Thermal error modeling · SVM neural network

1 Introduction
During the long-term high-speed operation of the machine tool, the internal heat dissipation of the high-speed electric spindle is insufficient and slow, which causes the internal temperature to rise rapidly, resulting in thermal deformation of the spindle and affecting the machining accuracy of the machine tool. Studies have shown that the core component of the precision machine tool is the high-speed electric spindle. In the field of finishing, the machining error caused by the heat generation of the spindle of the machine tool accounts for 40–70% of the total error of the machine tool. Therefore, the thermal deformation of the high-speed electric spindle is the main reason that affects the machining accuracy of the machine tool [1–3]. To ensure the machining accuracy of the machine tool, it is necessary to minimize the influence of the thermal error of the electric spindle on the machining accuracy. At this stage, there are three methods to reduce the thermal error of machine tools: thermal error prevention, temperature control, and thermal error compensation method [4]. Because the thermal error compensation method can compensate the thermal error without changing the spindle structure and improve the machining accuracy of the machine tool, it has been widely used at this stage.

The rationality of the selection of temperature measurement points is particularly important. Too many temperature measurement points increase the amount of calculation, and too few measurement points cannot fully reflect the temperature change of the electric spindle, which affects the accuracy and robustness of the error compensation model. The
commonly used method for selecting temperature measurement points is the combination of cluster analysis and gray relation analysis degree. Commonly used cluster analysis includes fuzzy clustering, K-means mean clustering, pedigree clustering, or a combination of the two clustering methods. For example, the following are groupings of used temperature measurement points. Yue et al. [5] used the fuzzy c-means clustering method to establish a temperature compensation model, which reduced the computational complexity of the neural network prediction model. Yang et al. [6] combined fuzzy clustering and correlation to screen temperature measurement points, which ensured the robustness of the prediction model and improved the accuracy of the model. Liu et al. [7] used the K-means clustering method to combine the temperature-sensitive points, and screened out the combination of the best temperature-sensitive points, which improved the accuracy of the prediction model. Fu et al. [8] used the K-means mean clustering method to select the combination of temperature-sensitive points of the electric spindle by setting the K value to ensure the robustness and accuracy of the prediction model. Huang et al. [9] combined the empirical compensation function with numerical modeling to accurately determine the location and number of temperature sensors. Cui and Wang [10] used the moving heat source method to establish a computational model of fluid–structure interaction, and combined it with the finite element method to analyze thermal characteristics under different working conditions. This method provides a basis for the analysis of bearings in the motorized spindle. Chen et al. [11] used the finite element analysis method to analyze the heat transfer mechanism of the electric spindle, and optimized the electric spindle according to the analysis results to make the temperature distribution of the optimized electric spindle more uniform. Li et al. [12] proposed a thermal–mechanical coupling calculation method for the deformation error of the electro-spindle to calculate the steady-state thermal deformation of the electro-spindle, and verified it with finite element analysis.

The accuracy of thermal error compensation technology depends on the accuracy and robustness of thermal error model establishment. Scholars at home and abroad have carried out a lot of research on thermal error modeling technology, among which thermal error modeling methods include long short-term memory (LSTM), convolutional neural network, multivariable linear regression, random forest method, BP neural network, kriging model, and gray prediction model [13, 14]. Long short-term memory (LSTM) is an improved algorithm of recurrent neural network (RNN). The LSTM prediction model established based on the relationship between the temperature and thermal error of the electrical spindle is a good ideal prediction model for error relationships. Wu et al. [15] adopted a high-precision and robust convolutional neural network prediction model, which is characterized by the fact that the larger the amount of data, the higher the prediction accuracy. Zhou and Wang [16] used regression analysis to establish a thermal error prediction model to improve the machining accuracy of the electric spindle. Zhu et al. [17] used iterative removal of unnecessary features and used random forest method to establish a prediction model, which requires less data, higher prediction accuracy, and better robustness. Liu and Li et al. [18, 19] used the bat algorithm and the improved particle swarm optimization to optimize the weights and thresholds of the BP neural network, and established a high-precision BP prediction model. Guo et al. [20] and others used particle swarm optimization (PSO) to optimize the gray neural network to avoid the local optimal problem of gray network prediction results, and the prediction model improved the convergence and prediction accuracy of the network model.

The support vector machine (SVM) has strong advantages in dealing with linear and nonlinear problems, and has good prediction performance, but SVM neural network has problems such as slow convergence speed and easy precociousness. To solve the above problems, the population algorithm with the SVM neural network to make better use of the diversity advantages of the population algorithm. Among them, particle swarm optimization (PSO) has the advantages of simple principle, few parameters, and searching for solutions in a large space. Therefore, this paper combines the particle swarm algorithm with the SVM neural network to establish a thermal error model. K-means clustering is used to perform cluster analysis on different temperature measurement points, and the gray correlation degree is used to calculate the correlation between the temperature and the thermal error of the electric spindle, and the temperature-sensitive points are screened out. Finally, particle swarm optimization optimizes the kernel function $g$ of the SVM and the penalty function $c$ builds the predictive model. In order to verify the performance of the PSO-SVM prediction model and compare it with the GA-SVM prediction model, the thermal error model at different speeds was verified on the A02 electric spindle in this experiment. The final results show the thermal error predicted by PSO-SVM. The error model has higher accuracy and better robustness, and provides a more reliable method for modeling the thermal error of the electric spindle; at the same time, this paper provides a new research idea for thermal error modeling of high-speed motorized spindle by combining thermal steady-state analysis, K-means, and gray relation analysis.

In summary, the research idea of this paper is to analyze the thermal characteristics of the motorized spindle first. According to the analysis results, it is judged that the heat transfer characteristics of the spindle are reasonably arranged for temperature measurement points. Then, the clustering results and the gray relation analysis results are combined to select 4 temperature-sensitive points. Finally, the temperature-sensitive points are combined with...
PSO-SVM to establish a thermal error prediction model. The process is shown in Fig. 1.

2 Temperature and thermal error experiment

2.1 Steady-state thermal analysis

The high-speed electric spindle has the characteristics of compact internal structure and high speed. The main heat source is the motor stator, rotor, and spindle bearing. If the motor does not dissipate heat in time under working conditions, its internal temperature will rise rapidly, which will greatly reduce the performance of the bearing and reduce the service life of the electric spindle. So the heat dissipation of the electric spindle is very important.

In this experiment, the A02 electric spindle produced by a CNC Technology Co., Ltd. in Ningbo is used as the research object. Under the simulated working condition of 8000 r/min, ANSYS is used to analyze the thermal steady state. Combined with the literature published by Li et al. [21], the calculation results of the heat transfer coefficient and heat generation rate are shown in Tables 1 and 2. The parameters are set in the ANSYS software, and the steady-state temperature field is obtained (Fig. 2).

Table 1 Convective heat transfer coefficient of electric spindle system

| Parameter name                                             | Convective heat transfer coefficient (w/(m² · k)) |
|------------------------------------------------------------|--------------------------------------------------|
| Convective heat transfer between stator and rotor gap gas   | 172                                              |
| Convective heat transfer between stator and cooling jacket  | 288                                              |
| Convective heat transfer between front bearing and compressed air | 327                                              |
| Convective heat transfer between rear bearing and compressed air | 247.86                                           |
| Convective heat transfer between the rotor end and the surrounding air | 149                                              |
| Convective heat exchange between electric spindle casing and surrounding air | 9.7                                               |

The temperature cloud diagram of the heat transfer of the electric spindle is shown in Fig. 1.

Observing the temperature cloud map, it can be seen that the highest temperature is distributed in the middle gap of the built-in motor, and the highest temperature is about 53.67 °C. The reason is that the motor of the high-speed electric spindle is the largest heat source, no cooling inside, and the interior is closed. Therefore, the temperature at the inner gap is the highest.

From Fig. 1, it can be concluded that the maximum temperatures of the components are about 35.03 °C, 39.82 °C, 26.49 °C, 33.08 °C, 36.83 °C, and 32.41 °C, respectively. The temperature transmitted in the radial direction is faster than that in the axial direction, so the arrangement of the temperature measuring points of the electric spindle should be arranged in the axial direction as much as possible.

2.2 Construction of temperature and thermal error experimental platform

In this paper, the thermal error experiment was carried out on the A02 electric spindle developed by the Harbin University of Science and Technology-Tiankong High-Speed Electric Spindle Joint Laboratory. According to the analysis in the previous section, the heat transfer of the high-speed
electric spindle is mainly in the axial direction. Therefore, in order to collect enough data, at the same time, combined with the literature published by Dai et al. [22], to meet the above requirements, we select 10 temperature measuring points and arrange as many as possible in the axial direction. The temperature of the surface is measured with a PT100 thermal resistance (The measurement range (−40 to approximately +80 °C) measurement accuracy is ±0.5 °C), while the temperature of the front and rear bearings is measured with a PT1000 thermal resistance (The measurement range (−50 to approximately +100 °C) measurement accuracy is ±0.3 °C), thermal displacement adopts LK-H020 displacement sensors with an accuracy of 0.02 μm. Although the electric spindle will generate axial and radial thermal displacement during operation, since the radial displacement is much smaller than the axial displacement, it can be ignored [23]. The layout of the temperature measuring points is shown in Table 3.

| Parameter name | Heat generation rate (w/m³) |
|----------------|-----------------------------|
| Stator heat generation rate | 516,569.11 |
| Rotor heat generation rate | 788,274.28 |
| Front bearing heat generation rate | 328,431 |
| Rear bearing heat generation rate | 355,510 |

According to the layout scheme of temperature measurement points in Table 3, the schematic diagram and on-site installation diagram are shown in Fig. 3.

2.3 Analysis of results

In the case of actual machining, high-speed electric spindles often work for a long time at a constant speed. The object studied in this experiment is A02-type high-speed electric spindle. The experiment is divided into 3 groups according to the rotation speed from 6000 to 10,000 r/min. The time for each set to run is 180 min, and the temperature and thermal displacement of the electric spindle are collected. In order to minimize the influence of indoor temperature on the collected experimental data, the temperature of the laboratory is controlled at 22 °C. Due to the compact structure of the electric spindle itself, it is not easy to dissipate heat, in order to reduce the mutual influence between the experimental data, a certain period of time is required after each set of experiments, and the next set of experiments should be carried out after the electric spindle is completely cooled. According to the above data, the scheme for collecting temperature and thermal displacement data is shown in Fig. 4.

It can be concluded from Fig. 3 that the temperature of each temperature measurement point has experienced a period of temperature rise and then maintained a state of dynamic balance. Since the temperature measuring points T5 and T10 are located at the front and rear bearings, the cooling water does not cool the rear bearings, so the temperatures of T5 and T10 are high, and the temperature of the rear bearing is higher than that of the front bearing. Since the rest of the temperature measuring points are placed on the surface of the electric spindle, the temperature is low. The thermal displacement of the electric spindle increases with the increase of the spindle speed. The temperature and thermal displacement data obtained by the experiment provide data support for the subsequent optimization of temperature measurement points and modeling of thermal errors. From the analysis of thermal characteristics in Sect. 2.1, it can be concluded that the front and rear bearings are one of the largest heat sources of the main shaft, which is verified by experiments to ensure the accuracy of the simulation experiments.
3 Optimization of temperature measurement points

The temperature-sensitive points of the electric spindle should reflect the distribution and temperature change of the temperature field, and also consider the correlation and collinearity between the temperature measuring points. This experiment combines the results of K-means clustering and gray relation analysis degree to find the temperature-sensitive points within the temperature measurement points. This method can eliminate the collinearity problem between the temperature measurement points and improve the accuracy and robustness of the prediction model.

3.1 K-means clustering

The core idea of K-means clustering is to divide the data set into K classes that minimize the objective function. The description of K-means clustering is as follows:

Assume that X = \{X_1, X_2, ..., X_n\} is a collection of n known temperature data. Among them, the K-means clustering method firstly finds the set of K cluster centers, that is, M = \{M_1, M_2, ..., M_k\}, so that the objective function is the Formula (1)

$$J(X, M) = \min \left( \sum_{i=1}^{k} \sum_{j=1}^{n} d(M_i, X_j) \right)$$  (1)

Among them, \(d(M_i, X_j)\) represents the Euclidean distance between the center \(M_i\) of the cluster and the temperature object \(X_j\), and its formula is as Formula (2)

$$d(M_i, X_j) = \sqrt{(m_{i1}-x_{j1})^2 + (m_{i2}-x_{j2})^2 + \ldots + (m_{iN}-x_{jN})^2}$$  (2)

Calculate the Euclidean distance between the data of the remaining temperature points and the cluster centers according to Formula (2), and divide the categories according to the principle of minimum distance, so that a round of clustering is done, and then the average value of each category is calculated. The K averages are used as new K cluster centers, and recalculate the Euclidean distance between the remaining temperature measurement points and the cluster center according to the above calculation method, and the cycle is repeated until the termination conditions are met.

3.2 Gray relation analysis

The correlation between thermal displacement and temperature measurement points is calculated through the gray relation analysis degree, and the points with high correlation degree are selected for thermal error modeling. The degree of similarity can be used to judge whether the connection is close. The specific analysis steps are as follows:

1. Determine the analytical sequence.
Taking the research in this paper as an example, thermal displacement is used as the reference sequence, and temperature is used as the comparison sequence, namely, the reference sequence \( Y = Y(k) | k = 1, 2 \ldots n \); the comparison sequence \( X_i = X_i(k) | k = 1, 2 \ldots n; i = 1, 2 \ldots m \).

2. Non-tempering of variables.

Because the data between each factor column in the system may not be able to draw the correct conclusion or it is difficult to draw the correct conclusion due to the difference of the dimension, so when performing the gray correlation analysis, it is necessary to perform dimensionless quantification processing on the data, such as Formula (3) shows.

\[
x_i(k) = \frac{X_i(k)}{X_i} \quad (3)
\]

In the formula, \( k = 1, 2 \ldots n; i = 0, 1, 2 \ldots m \). \( x_i(k) \) is the normalized data, and \( X_i(k) \) is the original data.

3. Calculate the correlation coefficient, as shown in Eq. (4).
The International Journal of Advanced Manufacturing Technology (2022) 121:7215–7227

\[
ξ_i(k) = \frac{\min_{i,k} \min |y(k) - x_i(k)| + \rho \max_{i,k} |y(k) - x_i(k)|}{|y(k) - x_i(k)| + \rho \max_{i,k} |y(k) - x_i(k)|}
\]

(4)

Note \(\Delta_i(k) = |y(k) - x_i(k)|\), the correlation coefficient is shown in Eq. (5).

\[
ξ_i(k) = \frac{\min_{i,k} \Delta_i(k) + \rho \max_{i,k} \Delta_i(k)}{\Delta_i(k) + \rho \max_{i,k} \Delta_i(k)}
\]

(5)

\(\rho \in (0, \infty)\), \(\rho\) is the resolution factor. \(\rho\) the smaller the resolution, the greater the resolution, usually \(\rho \in (0, 1)\), in general \(\rho=0.5\).

4. Calculate the correlation degree.

The method used to calculate the correlation coefficient is to average the correlation coefficients in each time period and use the average value as the correlation degree, as shown in Formula (6).

\[
r_i = \frac{1}{n} \sum_{k=1}^{n} ξ_i(k), k = 1, 2, ..., n
\]

(6)

The larger the correlation value, the greater the similarity between the column and the reference column.

3.3 Temperature measurement point extraction

The extraction of temperature measurement points is based on the temperature measurement points of 8000 r/min. Since the temperature measurement points have been optimally clustered in Sect. 3.1 of this paper, the clustering results are shown in Table 4.

According to the calculation formula of the gray relation analysis degree, the correlation coefficient between the temperature measuring point corresponding to 8000 r/min and the thermal displacement is calculated, and the result is shown in Fig. 5.

Combine the calculated data in the above figure with the optimal solution obtained by K-means, the temperature measurement points with high gray relation analysis degree are respectively selected from the 4 groups of classification, so the result obtained is the first group 1 chose T9, group 2 chose T5, group 3 chose T1, and group 4 chose T10 as the temperature-sensitive point of this study.

Table 4 Optimal clustering result

| Group number | Measuring point |
|--------------|----------------|
| 1            | T3, T9         |
| 2            | T5             |
| 3            | T1, T2, T4, T6, T7, T8 |
| 4            | T10            |

4. Thermal error modeling

This paper will use SVM, because support vector machine has good robustness and does not need to be fine-tuned. It is relatively simple in calculation, and it is relatively perfect in theory. It is a common method to solve practical problems. However, the randomness of the two core parameters of SVM: The way the penalty factor and kernel function are assigned is random assignment. Since the random assignment may cause the prediction result to fall into the local optimum, the prediction accuracy cannot be guaranteed. Therefore, this paper uses the particle swarm algorithm to optimize the two parameters of the SVM, and establishes the thermal error prediction model of the PSO-SVM motorized spindle and compare this model with GA-AVM and SVM prediction model respectively.

4.1 SVM neural network

Support vector machine is a kind of neural network with good regression prediction performance in dealing with small sample, linear and nonlinear problems. The special algorithm used by the SVM model is to find the optimal plane that can meet the special requirements in the high-dimensional decision space. Through the calculation of the SVM model, a unique and global solution can be obtained, which can effectively prevent the neural network from falling into local extreme values. The specific form of the support vector machine prediction model is as follows:
1. Suppose the training set.
   \[ T = \{(x_1, y_1), \ldots, (x_n, y_n)\} \in (X \times Y)^n \]  
   in the formula \( x_i \in X = R^n, y_i \in Y = \{1, -1\} (i = 1, 2, \ldots, n); x_i \) is the feature vector.

2. Choose a suitable kernel function \( K(x, x') \) and a better parameter C, and construct the formula for solving the optimal problem as follows.
   \[
   \min_{\alpha} \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} y_i y_j \alpha_i \alpha_j K(x_i, x_j) - \sum_{i=1}^{n} \alpha_i \tag{8}
   \]
   \[
   s.t. \sum_{i=1}^{n} y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i = 1, \ldots, n \tag{9}
   \]
   the optimal solution obtained is \( \alpha^* = (\alpha_1^*, \ldots, \alpha_n^*)^T \).

3. Choose \( \alpha^* \) positive component of \( 0 < \alpha_j^* < C \), to calculate the threshold as in Formula (10).
   \[
   b^* = y_i - \sum_{i=1}^{n} y_i \alpha_i^* K(x_i, x_j) \tag{10}
   \]

4. Construct the decision function as Formula (11).
   \[
   f(x) = \text{sgn} \left( \sum_{i=1}^{n} \alpha_i^* y_i K(x_i, x) + b^* \right) \tag{11}
   \]

The SVM architecture is shown in Fig. 6.

4.2 Particle swarm optimization–related parameter settings

PSO as a parallel algorithm simulates the predation of flocks of birds. The group uses the information sharing in the individual to make the whole group arrange the disordered information in an orderly manner in solving the spatial problem, thereby obtaining the optimal solution. Therefore, particle swarm optimization has attracted attention due to its advantages of high precision, fast convergence speed, and easy implementation. The specific process of its algorithm is as follows.

1. Set the population size, random position, and particle velocity.
2. Evaluate the fitness of each individual.
3. Compare the fitness value of each particle with the best position cbest and the best position gbest, respectively; if the comparison result is good, the position of the particle is used as the best position cbest and gbest.

4. If the result is not good, fine-tune the particle position according to Formulas (12) and (13).
   \[
   v_m = v_m + c_1 \times \text{rand()} \times (\text{cbest}_m - x_m) + c_2 \times \text{rand()} \times (\text{gbest}_m - x_m) \tag{12}
   \]
   \[
   x_m = x_m + v_m \tag{13}
   \]
   In the formula: \( m = 1, 2, \ldots, n \), \( n \) is the total number of particles in the population, \( v_m \): particle velocity, \( v_m \) the maximum value of is \( v_{\text{max}} \) and \( v_{\text{max}} > 0 \), if \( v_m > v_{\text{max}} \) then \( v_m = v_{\text{max}} \), \( \text{rand()} \in (0, 1) \) random number between, \( x_m \): the current position of the particle, \( c_1 \) and \( c_2 \): the learning factor, \( c_1 = c_2 = 2 \) in general.

5. Return to step 2 if the convergence condition is not reached.

4.3 PSO-SVM neural network thermal error model

Since the penalty factor and kernel function selected by the SVM neural network are randomly generated during the training process, and the two parameters selected by themselves are likely to have a greater impact on the results, therefore, the particle swarm algorithm is used to optimize the relevant parameters of the SVM, namely penalty factor and kernel function. The neural grid is trained according to these two optimized parameters. To a great extent, over-fitting and under-fitting problems caused by random core parameters are avoided.

The model building steps are as follows:

1. Retrieve the data to normalize the data, and select the training set and test set from it.
2. Initialize particle swarm optimization and related parameters. Set the parameters related to particle swarm optim-
mization, and the cross-validation folds and give the value range of the penalty factor \( c \) and the kernel function \( g \), etc.

3. Determine the fitness function, train the training set with SVM, and use the mean square error MSE as the fitness function. In fact, the fitness function is shown in Formula (14).

\[
MSE = \frac{1}{m} \sum_{i=1}^{m} (y_i - \hat{y}_i)^2
\]  

(14)

In the formula, \( m \) is the number of samples in the training set, and \( y_i \) and \( \hat{y}_i \) are the predicted value and the actual value of the PSO-SVM, respectively.

4. Initialize the population particles, calculate the fitness function, and store the two optimal values \( c_{best} \) and \( g_{best} \).

5. Find out the fitness function values \( c_{best} \) and \( g_{best} \). If the \( c_{best} \) and \( g_{best} \) on the particle information are better than those on the fitness function, the optimal value will update the optimal value.

6. Determine whether the optimal value of the fitness function has reached the previously set precision or the maximum number of iterations. If the conditions are met, go to step 7, otherwise continue to step 5.

7. Generate the optimal solution. Stop the iteration; \( c_{best} \) and \( g_{best} \) are recorded as the optimal penalty factor and kernel function of SVM.

According to the above discussion, the flow chart of PSO-SVM regression prediction is obtained, as shown in Fig. 7.

In order to verify the accuracy and performance of the PSO-SVM prediction model, contrast it with the GA-SVM prediction model.

4.4 Prediction and analysis of thermal error of electric spindle

In this paper, the experimental data of 8000 r/min will be used as the training set. The thermal error prediction model has high accuracy at the same speed, in order to prevent the predicted model from causing problems such as insufficient model accuracy at different speeds. Therefore, in order to test the accuracy and robustness of the prediction model, this paper will use 6000 r/min and 10,000 r/min data as the validation set to verify the model.

The prediction curves of each model under different rotational speeds are shown in Figs. 8 and 9. According to the prediction curve, the prediction accuracy of PSO-SVM and GA-SVM neural network at different speeds is higher than that of the standard SVM, but the accuracy of GA-SVM at different speeds is lower than that of PSO-SVM prediction accuracy. To sum up, it can be concluded that the optimized predicted value of the neural network is closer to the actual measured value, and at the same time, the accuracy of the prediction model of the SVM is greatly increased. According to the prediction curve, when the thermal displacement curve reaches a dynamic equilibrium, the fluctuation of the prediction curve is large, so the prediction error range of SVM, GA-SVM, and PSO-SVM at 6000 r/min is about \(-6.76 \) to \(6.68 \) \(\mu m\), \(-1.03 \) to \(8.40 \) \(\mu m\), and \(-3.26 \) to \(4.48 \) \(\mu m\). The prediction error range at 10,000 r/min is about \(-7.90 \) to \(1.96 \) \(\mu m\), \(-8.75 \) to \(4.29 \) \(\mu m\), and \(-4.04 \) to \(5.62 \) \(\mu m\). In summary, the models predicted by GA-SVM and PSO-SVM are higher in accuracy than those predicted by SVM neural network, but the models predicted by GA-SVM and PSO-SVM neural networks have similar accuracy. Therefore, in order to more accurately compare the two prediction models, the following indicators will be used to evaluate: coefficient of determination \( (R^2) \), root mean square error (RMSE), mean absolute error (MAE), and model accuracy \( (\eta) \). The results of each indicator are shown in Fig. 10.

\( R^2 \) is the coefficient of determination, which is used as an evaluation index for the degree of fit of the model.
The value range of the coefficient of determination is $(-\infty, 1]$. Theoretically, the closer the coefficient of determination is to 1, the higher the fitting degree of the model to the data, while for MAE and RMSE, the value range is $[0, +\infty)$; that is, when the value is closer to 0, the model is better; that is, the larger the prediction error, the larger the value. The closer the model accuracy is to 1, the higher the model accuracy. According to the figure above, the values of the evaluation indicators of the standard neural networks SVM, PSO-SVM, and GA-SVM at 10,000 r/min and 6000 r/min are as follows: the evaluation index MAE at 10,000 r/min is 3.69, 1.68, and 2.97; RMSE is 4.11, 2.08, and 3.69; $R^2$ is 0.99, 0.99, and 0.97; and model accuracy $\eta$ is 0.90, 0.96, and 0.92. The evaluation results of each model at 6000 r/min are $R^2$: 0.96, 0.99, and 0.98; MAE: 3.27, 1.34, and 2.57; RMSE: 3.82, 1.68, and 3.12; and $\eta$: 0.90, 0.96, and 0.92; It can be seen from the above data that although the coefficient of determination of SVM is slightly higher than that of the optimized neural network at a speed of 10,000 r/min, the data of 10,000 r/min fits well in the SVM prediction model, but this does not make the SVM prediction model to have higher accuracy, and the evaluation indicators of PSO-SVM and GA-SVM at different speeds are all better than the standard SVM. According to the comparison of the above data, it can be concluded that each evaluation index at 10,000 r/min is higher than that of GA-SVM and the coefficient of determination is increased by 0.02; MAE and RMSE are decreased by 1.29 and 1.61, respectively, while for PSO, the prediction accuracy of SVM is 96%. At 6000 r/min, the indicators are also higher than GA-SVM and the coefficient of determination is increased by 0.01, MAE and RMSE are decreased by 1.23 and 1.44, respectively, and the prediction model accuracy of PSO-SVM is also 96%.

In conclusion, the model predicted by PSO-SVM neural network is higher than the model predicted by GA-SVM and standard SVM in terms of accuracy and generalization ability. Therefore, the prediction model of PSO-SVM is more suitable for thermal error prediction of high-speed electric spindle.
**Fig. 9** Prediction curve and error of each model at 10,000 r/min. a) SVM prediction curve. b) GA-SVM prediction curve. c) PSO-SVM prediction curve. d) Prediction error of each model.

**Fig. 10** Evaluation index diagram of each model at different speeds. a) 6000 r/min, the evaluation index of each model. b) 10,000 r/min, the evaluation index of each model.
5 Conclusion

In this paper, the A02 high-speed electric spindle is taken as the research object, and the measurement experiment of temperature field and thermal displacement is carried out. According to the experimental data, K-means mean clustering is used to perform cluster analysis on each temperature point, and the temperature measurement points are grouped. SVM established the thermal error prediction model of the electric spindle. The conclusion is as follows:

1. According to the K-means mean clustering method, the 10 temperature measurement points are divided into 4 categories, and they are combined with the gray correlation results to select 4 temperature-sensitive points, which minimizes the collinearity between the temperature measurement point data; the resulting thermal error prediction model is guaranteed in both accuracy and robustness.

2. The thermal error model of the PSO-SVM motorized spindle was established according to the experimental data. In order to verify that the accuracy of the PSO-SVM is better than that of the GA-SVM and SVM prediction models, it was verified at different speeds. The results show that the accuracy values of thermal error models predicted by PSO-SVM, GA-SVM, and SVM are 95.76%, 92.44%, 89.02%, 95.80%, 92.14%, and 89.88% respectively at 6000 r/min and 10,000 r/min. Among them, the model accuracy of PSO-SVM is approximately 96%, and PSO-SVM is easier to apply by reducing parameter settings than GA-SVM, and PSO-SVM shows better prediction accuracy on thermal error prediction models. Therefore, the thermal error prediction model of PSO-SVM is more suitable for thermal error prediction of high-speed motorized spindles.

Author contribution All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Zhaolong Li, Wenming Zhu, Bo Zhu, Qinghai Wang, and Baodong Wang. The first draft of the manuscript was written by Wenming Zhu and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding This research was funded by the National Natural Science Foundation of China (grant number [52075134]).

Data availability Data is contained within the article.

Declarations

Competing interests The authors declare no competing interests.

References

1. Liu K, Han W, Wang YQ, Liu HB, Song L (2021) Review on thermal error compensation for axes of CNC machine tools. J Mech Eng 57:156–173
2. Li Y, Yu ML, Bai YM, Hou ZY, Wu WW (2021) A review of thermal error modeling methods for machine tools. Appl Sci 11(11):5216
3. Wang HT, Li TM, Wang LP, Li FC (2015) Review on thermal error modeling of machine tools. J Mech Eng 51:119–128
4. Dai Y, Tao XS, Li ZL, Zhan SQ, Li Y, Gao YH (2022) A review of key technologies for high-speed motorized spindles of CNC machine tools. Machines 10(2):145
5. Yue HT, Guo CG, Li Q, Zhao LJ, Hao GB (2020) Thermal error modeling of CNC milling machine tool spindle system in load machining: based on optimal specific cutting energy. J Braz Soc Mech Sci Eng 42(9):1–12
6. Yang J, Shi H, Feng B, Zhao L, Ma C, Mei XS (2015) Thermal error modeling and compensation for a high-speed motorized spindle. Int J Adv Manuf Technol 77(5):1005–1017
7. Liu JY, Cai YH, Zhang QL, Zhang HF, He H, Gao XD, Ding LT (2021) Thermal error analysis of taurenEDM machine tool based on FCM fuzzy clustering and RBF neural network. J Intell Fuzzy Syst 1–12
8. Fu GQ, Gong HW, Gao HL, Gu TD, Cao QZ (2019) Integrated thermal error modeling of machine tool spindle using a chicken swarm optimization algorithm-based radial basic function neural network. Int J Adv Manuf Technol 105(5):2039–2055
9. Huang Z, Liu YC, Du L, Yang H (2020) Thermal error analysis, modeling and compensation of five-axis machine tools. J Mech Sci Technol 34(10):4295–4305
10. Cui L, Wang QS (2018) Thermal properties analysis of compact motorized spindle considering fluid-solid thermal coupling. IOP Conf Ser Mater Sci Eng 389(1):012004
11. Chen B, Guan X, Cai DC, Li HL (2022) Simulation on thermal characteristics of high-speed motorized spindle. Case Stud Therm Eng 102144
12. Li YF, Zhang YJ, Zhao QY, Shi XJ (2021) Thermal-mechanical coupling calculation method for deformation error of motorized spindle of machine tool. Eng Fail Anal 128:105597
13. Liu YC, Li KY, Tsai YC (2021) Spindle thermal error prediction based on LSTM deep learning for a CNC machine tool. Appl Sci 11(12):5444
14. Liu PL, Du ZC, Li HM, Deng M, Feng XB, Yang JG (2021) Thermal error modeling based on BiLSTM deep learning for CNC machine tool. Adv Manuf 9(2):235–249
15. Wu CY, Xiang ST, Xiang WS (2021) Spindle thermal error prediction approach based on thermal infrared images: a deep learning method. J Manuf Syst 59:67–80
16. Zhou HM, Wang Z (2021) Cooling prediction of motorized spindle based on multivariate linear regression. J Phys Conf Ser 1820(1):012196
17. Zhu MR, Yang Y, Feng XB, Du ZC, Yang JG (2022) Robust modeling method for thermal error of CNC machine tools based on random forest algorithm. J Intell Manuf 1–14
18. Liu Y, Wang XF, Zhu XG, Zhai Y (2021) Thermal error prediction of motorized spindle for five-axis machining center based on analytical modeling and BP neural network. J Mech Sci Technol 35(1):281–292
19. Li ZL, Zhu B, Dai Y, Zhu WM, Wang QH, Wang BD (2021) Research on thermal error modeling of motorized spindle based on BP neural network optimized by beetle antennae search algorithm. Machines 9(11):286
20. Guo QJ, Fan S, Xu RF, Cheng X, Zhao GY, Yang JG (2017) Spindle thermal error optimization modeling of a five-axis machine tool. Chin J Mech Eng 30(3):746–753
21. Li ZL, Zhu B, Dai Y, Zhu, WM, Wang QH, Wang BD (2022) Thermal error modeling of motorized spindle based on Elman neural network optimized by sparrow search algorithm. Int J Adv Manuf Technol 1–18
22. Dai Y, Yin XM, Wei WQ, Wang G, Zhan SQ (2020) Thermal error modeling of high-speed motorized spindle based on ANFIS. Chin J Sci Instrum 41(6):50–58
23. Yue HT, Guo CG, Li Q, Zhao LJ, Hao GB (2020) Thermal error modeling of CNC milling machining spindle based on an adaptive chaos particle swarm optimization algorithm. J Braz Soc Mech Sci Eng 42:1–13

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.