

Article

Year-Round Thermal Error Modeling and Compensation for the Spindle of Machine Tools Based on Ambient Temperature Intervals

Xinyuan Wei ^{1,†}, Honghan Ye ^{2,†} and Xugang Feng ^{1,*}
¹ School of Electrical and Information Engineering, Anhui University of Technology, Ma'anshan 243032, China; weixy@ahut.edu.cn

² Department of Industrial and Systems Engineering, University of Wisconsin—Madison, Madison, WI 53705, USA; hye42@wisc.edu

* Correspondence: fxg@ahut.edu.cn; Tel.: +86-13965390996

† These authors contributed equally to this work.

Abstract: The modeling and compensation method is a common method for reducing the influence of thermal error on the accuracy of machine tools. The prediction accuracy and robustness of the thermal error model are two key performance measures for evaluating the compensation effect. However, it is difficult to maintain the prediction accuracy and robustness at the desired level when the ambient temperature exhibits strong seasonal variations. Therefore, a year-round thermal error modeling and compensation method for the spindle of machine tools based on ambient temperature intervals (ATIs) is proposed in this paper. First, the ATIs applicable to the thermal error prediction models (TEPMs) under different ambient temperatures are investigated, where the C-Means clustering algorithm is utilized to determine ATIs. Furthermore, the prediction effect of different numbers of ATIs is analyzed to obtain the optimal number of ATIs. Then, the TEPMs corresponding to different ATIs in the annual ambient temperature range are established. Finally, the established TEPMs of ATIs are used to predict the experimental data of the entire year, and the prediction accuracy and robustness of the proposed ATI model are analyzed and compared with those of the low and high ambient temperature models. The prediction accuracies of the ATI model are 20.6% and 41.7% higher than those of the low and high ambient temperature models, respectively, and the robustness is improved by 48.8% and 62.0%, respectively. This indicates that the proposed ATI method can achieve high prediction accuracy and robustness regardless of the seasonal temperature variations throughout the year.

Keywords: CNC machine tool; thermal error; ambient temperature interval; model robustness



Citation: Wei, X.; Ye, H.; Feng, X. Year-Round Thermal Error Modeling and Compensation for the Spindle of Machine Tools Based on Ambient Temperature Intervals. *Sensors* **2022**, *22*, 5085. <https://doi.org/10.3390/s22145085>

Academic Editors: Fang Cheng, Qian Wang, Tegoeh Tjahjowidodo and Ziran Chen

Received: 11 June 2022

Accepted: 5 July 2022

Published: 6 July 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

In the machining process, changes in the internal and external heat sources such as motor operation, friction, cutting heating, and environmental temperature result in thermal errors, which represent 40–70% of the total errors of machine tools [1,2]. With the development of high-precision CNC machine tools, the influence of thermal errors on the tool performance is gradually becoming dominant [3]. To reduce such influence on machine tools, there are two main approaches. The first approach is to establish the analytical model and then to simulate and analyze thermal error laws. Although numerical analysis is promising to compensate for thermal errors, it is extremely difficult for the numerical method to build an exact structural model in practice and simulate the thermal deformation of machine tools because of complicated deformation processes. Alternatively, software compensation methods are commonly used to reduce the influence of thermal errors on the accuracy of machine tools [4]. In this method, temperature sensors are first installed on various locations of a machine, and a thermal error prediction model is established based on the temperature information collected from those sensors. Then the established model is embedded into the CNC system to realize thermal error compensation in real-time.

Such a method, which is generally established by selecting relevant temperature-related variables and designing an appropriate modeling algorithm, offers high prediction accuracy and robustness [5].

First, the selected temperature variables are commonly called temperature-sensitive points (TSPs) in the literature. Currently, the most common TSP-selection algorithm is fuzzy clustering combined with the grey correlation degree algorithm [6–9]. The algorithm mainly involves classifying temperature variables and selecting one variable from each class as a TSP that has the greatest correlation with the thermal error. Based on the concept of this algorithm, researchers have proposed the rough set theory combined with the grey correlation degree algorithm [10] and fuzzy clustering combined with the correlation coefficient algorithm [11] to select the TSPs. Second, the algorithms commonly used to establish the thermal error prediction model (TEPM) include the multiple linear regression algorithm [6,12,13], time series algorithm [14,15], neural network algorithm [7,8,10,13], and support vector machine algorithm [16,17].

Recently, researchers have further studied the thermal error modeling algorithm of CNC machine tools [18–22] to improve the accuracy and robustness of thermal error prediction. However, the thermal error data in these studies had very small variations of ambient temperature, and the influence of ambient temperature was rarely considered. Unfortunately, their models perform poorly when the ambient temperature changes significantly, where the ambient temperature has a non-negligible influence on the thermal error modeling and compensation [1,12]. In fact, a significant change in the environmental temperature would lead to changes in the law of thermal deformation of the key parts of a machine tool. To verify this observation, we conducted the analysis of variance (ANOVA) of different sets of spindle thermal error data throughout a year. The results indicated a significant difference between the spindle thermal error laws under different ambient temperatures, i.e., the spindle thermal error law of a machine tool changes significantly with the ambient temperature. Thus, the constructed TEPMs from previous studies are difficult to use in accurately predicting spindle thermal errors over a wide range of ambient temperatures. In this regard, Zhang et al. [23] and Li et al. [24] used the finite element method to study thermal deformation laws of machine tools under the condition of time-varying environmental temperature. This method can be used to obtain the accurate thermal deformation of a machine tool under different ambient temperatures. However, it is a complex method, and the established model is not universal. More recently, Liu et al. [25] constructed thermal error models on both sides of a segment point of the ambient temperature. However, they only considered one segment point in their model and did not investigate the optimal number of segment points. Therefore, there exists a research gap, and the existing literature lacks a methodology that systematically analyzes the influence of ambient temperatures on the predictions and provides high accuracy and robustness performances.

Different from the existing methods [9,18,19] that have not considered the ambient temperatures in their models, the main aim of the research in this paper is to systematically investigate the influences of different ambient temperatures on thermal error prediction. The objective of this paper is to provide a thermal error modeling method, which takes ambient temperatures into consideration, so that the prediction accuracy and robustness can be further improved. To fill the aforementioned research gap, in this paper, we propose a year-round thermal error modeling and compensation method for the spindle of machine tools based on ambient temperature intervals (ATIs). Specifically, first, the ATIs applicable to the TEPMs are investigated based on experimental data collected over an entire year. Second, the C-means clustering algorithm is utilized to determine the ATIs. The influence of the number of ATIs on the prediction effects is further analyzed to determine the optimal number of ATIs. Then, the TEPMs are established with the optimal number of ATIs in the annual ambient temperature range. Finally, the prediction effects of the established TEPMs based on ATIs are analyzed and compared with those of the low and high ambient temperature models using the annual experimental data. As a result, the proposed method has two main contributions: (1) Different from the simple segmentation in [25], the C-means

clustering algorithm is embedded and generic to find the number of ATIs; (2) the proposed method based on the selected ATIs performs better than the models that do not consider ambient temperatures in terms of both prediction accuracy (over 20% improvements) and robustness (over 48% improvements). The proposed method provides an important reference for thermal error modeling and compensation for CNC machine tools in a large ambient temperature range, and it is of great significance for promoting the development and precision of intelligent manufacturing.

The remainder of this paper is organized as follows. Section 2 introduces the modeling algorithms used in this study. In Section 3, an analysis of the experimental data collected over one year is presented, including the significance test of thermal error difference and the mutual prediction effect analysis of the annual experimental data. Section 4 first introduces the overview of the proposed modeling and compensation method based on the ATIs, and then it provides the details of determining ATIs based on the C-Means clustering method and selecting the optimal number of ATIs; it finally analyzes the prediction effects of the proposed ATI method. Section 5 draws the conclusions along with their applications and discusses the future scope of the study.

2. Thermal Error Modeling Algorithms

The thermal error modeling theory generally includes TEPM establishment and TSP selection algorithms. Principal component regression (PCR) [26], ridge regression [20], and partial least squares [27,28] are the common partial regression algorithms that can effectively suppress the influence of collinearity between input variables. Based on our preliminary studies, we used the PCR algorithm to construct the TEPM and the correlation coefficient algorithm to select the TSPs, which are introduced in Sections 2.1 and 2.2, respectively.

2.1. Thermal Error Prediction Model (TEPM) Establishment Algorithm

Principal component analysis transforms a group of correlated variables into a group of linearly uncorrelated variables using orthogonal transformation, where the transformed variables are called principal components. Then, the principal components are further used as independent variables to establish the regression model.

In this study, temperature data were assumed as $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2 \dots \mathbf{x}_p)$, where p represents the number of temperature variables. The specific steps for establishing the thermal error PCR model are as follows.

1. Normalize the original data $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2 \dots \mathbf{x}_p)$ to obtain $\mathbf{X}^* = (\mathbf{x}_1^*, \mathbf{x}_2^* \dots \mathbf{x}_p^*)$. The standardized formula for this is as follows:

$$\mathbf{x}_i^* = \frac{\mathbf{x}_i - E(\mathbf{x}_i)}{\sqrt{Var(\mathbf{x}_i)}} (i = 1, 2, \dots, p), \quad (1)$$

where original data \mathbf{x}_i is the temperature information on variable i collected from sensor i , and those sensors are installed in various locations of the machine tool, which will be shown later in Figure 1 and Table 1. In addition, $E(\mathbf{x}_i)$ and $Var(\mathbf{x}_i)$ are the expected value (average) and the variance of the \mathbf{x}_i , respectively. Principal components Z_i can be obtained from the eigenvalues and corresponding eigenvectors of a correlation coefficient matrix P of the original data $\mathbf{X} = (\mathbf{x}_1, \mathbf{x}_2 \dots \mathbf{x}_p)$:

$$\begin{cases} Z_1 = e_{11} * \mathbf{x}_1^* + e_{21} * \mathbf{x}_2^* + \dots + e_{p1} * \mathbf{x}_p^* \\ Z_2 = e_{12} * \mathbf{x}_1^* + e_{22} * \mathbf{x}_2^* + \dots + e_{p2} * \mathbf{x}_p^* \\ \dots \\ Z_p = e_{1p} * \mathbf{x}_1^* + e_{2p} * \mathbf{x}_2^* + \dots + e_{pp} * \mathbf{x}_p^* \end{cases}, \quad (2)$$

where $e_i = (e_{i1}, e_{i2}, \dots, e_{ip})^T$ is the eigenvector of the correlation coefficient matrix P , and each entry ($P_{i,j}$) of the matrix P can be calculated as follows:

$$P_{i,j} = \frac{Cov(\mathbf{x}_i, \mathbf{x}_j)}{\sqrt{Var(\mathbf{x}_i)Var(\mathbf{x}_j)}},$$

where $i = 1, 2, \dots, p$, $j = 1, 2, \dots, p$, and $Cov(\mathbf{x}_i, \mathbf{x}_j)$ is the covariance between \mathbf{x}_i and \mathbf{x}_j . Note that all elements of the main diagonal of P (namely when $i = j$) have the value of one since the covariance of \mathbf{x}_i with itself is $Var(\mathbf{x}_i)$.

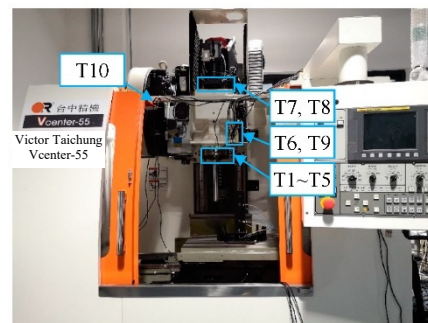


Figure 1. Experimental object.

Table 1. Installation sites and measurement functions of temperature sensors.

Sensors	Installation Site	Function
T1~T5	Front bearing of Spindle	Bearing temperature measurement
T7,T8	Spindle motor	Spindle motor temperature measurement
T6,T9	Spindle box	Spindle box temperature measurement
T10	Machine frame	Ambient temperature measurement

- Calculate the cumulative contribution rate V_g of the i th principal component; the formula is as follows:

$$V_g = \sum_{i=1}^g \lambda_i / \sum_{i=1}^p \lambda_i (g = 1, 2, \dots, p), \quad (3)$$

where λ_i is the eigenvalue of matrix P corresponding to eigenvector e_i . In this study, the principal components with a V_g of more than 95% were selected as TSPs to avoid a loss of information.

- Normalize the original data of thermal error y using Equation (1) to obtain y^* and then establish the regression model between y^* and the selected principal components:

$$\hat{y}^* = \hat{a}_1 Z_1 + \hat{a}_2 Z_2 + \dots + \hat{a}_g Z_g, \quad (4)$$

where $\hat{a}_1, \dots, \hat{a}_g$ are the coefficients of the regression model. Specifically, vector $\hat{a} = (\hat{a}_1, \dots, \hat{a}_g)^T$ can be obtained by calculating $Z^T Z^{-1} Z^T y^*$, where $Z = (Z_1, \dots, Z_g)$, since both Z and y^* are known. Please note that the intercept is not included here since all data are centered with mean 0 by normalization.

- Substitute the formula of each principal component in Equation (2) into regression Equation (4) to obtain the regression equation between standardized variables $\mathbf{x}_1^*, \mathbf{x}_2^*, \dots, \mathbf{x}_p^*$ and dependent variable \mathbf{y}^* :

$$\hat{y}^* = \hat{\beta}_1 \mathbf{x}_1^* + \hat{\beta}_2 \mathbf{x}_2^* + \dots + \hat{\beta}_p \mathbf{x}_p^*, \quad (5)$$

where $\hat{\beta}_i = (\hat{a}_1, \dots, \hat{a}_g)(e_{i1}, \dots, e_{ig})^T$ for $i = 1, 2, \dots, p$. The regression model between the original variables can be obtained from the relationship between estimated parameter $\hat{\beta}_i$ in Equation (5) and the original data regression model estimated parameter \hat{b}_i :

$$\hat{y} = \hat{b}_0 + \hat{b}_1 x_1 + \hat{b}_2 x_2 + \dots + \hat{b}_p x_p, \quad (6)$$

where the relationship between $\hat{\beta}_i$ and \hat{b}_i is as follows.

$$\begin{cases} \hat{b}_i = \hat{\beta}_i * \frac{\sqrt{Var(y)}}{\sqrt{Var(x_i)}} \\ \hat{b}_0 = E(y) - \sum_{i=1}^p \hat{b}_i * E(x_i) \end{cases} \quad (7)$$

2.2. TSP Selection Algorithm

As the PCR algorithm can suppress the influence of collinearity between input variables, the temperature variables did not have to be classified to reduce the correlation among the TSPs. We used Pearson's correlation coefficient to determine the relationship between the temperature variable and thermal error. The temperature variables with high correlation coefficients were selected as the TSPs. The equation of the Pearson correlation coefficient is as follows:

$$\rho_{x_i y} = \frac{Cov(x_i, y)}{\sqrt{Var(x_i)Var(y)}}, \quad (8)$$

where $\rho_{x_i y}$ is the correlation coefficient between the i th temperature variable x_i and the thermal error y . $Cov(x_i, y)$ is the covariance between x_i and y . $Var(x_i)$ and $Var(y)$ are the variances of x_i and y , respectively. The equations are as follows:

$$Cov(x_i, y) = \frac{\sum_{k=1}^n (x_{i,k} - E(x_i))(y_k - E(y))}{n-1}, \quad (9)$$

$$Var(x_i) = \frac{\sum_{k=1}^n (x_{i,k} - E(x_i))^2}{n-1}, \quad (10)$$

where n is the number of observations/samples in the data.

3. Experimental Setup and Data Analysis

Thermal error measurement experiments were conducted on a three-axis vertical machining center for one year. In this section, an analysis of variance was performed on each batch of the experimental data. Then, the correlation coefficient and PCR algorithm were used to establish the TEPM. The prediction effects of the established models were further analyzed to prove that it was difficult to maintain their robustness over a large range of ambient temperatures.

3.1. Thermal Error Measurement Experiments

The annual thermal error experimental object was a Vcenter-55 three-axis vertical machining center (Figure 1). The five-points measurement method was used for the spindle thermal error measurement according to the International Standard "Test code for machine tools—Part 3: Determination of thermal effects" (ISO 230-3:2020 IDT) [29]. Two displacement sensors each were used for the X and Y directions, and another one in the Z direction (Figure 2).

High-precision capacitance sensors were used to measure the spindle thermal error and their measurement accuracy was 1 μm . Ten Pt resistance temperature sensors were used to measure the temperature change at different positions of the machine tool. These temperature sensors are used to measure the main heat sources of the machine tool, such as spindle motor, spindle bearing, and so on. Ambient temperature is also measured by the sensor located at the machine tool frame. The distribution of the 10 temperature sensors is shown in Figure 1. The measurement accuracy of the temperature sensors was 0.1 $^{\circ}\text{C}$, and their installation sites and measurement functions are listed in Table 1.

According to the International Standard [29], the spindle of the machine tool should idle at constant speed, and the worktable should reciprocate at a constant feed speed along the X- and Y-axes. Temperature and thermal error data were collected every 5 min for

more than 4 h. In the annual experiment, the spindle speeds were set to 2000, 4000, and 6000 rpm, and the feed speed to 1500 mm/min. Each batch of experiments was performed after cooling the machine tool for 2 days.

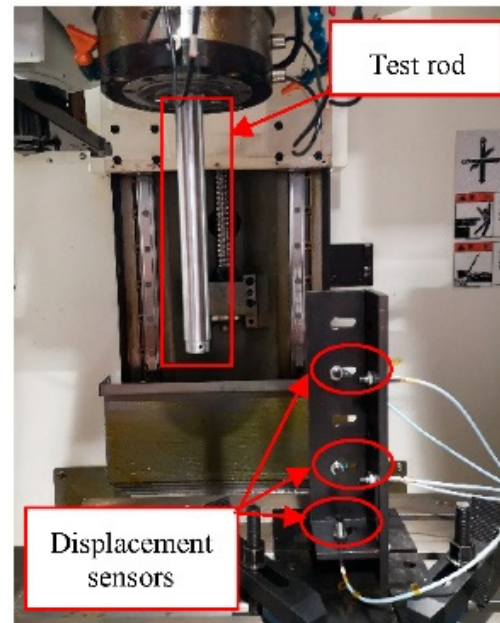


Figure 2. Distribution of displacement sensors.

3.2. Experimental Data Analysis

A total of 46 batches of effective experimental data were collected over an entire year and were sorted according to the initial ambient temperature recorded as K1–K46. Figure 3 graphically plots the temperature measurement results of the K1 experiment, which was conducted in winter, with the lowest initial ambient temperature and K46, which was conducted in summer, with the highest initial ambient temperature. As observed, the temperature change curves of K1 and K46 have significant differences.

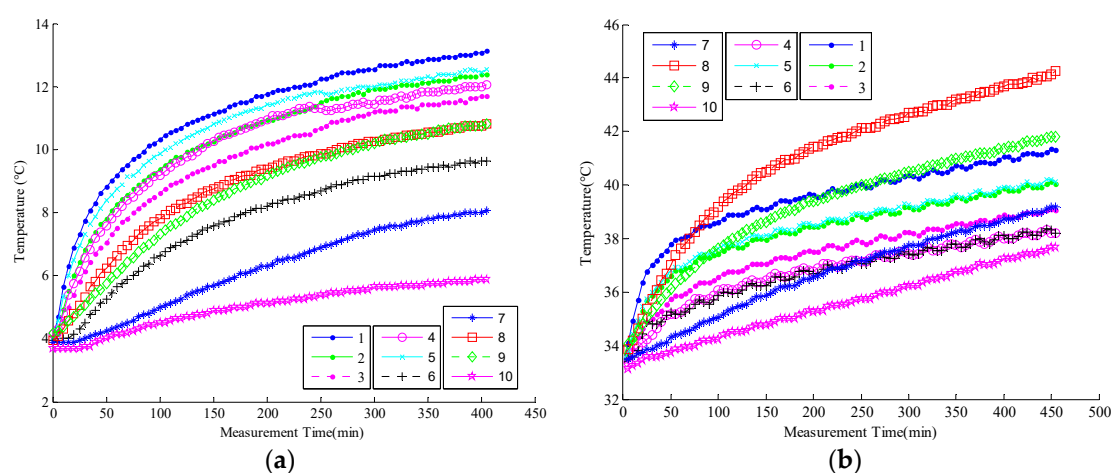


Figure 3. (a) Temperature change curves of experimental batches K1; (b) temperature change curves of experimental batches K46.

The thermal error curves of the 46 batches can be drawn similarly. Only the thermal error curves of batches K1–K5 and K42–K46 could be drawn, due to the large volume of data (Figure 4). Figure 4 reveals significant differences in the thermal error law between the low and high ambient temperature experiments. The thermal error in the low temperature

environment was larger because its gradient began to increase by a greater margin after the machine tool had run for a certain time.

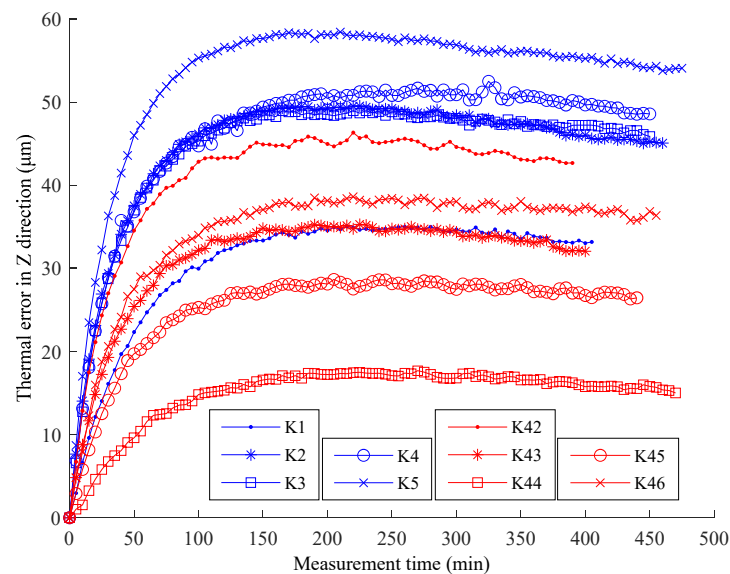


Figure 4. Thermal error curves of ten batches of experiments.

According to the exploratory analysis of the experimental data, the initial ambient temperature range was 4.1–32.2 °C, and the range of temperature increase in the machine tool was 5.9–16.8 °C. Figure 5 illustrates the temperature results of each batch of experimental data.

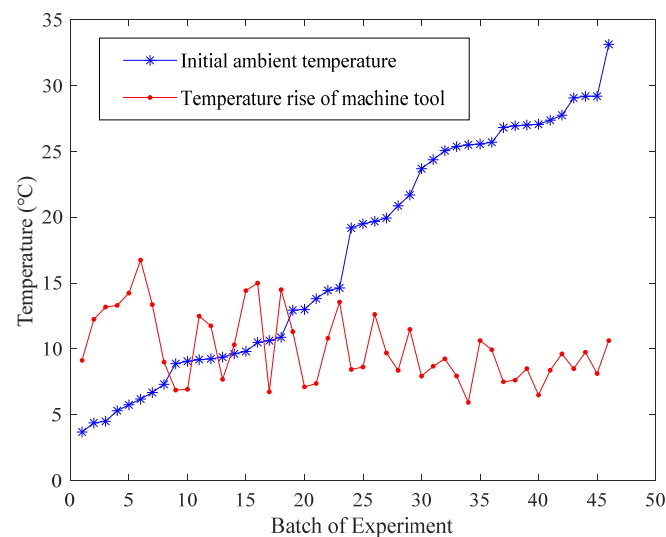


Figure 5. Statistical results of temperature changes of 46 batches of experiments.

3.3. Significance Analysis of Thermal Error Data

ANOVA [30] is mainly used to test the significance of the difference between the mean values of two or more samples. According to the number of influencing factors, it includes one-way, two-way, and multi-factors ANOVA. We used one-way ANOVA to test whether the ambient temperature had a significant influence on the thermal error variation law of the machine tool.

If we assume that there are r batches of thermal error experimental data under different ambient temperatures and n thermal error measurement data in each batch

of the experiment, then all the thermal error experimental data would be recorded as y_{ij} ($i = 1, 2, \dots, r; j = 1, 2, \dots, n$). The thermal error average value of each batch was $\bar{y}_i = \sum_{j=1}^n y_{ij}$, and the thermal error average value of all batches was $\bar{y} = \sum_{i=1}^r \bar{y}_i$. The basic steps of one-way ANOVA are as follows.

1. Calculation of test statistics. The calculation formulas of the sum of squares of the total deviation S_T , the sum of squares of the intra-group deviation S_E , and the sum of squares of the inter-group deviations S_A , respectively, are shown below.

$$\begin{cases} S_T = \sum_{i=1}^r \sum_{j=1}^n (y_{ij} - \bar{y})^2 \\ S_E = \sum_{i=1}^r \sum_{j=1}^n (y_{ij} - \bar{y}_i)^2 \\ S_A = n \sum_{i=1}^r (\bar{y}_i - \bar{y})^2 \end{cases} \quad (11)$$

The degrees of freedom (DoF) of the three statistical parameters are $df_T = n * r - 1$, $df_E = n * r - r$, and $df_A = r - 1$.

2. *F*-test statistics. The statistics in step (1) were divided by the corresponding DoF to calculate the mean square of the sum of squares to eliminate the interference caused by different DoF. This will be conducive to the comparison of the sum of squares of the deviations of each group of data. Then, the ratio *F* of the average of the squares of the inter-group deviations S_A to the average of the squares of intra-group deviations S_E was used to test the original hypothesis.

$$F = \frac{S_A/df_A}{S_E/df_E} \quad (12)$$

3. Determine the significance. Test statistic *F* follows the *F* distribution with df_A and df_E . The larger the value of *F*, the more inclined it is to reject the original hypothesis. Therefore, the rejection domain of the test is as follows:

$$W = \{F \geq F_{1-\alpha}(df_A, df_E)\} \quad (13)$$

where α is the significance level, and $F_{1-\alpha}(df_A, df_E)$ is calculated through the *F* distribution table. If $F \geq F_{1-\alpha}(df_A, df_E)$, the thermal error data of each batch have significant differences. Otherwise, the difference is not significant.

The *p* value of the test is obtained by the density function of the *F* distribution.

$$p = P(Y \geq F) \quad (14)$$

In general, if $p \leq 0.001$, each batch of thermal error data is considered to be significantly different; if $p \geq 0.05$, the difference is not significant.

One-way ANOVA for the thermal error data in the *Z* direction of each batch was performed following the aforementioned steps, and the results are summarized in Table 2. In Table 2, S_s is the sum of squares of deviations, df is the DoF, M_s is the mean square, *F* is the statistic to be tested, *p*-value is the density function *p* of the statistic *F* to be tested in the *F* distribution, and F_{crit} is the critical value of the *F* statistic when the confidence level $\alpha = 0.001$.

Table 2. ANOVA and significance test results.

Source of Difference	S_s	df	M_s	<i>F</i>	<i>p</i> -Value	F_{crit}
Type	388,844	46	8640.978	113.5396	Close to 0	1.790356
Error	231,055.9	3036	76.10538			
Total	619,900	3081				

According to the analysis results in Table 2, the value of F is significantly larger than the critical value F_{crit} of 1.79, with a DoF of (46, 3036), at the confidence level of $p = 0.001$. Meanwhile, the p -value corresponding to F is close to 0, which is markedly less than 0.001, indicating that the ambient temperature has a very significant impact on the thermal error data.

3.4. Analysis of Mutual Prediction Results with Annual Experimental Data

Each of the 46 batches of experiments was used to establish a TEPMs and predict the remaining batches of experiments. Thus, the mutual prediction results of 46 batches of experiments can be obtained. First, TSPs were selected for each batch of experiments with the use of the TSP selection algorithm presented in Section 2.2. Two TSPs are selected for each batch of experiments in this study with reference to [18]. The experimental data of K1 batch are taken as an example to illustrate TMP selection. The correlation coefficient between the thermal error and each temperature variable is calculated by Equation (8), and the results are shown in Table 3.

Table 3. Correlation coefficient between temperature variable and thermal error of K1.

Sensors	T_1	T_2	T_3	T_4	T_5	T_6	T_7	T_8	T_9	T_{10}
Correlation coefficient	0.92	0.92	0.91	0.92	0.92	0.85	0.77	0.86	0.86	0.70

It can be observed from Table 3 that the correlation coefficients of T_1 and T_5 are the largest, both of which are 0.92. Thus, T_1 and T_5 are selected as the TSPs, which are used to measure the heating of the spindle motor. In the same manner, the TSPs of each batch of experiments are selected, as listed in Table 4.

Table 4. Selection results of TSPs for each batch.

Batch	K1	K2	K3	K4	K5	K6	K7	K8	K9	K10
Results	T1, T5	T1, T5	T1, T3	T1, T5	T1, T5	T1, T4	T1, T2	T1, T4	T1, T5	T1, T5
Batch	...	K38	K39	K40	K41	K42	K43	K44	K45	K46
Results	...	T1, T8	T1, T5	T1, T8	T1, T5	T1, T5	T1, T5	T8, T9	T1, T5	T1, T5

Based on the TSPs selection results, the TEPM of each batch of experiments was established using the PCR algorithm in Section 2.1. For example, the TEPMs of K1 and K46 are as follows.

$$Z_1 = 1.8767 + 1.8147T_5 + 4.9211T_1 \quad (15)$$

$$Z_{46} = 2.1280 + 2.3108T_1 + 9.6814T_5 \quad (16)$$

Furthermore, the prediction results of each TEPM were calculated. The predicted residual standard deviation (Rsd) was used to represent prediction accuracy. The formula of the predicted Rsd is as follows:

$$Rsd = \sqrt{\frac{\sum_{j=1}^n (\hat{y}_j - y_j)^2}{n-1}}, \quad (17)$$

where y_j is the j th measured thermal error value, and \hat{y}_j is the j th predicted thermal error value.

Finally, the mutual prediction results of the 46 batches of experiments were calculated and are presented in Figure 6, depicting the differences.

The mutual prediction accuracy is high when the ambient temperature is within a certain range, as shown in Figure 6. However, when the ambient temperature exceeded a certain range, prediction accuracy decreased; i.e., the robustness decreased. In addition, the model based on experimental data under low ambient temperature has a better prediction effect than the one based on the high-temperature data.

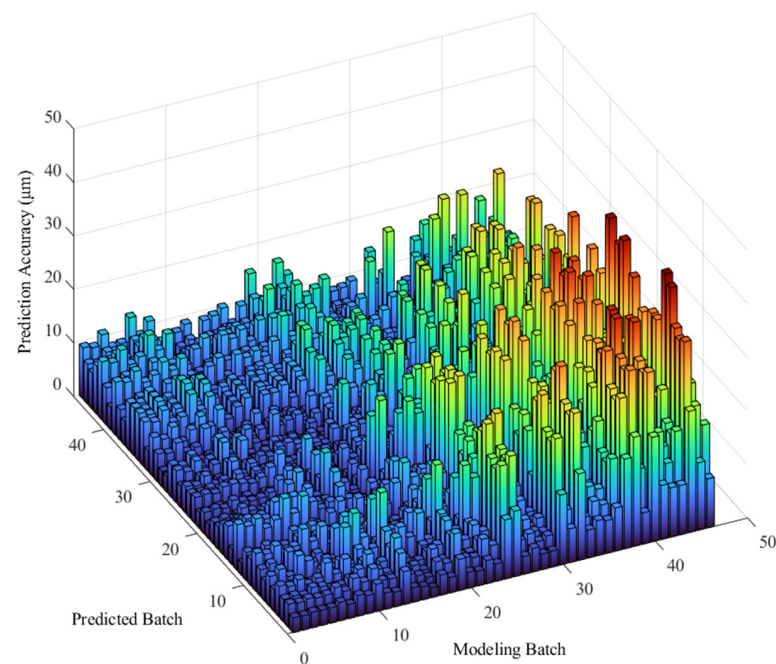


Figure 6. Mutual prediction results of 46 batches of experiments.

When using a TEPM to predict the thermal error data under different ambient temperatures, the method can only maintain high prediction accuracy for data within a certain ambient temperature range. We call this temperature range the “applicable ATI of TEPM”. For each TEPM, experiment batches with an Rsd of less than $5\ \mu\text{m}$ were grouped according to the mutual prediction results presented in Figure 6. Then, the ambient temperature range of these experiment batches was obtained as the applicable ATI of the prediction model. Thus, the applicable ATI of the TEPM of experimental data from each batch was obtained (Figure 7). The horizontal axis in the figure is the batch of the modeling data. The vertical axis is the ambient temperature range. The blue dotted lines are the upper and lower boundaries of all applicable ATIs. The upper and lower boundaries of the ATIs were fitted as the red lines.

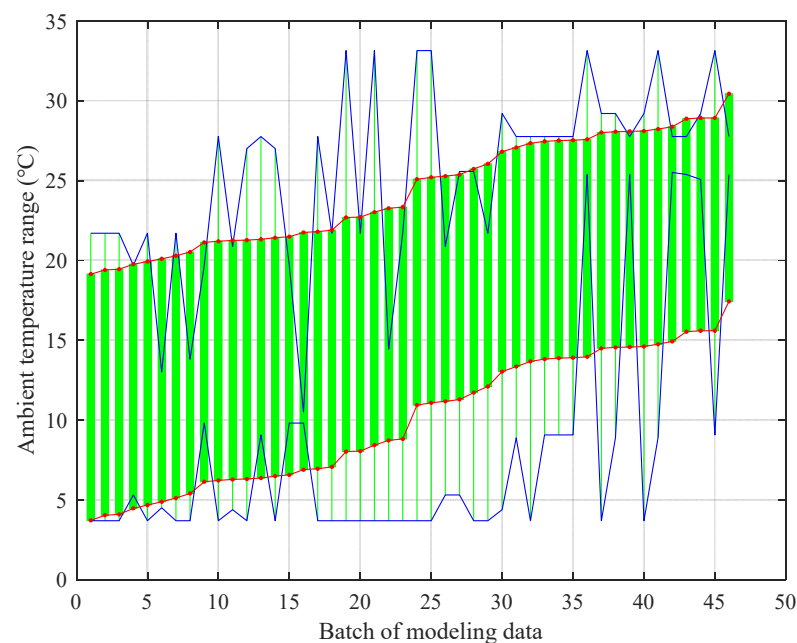


Figure 7. Applicable ATIs of TEPMs of 46 batches of experiments.

The applicable ATIs fluctuated by a margin due to some random factors, such as measurement error. It is obvious that the established TEPMs have different applicable ATIs. This proves the necessity of the proposed modeling method based on ATIs for compensation through an entire year.

4. Thermal Error Modeling and Compensation Method Based on ATIs

To solve the problem of decreased thermal error prediction effects over the annual ambient temperature range, we proposed a year-round thermal error modeling and compensation method based on ATIs. In this section, an overview of the proposed modeling and compensation method is first introduced in Section 4.1. In Section 4.2, the ATIs are determined using the C-Means clustering algorithm and the optimal number of ATIs is obtained based on the prediction results of TEPMs corresponding to the number of ATIs. Finally, the prediction effects of the established models are analyzed based on the experimental data collected over one year in Section 4.3.

4.1. Overview of the Proposed Method Based on ATIs

As the TEPM established by the experimental data under a fixed ambient temperature can only maintain high prediction accuracy in its applicable ATI, it is necessary to establish the corresponding TEPMs for different ambient temperature ranges, namely the modeling and compensation method based on ATIs.

Figure 8 presents the flowchart of the proposed method that contains two processes. First, in the modeling process, ATIs are divided and the ambient temperature corresponding to each ATI is determined based on the C-Means clustering algorithm. Subsequently, TEPMs are constructed for each determined ATI. In the compensation process, ATI would be determined according to the initial ambient temperature value. Then, the corresponding TEPM is invoked to predict the thermal errors. Finally, thermal error compensation would be performed according to the predicted value. Thus, the high prediction accuracy and robustness can be realized for the thermal error within the annual ambient temperature range.

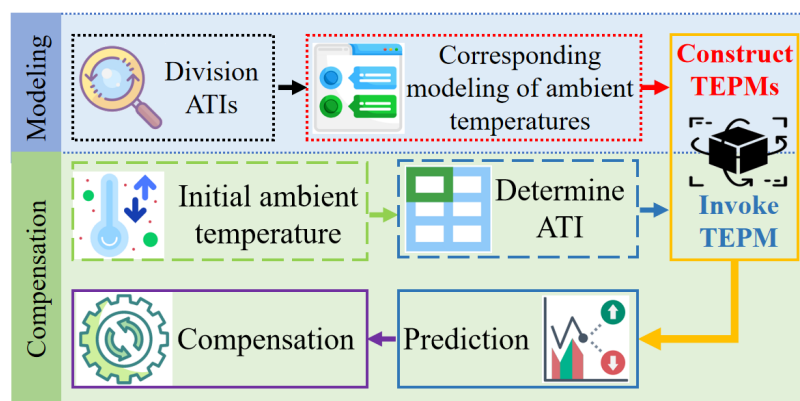


Figure 8. Flowchart of the modeling and compensation method based on ATIs.

4.2. Determination of the ATIs

The key to the success of the proposed modeling and compensation method is determining ATIs reasonably. To achieve that, the C-Means clustering algorithm is utilized to determine the ATIs. As the number of ATIs is not unclear, the prediction effects are analyzed to obtain the optimal number of ATIs. In this subsection, the C-Means clustering algorithm is first introduced and then the optimal number of ATIs is analyzed and obtained.

4.2.1. C-Means Clustering Algorithm

C-Means clustering is a typical algorithm for dynamic clustering. The ambient temperature variables of 46 batches of experiments are clustered using the C-Means clustering algorithm. According to the clustering results, the initial ambient temperature variation

range of the ambient temperature variables of the same class is considered as one ATI. The basic steps of dividing the ATIs based on the C-means clustering algorithm are as follows.

1. Given the number of clusters M and randomly selecting M initial cluster centers C_1, C_2, \dots, C_M
2. Calculate the distance from each ambient temperature variable to each cluster center, and the calculation formula is as shown:

$$d(x, C) = \sum_{k=1}^n |x_k - C_k|, \quad (18)$$

where x represents the ambient temperature variable, and n represents the number of sample data in temperature variable.

3. Divide the ambient temperature variables closest to the cluster center into this cluster;
4. Recalculate cluster center according to the clustering results, that is, take the average value of the ambient temperature variables belonging to the same cluster as the new cluster center;
5. Repeat steps 2 to 4 until the cluster results remains unchanged or the maximum number of iterations is reached;
6. Consider the initial ambient temperature range contained in the same class as one ATI.

For example, suppose the number of clusters is 4; then, the 46 ambient temperature variables of the 46 batches of experiments are grouped into four classes, as shown in Table 5. The demarcation of the two adjacent ATIs is calculated by averaging the boundaries of the two adjacent initial ambient temperature ranges. For example, the demarcation of the first and second ATI is $(7.0 + 9.1)/2 \approx 8.1$.

Table 5. Clustering results with four clusters.

Class	Results	Initial Ambient Temperature Range	ATI
Class 1	1–8	4.1–7.0	[4.1–8.1]
Class 2	9–23	9.1–14.5	(8.1–17.2]
Class 3	24–30	19.8–23.4	(17.2–24.0]
Class 4	31–46	24.6–32.2	(24.0–32.2]

The TEPM established by the experimental data at the ambient temperature at the left end of the ATI is used as the TEPM of the ATI. In this way, the TEPMs corresponding to the four ATIs in Table 5 can be obtained, as shown below.

$$\begin{cases} Y_1 = 1.8767 + 1.8147T_5 + 4.9211T_1 & 4.1 \leq T_0 \leq 8.1 \\ Y_9 = 2.2417 + 2.2222T_1 + 3.4832T_5 & 8.1 < T_0 \leq 17.2 \\ Y_{24} = 1.8492 + 1.8419T_1 + 7.2712T_5 & 17.2 < T_0 \leq 24.0 \\ Y_{31} = 2.6280 + 2.7416T_1 + 0.2192T_5 & 24.0 < T_0 \leq 32.2 \end{cases} \quad (19)$$

The coefficients in the regression models, as shown in Equation (19), are all positive, which show that when the temperature rises, the thermal error increases as well. This is consistent with the physical law in general. It is also noteworthy that all 46 batches of data (only temperature information) are used as initial batches for determining the ATIs. In practice, a general rule to reduce the number of initial batches is that it is suitable to reduce the batches when the batches of experiments are conducted within a similar ambient temperature interval, but reducing them is not suitable when they are conducted with significant different ambient temperatures. As a result, the number of initial batches for determining the ATIs can be reduced while the information contained in the data is not lost. In addition, the number of initial batches (both temperature and thermal error information) for modeling is the same as the number of ATIs, and the effects of different numbers of ATIs on the performance will be investigated in the next subsection.

4.2.2. Modeling Effect Analysis of Different Numbers of ATIs

The number of ATIs needs to be given before dividing the ATIs using the C-Means clustering algorithm. Thus, the number of ATIs has an important influence on the prediction effects. In order to obtain the optimal number of ATIs, the TEPMs of different ATIs are established. Then, the established TEPMs are used to predict the experimental data of an entire year. According to the prediction effects, the optimal number of ATIs can be obtained using the elbow method.

As for the established TEPMs of ATIs, the mean and standard deviations [31] of the Rsd of the 46 batches of experiments can be calculated as follows:

$$S_M = \frac{1}{K} \sum_{k=1}^K Rsd_k, \quad (20)$$

$$S_D = \sqrt{\frac{\sum_{k=1}^K (Rsd_k - S_M)^2}{K-1}}, \quad (21)$$

where Rsd_k is the Rsd value of the k th batch of experiments, and they can be calculated by Equation (17). $K = 46$ is the total number of batches of the predicted experiments. S_M and S_D are used to characterize the prediction accuracy and robustness of the model.

In this study, the number of ATIs ranges from 2 to 6 with a step size of 1. Then, the prediction accuracy and robustness with different number of ATIs are calculated using Equations (20) and (21). The calculation results are plotted in Figure 9 to visually show the influence of the number of ATIs on the prediction effects. It can be observed from Figure 9 that the prediction effects gradually improved as the number of ATIs increases. Marked by the black circle in Figure 9, when the ATIs reaches four, the prediction effects gradually stabilize. Therefore, the number of ATIs is determined as four according to the elbow method. As a side note, the method in [25] with one segmentation point is only one special case of our generic method when the number of ATIs is two.

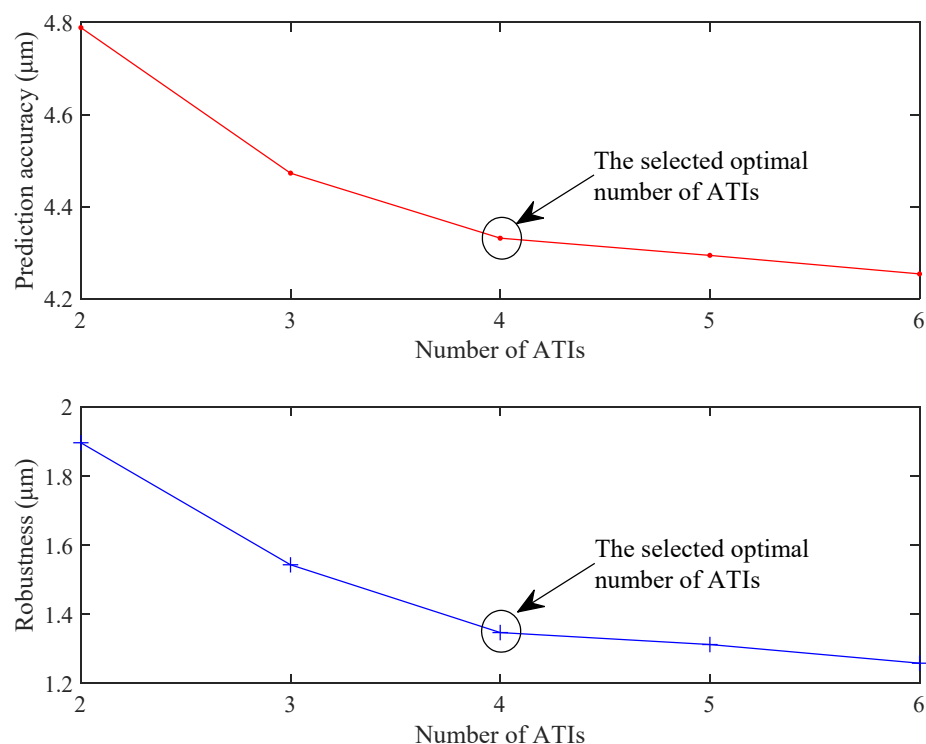


Figure 9. Prediction effects with different numbers of ATIs.

4.3. Prediction Performance Analysis with Annual Experimental Data

The TEPMs corresponding to 4 ATIs are shown in Equation (19), which are used to predict the annual thermal error data. The low and high ambient temperature models Y_1 and Y_{46} were also used to predict the annual thermal error data. The prediction results Rsd of these three models were calculated and plotted in Figure 10. Note that Figure 10 is plotted based on 46 batches of data, which were collected over an entire year, as mentioned before.

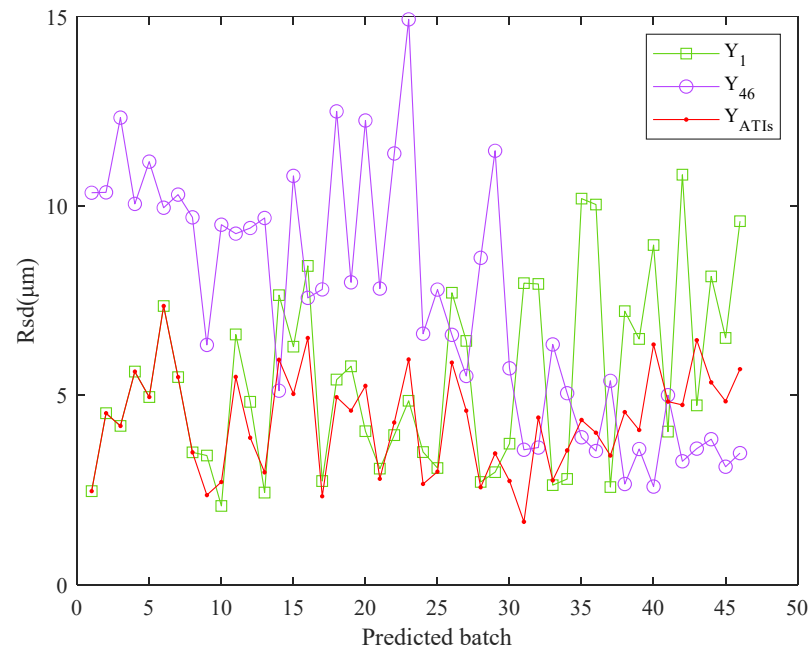


Figure 10. Prediction results of the three models for the annual experiments.

As shown in Figure 10, the ATI model has higher prediction accuracy than the low and high ambient temperature models. This is because the proposed year-round thermal error modeling method not only takes the ATIs into consideration but also determines the optimal number of ATIs to achieve better performance, which is a key difference from the existing literature. Furthermore, the prediction results of three models are summarized in Table 6. Table 6 reveals that the TEPMs established based on the ATIs exhibit high prediction accuracy and robustness throughout the year. Further calculation showed that the prediction accuracy of the proposed ATI model was 20.6% and 41.7% higher than those of the Y_1 and Y_{46} , respectively, and robustness was improved by 48.8% and 62.0%, respectively. In the thermal error compensation process, the TEPM corresponding to the ATI could be invoked for prediction and compensation, according to the actual initial ambient temperature.

Table 6. Prediction results of the three models (Unit: μm).

Model	ATI Model	Y_1	Y_{46}
Mean of Rsd	4.33	5.45	7.43
Standard deviation of Rsd	1.24	2.42	3.26

Mean of Rsd Standard deviation of Rsd

5. Conclusions

In this study, the one-way ANOVA was first conducted to show that the ambient temperature has a significant impact on the spindle thermal error data at a significant level of 0.001. Therefore, a thermal error modeling and compensation method for spindle of

machine tool based on ATIs was proposed in this paper. Based on the annual experimental data, the ATIs applicable to the TEPMs were investigated, and the number of ATIs based on the C-Means clustering algorithm was determined as 4. The prediction accuracy and robustness of the proposed method were 4.33 μm and 1.24 μm , respectively, demonstrating the high prediction accuracy and robustness of the method throughout the year. In addition, the prediction accuracy of the proposed ATI model was 20.6% and 41.7% higher than those of the Y_1 and Y_{46} , respectively, and robustness improved by 48.8% and 62.0%, respectively. This study can serve as an important reference for machine tools of thermal error modeling and compensation.

Obtaining the applicable ATIs is the basis of the proposed method. In fact, the number of ATIs is related to the range of the applicable ATI of the TEPM, which must be determined according to the actual scenario. Factors affecting the range of the applicable ATI, such as the modeling algorithm and the machine tool itself, should be the focus of future research.

Author Contributions: Conceptualization, X.W. and H.Y.; methodology, X.W. and H.Y.; validation, X.F. resources, X.W. and H.Y.; writing—original draft preparation, X.W. and H.Y.; writing—review and editing, X.F.; funding acquisition, X.W. and X.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Anhui Provincial Key Research and Development Project of China (grant number 2022f04020005), the General project of Anhui Natural Science Foundation (grant number: 1908085ME134), and the Major natural science research projects in Anhui Universities (grant number KJ2021ZD0042).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors thank Xindong Zhuang for assistance with the investigation and visualization.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Mayr, J.; Jedrzejewski, J.; Uhlmann, E.; Donmez, M.A.; Knapp, W.; Härtig, F.; Wendt, K.; Moriwaki, T.; Shore, P.; Schmitt, R.; et al. Thermal issues in machine tools. *CIRP Ann.* **2012**, *61*, 771–791. [\[CrossRef\]](#)
2. Gebhardt, M.; Mayr, J.; Furrer, N.; Widmer, T.; Weikert, S.; Knapp, W. High precision grey-box model for compensation of thermal errors on five-axis machines. *CIRP Ann.* **2014**, *63*, 509–512. [\[CrossRef\]](#)
3. Ramesh, R.; Mannan, M.; Poo, A. Error compensation in machine tools—A review Part II: Thermal errors. *Int. J. Mach. Tools Manuf.* **2000**, *40*, 1257–1284. [\[CrossRef\]](#)
4. Ma, C.; Liu, J.; Wang, S. Thermal error compensation of linear axis with fixed-fixed installation. *Int. J. Mech. Sci.* **2020**, *175*, 105531. [\[CrossRef\]](#)
5. Volk, W.; Groche, P.; Brosius, A.; Ghiotti, A.; Kinsey, B.L.; Liewald, M.; Madej, L.; Min, J.; Yanagimoto, J. Models and modelling for process limits in metal forming. *CIRP Ann.* **2019**, *68*, 775–798. [\[CrossRef\]](#)
6. Miao, E.; Gong, Y.; Niu, P.; Ji, C.; Chen, H. Robustness of thermal error compensation modeling models of CNC machine tools. *Int. J. Adv. Manuf. Technol.* **2013**, *69*, 2593–2603. [\[CrossRef\]](#)
7. Yang, J.; Shi, H.; Feng, B.; Zhao, L.; Ma, C.; Mei, X. Applying neural network based on fuzzy cluster pre-processing to thermal error modeling for coordinate boring machine. *Procedia CIRP* **2014**, *17*, 698–703. [\[CrossRef\]](#)
8. Abdulshahed, A.M.; Longstaff, A.P.; Fletcher, S.; Potdar, A. Thermal error modelling of a gantry-type 5-axis machine tool using a grey neural network model. *J. Manuf. Syst.* **2016**, *41*, 130–142. [\[CrossRef\]](#)
9. Wei, X.; Miao, E.; Liu, H.; Liu, S.; Chen, S. Two-dimensional thermal error compensation modeling for worktable of CNC machine tools. *Int. J. Adv. Manuf. Technol.* **2019**, *101*, 501–509. [\[CrossRef\]](#)
10. Cheng, Q.; Qi, Z.; Zhang, G.; Zhao, Y.; Sun, B.; Gu, P. Robust modelling and prediction of thermally induced positional error based on grey rough set theory and neural networks. *Int. J. Adv. Manuf. Technol.* **2015**, *83*, 753–764. [\[CrossRef\]](#)
11. Zhang, T.; Ye, W.; Shan, Y. Application of sliced inverse regression with fuzzy clustering for thermal error modeling of CNC machine tool. *Int. J. Adv. Manuf. Technol.* **2016**, *85*, 2761–2771. [\[CrossRef\]](#)
12. Tan, B.; Mao, X.; Liu, H.; Li, B.; He, S.; Peng, F.; Yin, L. A thermal error model for large machine tools that considers environmental thermal hysteresis effects. *Int. J. Mach. Tools Manuf.* **2014**, *82–83*, 11–20. [\[CrossRef\]](#)

13. Shi, H.; Jiang, C.; Yan, Z.; Tao, T.; Mei, X. Bayesian neural network-based thermal error modeling of feed drive system of CNC machine tool. *Int. J. Adv. Manuf. Technol.* **2020**, *108*, 3031–3044. [[CrossRef](#)]
14. Mayr, J.; Blaser, P.; Ryser, A.; Hernandez-Becerro, P. An adaptive self-learning compensation approach for thermal errors on 5-axis machine tools handling an arbitrary set of sample rates. *CIRP Ann.* **2018**, *67*, 551–554. [[CrossRef](#)]
15. Chen, Y.; Chen, J.; Xu, G. A data-driven model for thermal error prediction considering thermoelasticity with gated recurrent unit attention. *Measurement* **2021**, *184*, 109891. [[CrossRef](#)]
16. Zhang, Z.P.; Liu, K.; Guo, F. Prediction of machine tool thermal error compensation based on SVMR and ARM11. *Appl. Mech. Mater.* **2015**, *740*, 120–126. [[CrossRef](#)]
17. Zimmermann, N.; Breu, M.; Mayr, J.; Wegener, K. Autonomously triggered model updates for self-learning thermal error compensation. *CIRP Ann.* **2021**, *70*, 431–434. [[CrossRef](#)]
18. Liu, H.; Miao, E.M.; Wei, X.Y.; Zhuang, X.D. Robust modeling method for thermal error of CNC machine tools based on ridge regression algorithm. *Int. J. Mach. Tools Manuf.* **2017**, *113*, 35–48. [[CrossRef](#)]
19. Liu, H.; Miao, E.; Zhuang, X.; Wei, X. Thermal error robust modeling method for CNC machine tools based on a split unbiased estimation algorithm. *Precis. Eng.* **2018**, *51*, 169–175. [[CrossRef](#)]
20. Wei, X.; Ye, H.; Miao, E.; Pan, Q. Thermal Error Modeling and Compensation Based on Gaussian Process Regression for CNC Machine Tools. *Precis. Eng.* **2022**, *77*, 65–76. [[CrossRef](#)]
21. Zimmermann, N.; Lang, S.; Blaser, P.; Mayr, J. Adaptive input selection for thermal error compensation models. *CIRP Ann.* **2020**, *69*, 485–488. [[CrossRef](#)]
22. Wu, C.; Xiang, S.; Xiang, W. Spindle thermal error prediction approach based on thermal infrared images: A deep learning method. *J. Manuf. Syst.* **2021**, *59*, 67–80. [[CrossRef](#)]
23. Zhang, C.; Gao, F.; Li, Y. Thermal error characteristic analysis and modeling for machine tools due to time-varying environmental temperature. *Precision Eng.* **2017**, *47*, 231–238. [[CrossRef](#)]
24. Li, F.; Li, T.; Jiang, Y.; Wang, H.; Ehmann, K.F. Explicit error modeling of dynamic thermal errors of heavy machine tool frames caused by ambient temperature fluctuations. *J. Manuf. Process.* **2019**, *48*, 320–338. [[CrossRef](#)]
25. Liu, Y.; Miao, E.; Liu, H.; Chen, Y. Robust machine tool thermal error compensation modelling based on temperature-sensitive interval segmentation modelling technology. *Int. J. Adv. Manuf. Technol.* **2019**, *106*, 655–669. [[CrossRef](#)]
26. Miao, E.; Liu, Y.; Liu, H.; Gao, Z.; Li, W. Study on the effects of changes in temperature-sensitive points on thermal error compensation model for CNC machine tool. *Int. J. Mach. Tools Manuf.* **2015**, *97*, 50–59. [[CrossRef](#)]
27. Shen, J.H.; Yang, J.G. Application of partial least squares neural network in thermal error modeling for CNC machine tool. *Key Eng. Mater.* **2009**, *392–394*, 30–34. [[CrossRef](#)]
28. Fan, L.; Jing, X.; Chen, J.; Zhang, K.; Li, D.; Shi, H. Research on thermal deformation modeling of high-speed motorized spindle based on PLS. In Proceedings of the 2018 Chinese Control And Decision Conference (CCDC), Shenyang, China, 9–11 June 2018.
29. ISO 230-3; Test Code for Machine Tools—Part 3: Determination of Thermal Effects. International Organization for Standardization: Geneva, Switzerland, 2020; p. 2020.
30. Anders, K. Resolution of Students t-tests, ANOVA and analysis of variance components from intermediary data. *Biochemia Medica.* **2017**, *27*, 253–258. [[CrossRef](#)]
31. Wei, X.; Feng, X.; Miao, E.; Qian, M.; Pan, Q. Sub-regional thermal error compensation modeling for CNC machine tool worktables. *Precision Eng.* **2022**, *73*, 313–325. [[CrossRef](#)]