Selection of key temperature measuring points for thermal error modeling of CNC machine tools

Quan LIU*, Yuting CHEN*, Ping LOU*, Junwei YAN*, Jiwei HU* and Angran XIAO**

*School of Information Engineering, Wuhan University of Technology
122 Ruoshi Road, Hongshan District, Wuhan, Hubei 430070, China
E-mail: Louping@whut.edu.cn

**Department of Mechanical Engineering Technology, New York City College of Technology, City University of New York
Brooklyn, NY 11201, US.

Received: 28 April 2018; Revised: 18 September 2018; Accepted: 12 November 2018

Abstract
Thermal errors account for more than 40% of the machining errors in CNC machine tools, and even reach 70% in precision and ultra-precision machine tools. One approach to reducing thermal errors is to build thermal error prediction models by monitoring the temperature field of machine tools with the data-driven modeling approaches. Usually the data-driven approaches have over-fitting and under-fitting problems. The prediction performances of the data-driven models are greatly dependent on the input of the models, namely the number of temperature measuring points (TMPs) and their locations. In this paper, a selection method of key TMPs is presented to improve the accuracy and generalization of predictive models. In this selection method, correlation analysis is used to eliminate the uncorrelated or weakly correlated TMPs to thermal errors at first; then the Minimum Redundancy Maximum Relevance (MRMR) is presented to narrow the searching scope of TMPs; finally Wrapper method is used to test the candidate set of TMPs with cross-validation accuracy to find the key TMPs. To validate this method, a test platform is built on the CNC gantry drilling machine ZK5540A. 56 Fiber Bragg Grating (FBG) temperature sensors are amounted on the body, column, spindle, and base of this machine. 7 key TMPs was selected from the 56 ones with this method. And Multiple linear regression (MLR) approach is used to build the thermal error prediction model with the 7 key TMPs, 56 TMPs and key TMPs selected by Miao's method(Miao et al., 2014) respectively. The result shows that the prediction model built with MLR using the 7 key TMPs is much more accurate and has stronger generalization.

Keywords: CNC machining tools, Thermal error prediction model, Temperature Measuring Point(TMP), Minimum Redundancy Maximum Relevance (MRMR), Wrapper

1. Introduction

As the major contributor of machining errors in CNC, thermal errors are attracting great research attentions(Wang, 2006, Li et al., 2015, Dai et al., 2017). Statistically, thermal errors can account for around 40-70% of the overall machining errors(Bryan, 1990, Weck et al., 1995), and are even more significant in precision and ultra-precision machining. There are several solutions to reduce thermal errors, including temperature control in the machining process, thermal stable design of machine tools, and real-time thermal error compensation system(Ni, 1997). For instance, thermal equilibrium can be achieved and controlled during machining process using cooling systems; machine tools can be designed with symmetric structures and manufactured using materials with high thermal stiffness. Most of these methods are capable of reducing thermal errors but cannot eliminate them, and usually increase the manufacturing cost of the machine tools(Jedrzejewski and Modrzycki, 1992, Ramesh et al., 2000). The thermal error compensation method, on the other hand, is much more convenient and cost-efficient(Aguirre et al., 2014), because it can be applied directly without requiring any expensive modification to the machine tools.

A thermal error model is the mathematical formula representing the relationships between the thermal deformations of a machine tool and the temperature variables at different locations of the machine and its working environment, also
referred to as temperature field. Thermal error data is usually collected using displacement sensors. Temperature field data is usually collected by placing temperature sensors on the body of machine tools. During machining, the predicted deformation values are used to adjust the machine so that the resulting machining errors of the product can be greatly reduced (White et al., 2001). Apparently, the success of the entire compensation process relies greatly on the accuracy of the thermal error model, which depends on the key TMPs. Intuitively, the more TMPs, the better chance we can capture the temperature field accurately. However, besides incurring higher purchasing and installation costs, more TMPs do not necessarily result into more accurate thermal error model (Li et al., 2008). Some researchers select key TMPs, and use only the data from key TMPs to construct thermal error model. Veldhuis et al. select key TMPs using correlation analysis. If a TMP’s data is strongly correlated with the values of a thermal error, it is considered a key TMP (Veldhuis and Elbestawi, 1995). This method, however, does not consider the correlation among TMPs which can greatly affect the accuracy of the predicative model. Clustering analysis method is used to group all TMPs based on their similarities. Within a group, the TMP having the highest correlation coefficient with thermal error is chosen as the key TMP of the group. Since TMPs across different groups have relatively weaker similarities, the problem of collinearity can hence be greatly reduced. Han, et al. use fuzzy c-means algorithms to select a key TMP within a group; they select 4 key TMPs from 32 in the case study (Han et al., 2012). Zhang, et al. present an optimal clustering threshold for key TMPs selection; they select 7 key TMPs out of 29 in the case study (Zhang et al., 2013). Although the method is “practical and effective” at handling correlation among TMPs, it is very difficult to determine an appropriate threshold, which is critical in the method. Li, et al. combine clustering method with genetic algorithm (Li et al., 2017), Miao et al. combined fuzzy clustering, gray correlation and stepwise regression to select key TMPs (Miao et al., 2014). Yan et al. used grey correlation (Yan and Yang, 2009, Wang, 2006). As a numeric analysis method, finite element method (FEM) was used to identify key TMPs (Kim and Cho, 1997, Li et al., 2017), but the boundary conditions are difficult to determine. Other methods, such as thermal error sensitivity analysis (Liu et al., 2016) and rough sets theory (Cheng et al., 2016), are also effective to some extent. Generally, The current TMPs selection methods are mainly based on eliminating redundant TMPs by clustering or eliminating noise TMPs by analyzing correlation between TMPs and thermal error. Such methods usually involve setting parameters, such as clustering threshold. And these parameters can affect the number of selected TMPs. And the values of these parameters are mainly determined by researchers’ experience, although the number of selected TMPs may be crucial to thermal error model (Zhang et al., 2013). Too less TMPs may lead to poor prediction accuracy while too many may have a negative influence on a model’s robustness (Abdulshahed et al., 2016). Therefore, in this research, except for considering eliminating noisy TMPs and redundant TMPs, potential subsets searching based on MRMR and wrapper method are combined to determine the optimal number of selected TMPs, which can contribute to the performance of thermal error model.

In the following section, the key TMPs selection method is introduced in detail. In Section 3, a ZK5540A CNC gantry drilling machine with 56 TMPs is used as case study. To verify the efficaciousness of the proposed method, three thermal error models representing the positioning of the drill bit vs. different spindle speeds are constructed using the selected key TMPs, all 56 TMPs and key TMPs selected by Miao’s method (Miao et al., 2014) respectively. Comparing with the measured thermal deformations shows that the key TMPs selection method results into a more accurate thermal error model. The contents of the manuscript.

2. The Selection Method of Key TMPs

The key TMPs selection method has three steps as shown in Fig. 1. The first step is preliminary selection. All TMPs are screened to remove the ones that are not or weakly correlated with the thermal error. The second step is potential subsets searching. Subset searching method based on MRMR is used to search the potential subsets of TMPs. The last step, all potential subsets are assessed by wrapper method to identify the optimal one.
2.1 Preliminary Selection

A thermal error prediction model actually builds the functional relationship between thermal errors and the temperature field. That uncorrelated or weakly correlated TMPs are used as the input of the prediction model will incur the performance to be degraded, hence they need to be removed. Suppose all TMPs to be represented as a set $T$,

$$ T = \{ t_1, t_2, ..., t_n \} $$

where $t_i$ is the $i$th TMPs, and the total number of all TMPs is $n$.

Each TMP $t_i$ has a sequence of measured temperature values, denoted as:

$$ t_i = \{ t_i(1), t_i(2), ..., t_i(m) \} (i = 1, 2, ..., n) $$

where $t_i(j)$ is the $j$th measurement of TMPs $t_i$ and $m$ is the total number of collected temperature values in each sequence.

Correspondingly, the thermal error can be denoted as a set $E$.

$$ E = \{ E(1), E(2), ..., E(m) \} $$

where $E(j)$ is the $j$th measurement of $E$ and $m$ is the total number of collected thermal error values because every time when measurement is taken, both temperatures and thermal error values are collected.

Then the correlation coefficient between each TMP and the thermal error is calculated using Eq. (4).

$$ \rho(t_i, E) = \frac{\sum_{j=1}^{m} (t_i(j) - \bar{t}_i)(E(j) - \bar{E})}{\sqrt{\sum_{j=1}^{m} (t_i(j) - \bar{t}_i)^2 \sum_{j=1}^{m} (E(j) - \bar{E})^2}} $$

where $\bar{t}_i$ denotes the average of TMPs $t_i$, $\bar{E}$ represents the average of thermal error $E$. Suppose TMPs satisfying to $\rho(t_i, E) \geq \alpha (0 \leq \alpha \leq 1)$ to be selected the preliminary set. $\alpha$ is defined empirically. Usually the value of $\alpha$ can be set as 0.7, because $\rho > 0.7$ indicate strong correlation (Hsiung et al., 2002, Sabir et al., 2017)). The value of $\alpha$ can be modified according to specific conditions.

After the preliminary selection, only the TMPs strongly correlate to the thermal error are selected as candidate TMPs for further consideration. The TMPs having $\rho < \alpha$ are weakly correlated, meaning that temperature variables at those locations do not contribute significantly to the thermal error; hence do not need to be included in the model.
2.2 Potential Subsets Searching based on MRMR

Suppose there are $p$ candidate TMPs at the preliminary selection, there will be $2^p - 1$ possible subsets. If the value $p$ is large, taking all possible subsets into consideration will bring in extensive calculation. Thus, a potential subset searching scheme is made to narrow the search scope. The MRMR is used to search the potential subset (Battiti, 1994, Amiri et al., 2011, Yang and Beddoe, 2010). The searching process is depicted in Fig. 2.

At first, the preliminary TMPs are initialized, defined as set $T_1$. The searching procedure is as follows.

$$T_1 = \{ t'_1, t'_2, t'_3, ..., t'_p \}$$

where $t'_i$ is the $i$th preliminary TMPs, $p$ is the number of preliminary TMPs.

The measured temperature values of $t'_i$ are expressed by the sequence:

$$t'_i = \{ t'_i (1), t'_i (2), ..., t'_i (j), ..., t'_i (m) \} \quad (i = 1,2,..., p)$$

where $t'_i (j)$ is the $j$th measurement of TMP $t'_i$, $m$ is the number of temperature values collected from each TMP.

This searching method for potential subset is a forward selection strategy (Amiri et al., 2011). The search starts with an empty set. At first, the most correlated TMPs $t'_i$ to thermal error is found from the candidate TMPs, and then put it into the set $S_1$; From the remaining TMPs in the pool, the TMPs with the biggest value of the evaluation is selected with Eq. (8); Then, we put the selected TMPs into $S_1$ to generate set $S_2$; The procedure will be continued until the candidate TMPs are empty. Therefore, according to the order in which the potential subsets are generated, the potential subset is defined as $S_k (k = 1,2,...,p)$. In subset $S_k$, there are $k$ TMPs. The inclusion relation among these potential subsets is $S_1 \subseteq S_2 \subseteq S_3 \subseteq \cdots \subseteq S_p$.

$$\rho(t'_i, E) = \frac{1}{k} \sum_{\substack{t'_j \in S_k \setminus t'_i}} \rho(t'_i, t'_j), (t'_i, t'_j) \in T_1, t'_j \in S_k$$

Initialization:

- $T_1 = \{ t'_1, t'_2, t'_3, ..., t'_p \}$: Pool of candidate TMPs;
- $E$: Thermal error;
- $S_k = \emptyset (k = 1,2,\ldots,p)$: Potential subsets.

For each TMPs in $T_1$, compute $\rho(t'_i, E)$

Searching the first subset:

Finding the TMPs $t'_i$ with the biggest value of $\rho(t'_i, E)$;

Set $k = 1$, perform $S_1 \leftarrow t'_i$, $T_i \leftarrow T_1 \setminus \{ t'_i \}$

For each TMPs in $S_k$ Compute $\rho(t'_i, t'_j)$, ($t'_i \in T_i, t'_j \in S_k$)

Searching the next subset:

Finding the TMPs $t'_i$ with the biggest value of $\rho(t'_i, E)$

Set $k \leftarrow k + 1$, perform $S_k \leftarrow S_{k-1} \cup t'_i$, $T_i \leftarrow T_1 \setminus \{ t'_i \}$

$k < p$ Yes

No End

Fig. 2 Process of Potential Subsets Searching

[DOI: 10.1299/jamdsm.2018jamdsm0131] © 2018 The Japan Society of Mechanical Engineers
2.3 The Optimal Subset Identification based on Wrapper Method

Here, wrapper method is used to identify the optimal TMPs from the potential subsets. Wrapper method is proposed by Kohavi and John (Kohavi and John, 1997), it offers an effective and efficiency method to solve the difficulty of variable or feature selections. Wrappers utilize the learning machine of interest as a black box to score subsets of variable according to their predictive power (Guyon et al., 2003). The MLR model has been widely used in thermal error modeling (Wu and Kung, 2006, Li et al., 2017) and it has characteristics of simple structure and quick learning. Thus, the MLR thermal error prediction model is built to be used to assess all potential subsets. The v-fold cross-validation method is used as the estimation method to control over-fitting (Hsu, 2004). The optimal subset of TMPs is identified on the basis of RSS value obtained by V-fold cross-validation method.

Given \( \{s_1,s_2,...,s_p\} \) as the potential subsets selected from \( T_1 \), subset \( s_k \) is defined as

\[
s_k = \{t'_a,t'_b,...,t'_c\} (k=1,2,...,p)
\]  

(8)

The measured experimental data is defined as

\[
D(j) = \{t'_1(j), t'_2(j),...,t'_p(j); E(j)\} (j = 1,2,...,m)
\]

(9)

where \( t'_i(j) \) is the \( jth \) measurement of TMPs \( t'_i \). \( E(j) \) is the \( jth \) measurement of \( E \). the \( m \) is the number of collected data.

At first, the number of folds \( V \) is decided. Each subset of the TMPs will be tested using V-fold cross-validation method. That is, the experimental data is divided into \( V \) mutually exclusive and equal sized folds(Kohavi 1995).

For a subset \( s_k = \{t'_a,t'_b,...,t'_c\} \), data set \( D_k(j) \) is extracted from \( D(j) \).

\[
D_k(j) = \{t'_a(j), t'_b(j),...,t'_c(j); E(j)\} (j = 1,2,...,m)
\]

(10)

Divide \( D_k(j) \) into \( V \) mutually exclusive and equal sized folds \( D_k^{(1)}, D_k^{(2)},..., D_k^{(V)} \) and One folder is expressed as \( D_k^{(v)} \):

\[
D_k^{(v)} = \{D_k((v-1)*\frac{m}{V} + 1), D_k((v-1)*\frac{m}{V} + 2),..., D_k(v*\frac{m}{V})\}, (v = 1,2,...,V)
\]

(11)

These folders are used to train and validate the MLR thermal error model \( V \) times. That is, each time, one folder is used as validation data and the rest of are used to create and train a thermal error model. Each time, the RSS of the model is calculated. The overall RSS of the subset \( s_k \) is the average of the \( V \) RSS values. For instance, given 10 folders of data, in the first round \( D_k^{(1)} \) is used as validation data. \( D_k^{(2)} - D_k^{(10)} \) are used to create a thermal model and \( rss_k^{(1)} \) is calculated using \( D_k^{(1)} \). In the next round, \( D_k^{(2)} \) is used as validation data to calculate \( rss_k^{(2)} \). The overall RSS is the average of \( rss_k^{(1)}, rss_k^{(2)} \) and \( rss_k^{(10)} \).

Each time \( v = 1,2,...,V \), a MLR thermal error model \( f_k^{(v)} \) based on \( s_k = \{t'_a,t'_b,...,t'_c\} \) and \( E \) is build using construction data \( D_k^{(1)} \cup D_k^{(2)} \cup D_k^{(3)} \cup ... \cup D_k^{(V)} \). And the RSS of the model \( f_k^{(v)} \), which is represented by \( rss_k^{(v)} \) is calculated based on validation data \( D_k^{(v)} \)

\[
E = f_k^{(v)}(t'_a,t'_b,...,t'_c)
\]

(12)

The \( rss_k^{(v)} \) is calculated using Eq. (13).

\[
rss_k^{(v)} = \frac{1}{N} \sum_{j=1}^{N} (E(j) - E_k^{(v)}(j))^2
\]

(13)

\[
E_k^{(v)}(j) = f_k^{(v)}(t'_a(j),t'_b(j),...,t'_c(j))
\]

(14)

Where \( E(j) \) is \( jth \) measured value of thermal error. \( E_k^{(v)}(j) \) denotes \( jth \) predicted value by model \( f_k^{(v)} \). The average value of RSS is calculated using Eq. (15).

\[
RSS_{\text{mean}}(k) = \frac{\sum_{v=1}^{V} rss_k^{(v)}}{V}
\]

(15)
Finally, all $RSS_{mean}(k)$ values of all potential subsets are calculated. Since $RSS_{mean}(k)$ represents the predictive accuracy, the subset of TMPs resulting into the least $RSS_{mean}(k)$ value is chosen as optimal subset, which is capable of creating the most accurate thermal error model.

**Step1:** Decide the number of folds, $V$.

**Step2:** For each subset $S_k$, perform the procedure as below.

- Extract experimental data
- Build MLR thermal error model with construction data.
- Determine prediction error with validation data.
- Calculate the average value
- $RSS_{mean}(k)$

**Step3:** Subset with the lowest value of $RSS_{mean}(k)$ is the optimal subset.

**Fig. 3 Process of optimal Subset Identification**

### 3. Case Study

#### 3.1 The Test Platform

A ZK5540A CNC gantry drilling machine is used to demonstrate the efficaciousness of the key TMP selection method. To ensure the practicality, the whole experiment is conducted in a real production workshop. The experiment system contains a temperature measuring subsystem based on FBG sensors, and a thermal error measuring subsystem based on Charge Coupled Device (CCD) laser displacement sensors, as shown in Fig. 4. FBG sensor has many advantages such as small size, anti-electromagnetism, and easy deployment. 56 FBG sensors are installed systematically to acquire the whole picture of the temperature field, as shown in Table 1 and Fig. 5. Obviously, sensors distributed on the spindle are relatively more, because spindle rotation is the main heat source in this experiment. Reasons for the simplification is listed as below: (1) ZK5540A has a bilaterally symmetric structure, spindle is the main heat source. (2) ZK5540A is a heavy duty machine. This kind of big machines in which positioning speeds are low, temperature increase due to friction or deformation of guideways does not significantly affect to the thermal state of the machine. (3) By monitoring the temperature change of the main parts of ZK5540A in the actual machining, it is found that the temperature change of the spindle is most significant while other parts are not obvious.
Table 1 FBG Sensors Mounting Schema in the Case Study

| Position       | FBG sensors |
|----------------|-------------|
| spindle        | $t_1, t_2, \ldots, t_{23}$ |
| Beam           | $t_{24}, t_{25}, t_{26}, t_{27}, t_{28}$ |
| Right column   | $t_{29}, t_{30}, t_{31}, t_{32}, t_{33}$ |
| Left column    | $t_{34}, t_{35}, t_{36}, t_{37}, t_{38}, t_{39}, t_{40}$ |
| Left bed       | $t_{41}, t_{42}, t_{43}, t_{44}, t_{45}, t_{46}, t_{47}, t_{48}$ |
| Right bed      | $t_{49}, t_{50}, t_{51}, t_{52}$ |
| Environment    | $t_{53}, t_{54}, t_{55}, t_{56}$ |

The thermal errors are measured from the tip of the spindle of the drilling machine using three CCD laser displacement sensors in three directions, $x$, $y$ and $z$, as shown in Fig. 6. The values of thermal error and the corresponding temperatures are collected simultaneously. One set of data is read per minute.
In order to simulate actual machining processes, the spindle speed is changed every 15 minutes as shown in Fig. 7. It starts with 750 rpm and continues for 15 minutes. Then spindle speed is then changed to 3000 rpm. The experiment takes about 4.5 hours, 268 sets of data are collected. The data is divided into two groups: the first 100 sets are training data used for key TMPs selection and thermal error modeling; the rest 168 sets are validation data used for testing and validating the key TMPs selection method.

3.2 Thermal Error and TMPs Analysis

The thermal errors in the direction x, y and z are shown in Fig. 8. The thermal errors in the direction x and y are smaller than that in the direction z. So the thermal error of the direction z is considered in the test. Fig. 9 shows the monitoring result of the temperature field.

3.3 Key TMPs Selection

In the preliminary selection step, as introduced in Section 2.1, the correlation coefficient between thermal error and TMPs are calculated using Eq.(4). As stated before, the threshold of the correlation coefficient $\alpha$ is set to 0.7. In this case, 30 preliminary TMPs satisfy $\rho(t_i,E) \geq 0.7$, hence are selected from the original 56. The 30 TMPs are: $t_1$, $t_2$, $t_3$, $t_4$, $t_5$, $t_6$, $t_7$, $t_8$, $t_9$, $t_{10}$, $t_{11}$, $t_{12}$, $t_{13}$, $t_{14}$, $t_{15}$, $t_{16}$, $t_{17}$, $t_{18}$, $t_{19}$, $t_{20}$, $t_{21}$, $t_{22}$, $t_{23}$, $t_{24}$, $t_{25}$, $t_{26}$, $t_{27}$, $t_{28}$, $t_{29}$, $t_{30}$, $t_{31}$, $t_{32}$, $t_{33}$, $t_{34}$, $t_{35}$, $t_{36}$, $t_{37}$, $t_{38}$, $t_{39}$, $t_{40}$, $t_{41}$, $t_{42}$, $t_{43}$, $t_{44}$, $t_{45}$, $t_{46}$, $t_{47}$, $t_{48}$, $t_{49}$, $t_{50}$, $t_{51}$, $t_{52}$, $t_{53}$, $t_{54}$, the preliminary TMPs, as shown in Fig. 10. Referring back to Table 1, it can be
seen that most of the TMPs are on the spindle. It shows that the preliminary selection step is capable of removing most of the irrelevant TMPs.

In the potential subset searching step, as introduced in Section 2.2, potential subsets were identified from the preliminarily TMPs. In this case, after using the forward selection method, 30 potential subsets $S_1, S_2, \ldots, S_{30}$ are selected. The search procedure is presented in Fig. 11.

The last step is identified the key optimal subset by wrapper method, as introduced in Section 2.3. The potential subsets $S_1, S_2, \ldots, S_{30}$ were tested and evaluated by the tester based on MLR model. And 10-fold cross-validation is commonly used (Bengio and Gr, 2003, Davison and Hinkley, 1997), so the training data is divided into 10 even folds in the case study. The RSS mean value of the 30 potential subsets are shown in Fig. 12. Since the RSS mean value of $S_7$ is the smallest, it is determined as the optimal subset. The 7 TMPs in subset $S_7$ are $t_{10}, t_{20}, t_{22}, t_{1}, t_{10}, t_{14}, t_{21}$. As shown in Fig. 13, there are 6 TMPs distributed on the spindle and one distributed on the column.

![Graphs showing temperature variation over sampling number for different TMPs](image)

Fig. 9 Temperature of part TMPs
Fig. 10 TMPS after primary selection

Initialization

\[ T_1 = \{t_1, t_2, \ldots, t_{27}, t_{30}, t_{32}, t_{39}, t_{40}, t_{46}, t_{53}, t_{54} \} \]

\[ s_k = \emptyset (k = 1, 2, 3, \ldots, 30) \]

Searching the potential subsets

```
\begin{align*}
\text{s}_1 \leftarrow t_{19} & \quad \text{s}_{12} \leftarrow s_{11} + t_{34} \Rightarrow s_{13} \leftarrow \text{s}_{12} + t_{40} \\
\downarrow & \quad \uparrow & \quad \downarrow & \quad \uparrow & \quad \downarrow & \quad \uparrow & \quad \downarrow & \quad \uparrow & \quad \downarrow & \quad \uparrow
\vdots
\end{align*}
```

End

Fig. 11 Potential Subsets Search Process in the Case Study
3.4 Predictive Thermal Error Modeling

MLR model is a typical kind of static model. It has merits of simple structure, fast training speed and lower computation. Many previous works (Han et al., 2012, Miao et al., 2014) indicate that MLR model can perform well in thermal error compensation. In addition, MLR model can also give a satisfactory performance when applied in real time compensation (CHEN and WANG, 2014). Thus, MLR model is used for thermal error compensation modeling to verify the effectiveness of the proposed TMPs selection method. Considering comparison is necessary for validating the effectiveness of the proposed method, MLR models of the thermal error based on the 7 key TMPs, original 56 TMPs, and key TMPs selected by Miao’s method (Miao et al., 2014) are made respectively.

The rest 168 sets of data are used to validate the model, as shown in Fig. 13, in which the solid curve represents the actual measured thermal error, while the dash curve represents the calculated errors from $MLR_1$. They are very close. The thermal error after compensation is represented using dotted curve. It is close to 0.

\[
\hat{E}_1 = \hat{b}_1 [1, t_{19}, t_{39}, t_{22}, t_{1}, t_{10}, t_{14}, t_{21}]'
\]

\[
\hat{b}_1 = [-0.5692 \quad 0.1121 \quad 0.0012 \quad 0.0003 \quad -0.0841 \quad -0.0244 \quad -0.0103 \quad 0.0048]
\]

$MLR_2$ is created using the original 56 TMPs.

\[
\hat{E}_2 = \hat{b}_2 [1, t_1, t_2, t_3, \ldots, t_{56}]'
\]
The predicated and actual thermal errors are presented in Fig. 14, in which the solid curve represents the actual measured thermal error, and the dash curve represents the calculated errors from the predicative model. It can be observed that at first the calculated error is very close to actual measured value, but it deviates the actual value pretty quickly. The accuracy and generalization ability of \( \text{MLR}_2 \) are weak.

\[ \hat{b}_{12} = [-1.1664, 0.0083, -0.0108, -0.0309, -0.0616, 0.0657, -0.0265, 0.0451, 0.0950, -0.0435, -0.003, 0.0058, 0.0293, -0.0332, -0.0001, 0.0016, -0.0212, -0.1210, 0.0015, -0.0031, -0.0051, 0.0037, 0.0001, 0.0023, -0.0163, -0.0193, -0.0582, 0.0483, 0.003, -0.0065, -0.0033, 0.0427, 0.0161, 0.0372, 0.0470, -0.0348, -0.0248, 0.002, 0.0339, -0.0167, -0.0330, -0.0122, 0.0117, 0.0033, 0.0318, 0.0142, 0.0465, -0.0386, 0.0469, -0.0344, 0.0014, -0.0273, -0.0139, 0.0004, -0.0015, 0.0015, 0.0016] \tag{19} \]

\[ \hat{E}_3(j) = \hat{b}_j [1, t_{10}, t_{14}, t_{20}, t_{24}, t_{25}, t_{41}, t_{47}, t_{48}, t_{49}, t_{54}, t_{55}] \] \tag{20} \]

\[ \hat{b}_j = [-4.9976, -0.0062, 0.0161, 0.0066, -0.0791, 0.2385, 0.2284, -0.0686] \tag{21} \]

In order to verify the effectiveness of proposed TMPs selection method, comparison should be made between the proposed method and other’s methods. So we follow Miao’s method to select TMPs (Miao et al., 2014). Firstly, the F-test is made to identify the optimal clustering threshold from 0.9659, 0.9935 and 0.9971. Through calculation, the optimal threshold is 0.9659 and the TMPs are clustered into 12 groups. Then, according to grey correlation, representative TMP of each cluster is identified. The representative TMPs are \( t_{10}, t_{14}, t_{20}, t_{24}, t_{25}, t_{41}, t_{47}, t_{48}, t_{49}, t_{54}, t_{55} \) respectively. Finally, 7 key TMPs are identified by stepwise regression. The 7 TMPs are: \( t_{10}, t_{14}, t_{20}, t_{24}, t_{25}, t_{41}, t_{47} \). In Miao’s research, \( \text{MLR}_3 \) model is also used to test the performance of TMPs selection methods. So \( \text{MLR}_3 \) is created using the key TMPs selected by Miao’s method.
The final thermal errors after compensation using these three models, $MLR_1$, $MLR_2$ and $MLR_3$ are listed in Table 2. After compensation with $MLR_1$, the average thermal error is reduced by 94.41%; the thermal error standard deviation is reduced by 41.89%; the maximum thermal error value is reduced by 68.08%. Meanwhile, after compensation with $MLR_2$, the average thermal error is reduced by only 23.06%; standard deviation and maximum error value even increase. After compensation with $MLR_3$, the average thermal error is reduced by 83.16%; the thermal error standard deviation is reduced by 35.92%; the maximum thermal error value is reduced by 78.11%. From the performance comparison of $MLR_1$ and $MLR_3$, it is concluded that the proposed TMPs selection method improves the accuracy of thermal error compensation. Two reasons can accounts for $MLR_2$ poor performance: firstly, the input of $MLR_2$ may include temperature variables which is uncorrelated or weak related to thermal error, the noise carried by these temperature has a negative influence on the robustness of thermal error model(Abdulshahed et al. 2016); secondly, the strong correlation among temperature variable will cause multicollinearity, which may also lead to the inaccuracy of prediction(El-Dereny and Rashwan, 2011). Comparing $MLR_1$ and $MLR_3$, it is found that the proposed TMPs selection method performs better than Miao’s method in reducing average error and reducing standard deviation. But the proposed method’s performance is slightly worse in reducing max error. Generally, the proposed method performs better than Miao’s. We deduce the reason is that in Miao’s method, clustering threshold is decided by a F-test comparison among several values which are set by experience. And in the proposed method, the number of keyTMPs is optimized by potential subsets searching based on MRMR and wrapper method. Apparently, the effectiveness of the proposed TMPs selection method is proved. Although the effectiveness of proposed method is only verified in an experiment taking spindle rotation as main heat source, it can be inferred the proposed method can still be effective when it comes to other heat source. The main reason is that the proposed method is based on data analysis. It does not depend on the location of heat sources. So the change of heat sources will not affect the performance of the proposed TMPs selection method. Besides, by applying the proposed TMPs selection method, most of the selected TMPs is located in spindle which is the main heat source in the experiment. Such result indicates that the method can effectively identify TMPs located in heat sources without any location information about heat sources.

Table 2 Compensation effect of $MLR_1$, $MLR_2$ and $MLR_3$

|                         | Absolute average (um) | Standard deviation (um) | Maximum value (um) |
|-------------------------|-----------------------------|--------------------------|----------------------|
| Original error          | 115.4892                   | 15.1526                  | 150.1895             |
| Error after compensation by $MLR_1$ | 6.4582                   | 8.8052                  | 47.9451              |
| Error after compensation by $MLR_2$ | 88.8584                   | 74.3844                  | 217.1284             |
| Error after compensation by $MLR_3$ | 19.4474                   | 9.7084                  | 32.8814              |

4. Conclusion

A key TMPs selection method combine subsets searching and wrapper method is presented in this paper. An experiment was carried out using a ZK5540 CNC gantry drilling machine. Using this method, 7 TMPs are selected from 56 ones. The accuracy of the resultant thermal error model and the result of compensation are acceptable, hence demonstrates the effectiveness of the presented method. Meanwhile, additional research works are being carried out to test the key TMPs selection method in actually machining processes, as well as other machine tools.

References

Abdulshahed, A. M., A. P. Longstaff, S. Fletcher and A. Potdar, Thermal error modelling of a gantry-type 5-axis machine tool using a Grey Neural Network Model, Journal of Manufacturing Systems, Vol. 41, (2016), pp. 130-142.

Aguirre, G., A. P. D. Nanclares and H. Urreta, Thermal Error Compensation for Large Heavy Duty Milling-boring Machines, Euspen Special Interest Group Meeting: Thermal Issues (2014).

[DOI: 10.1299/jamdsm.2018jamdsm0131] © 2018 The Japan Society of Mechanical Engineers
Amiri, F., M. M. R. Yousefi, C. Lucas, A. Shakery and N. Yazdani, Mutual information-based feature selection for intrusion detection systems, Journal of Network and Computer Applications, Vol. 34, No. 4 (2011), pp. 1184-1199.
Battiti, R., Using mutual information for selecting features in supervised neural net learning, IEEE Transactions on Neural Networks, Vol. 5, No. 4 (1994), pp. 537-550.
Bengio, Y. and Y. Gr, No Unbiased Estimator of the Variance of K-Fold Cross-Validation, Journal of Machine Learning Research, Vol. 5, No. 2003s-22 (2003), pp. 1089-1105.
Bryan, J., International Status of Thermal Error Research, CIRP Annals - Manufacturing Technology, Vol. 39, No. 2 (1990), pp. 645-656.
CHEN Song, WANG Yong-qing, Study on CNC Machine Tool Thermal Error Measurement Point Optimization Model Prediction and Real-time Compensation, Modular Machine Tool and Automatic Manufacturing Technique, No. 6(2014), pp. 88-91.
Cheng, Q., Z. Qi, G. Zhang, Y. Zhao, B. Sun and P. Gu, Robust modelling and prediction of thermally induced positional error based on grey rough set theory and neural networks, International Journal of Advanced Manufacturing Technology, Vol. 83, No. 5-8 (2016), pp. 753-764
Dai, H., S. Wang, X. Xiong, B. Zhou, S. Sun and Z. Hu, Thermal error modelling of motorised spindle in large-sized gear grinding machine, Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture, Vol. 231, No. 5 (2017), pp. 768 - 778.
Davison, A. C. and D. V. Hinkley, Bootstrap methods and their application, Cambridge University Press, (1997).
El-Dereny, M, and N. I. Rashwan, Solving multicollinearity problem using ridge regression models, Int. J. Contemp. Math. Sciences, Vol. 6, No. 12 (2011), pp.585 - 600
Guyon, I., A. Elisseeff and L. P. Kaelbling, An introduction to variable and feature selection, Journal of Machine Learning Research, Vol. 3, No. 6 (2003), pp. 1157-1182.
Han, J., L. Wang, H. Wang and N. Cheng, A new thermal error modeling method for CNC machine tools, International Journal of Advanced Manufacturing Technology, Vol. 62, No. 1-4 (2012), pp. 205-212.
Hsiung, M. W., L. Pai and H. W. Wang, Correlation between voice handicap index and voice laboratory measurements in dysphonic patients, European Archives of Oto-Rhino-Laryngology, Vol. 259, No. 2 (2002), pp. 97 - 99.
Hsu, W. H., Genetic wrappers for feature selection in decision tree induction and variable ordering in Bayesian network structure learning, Information Sciences, Vol. 163, No. 1 (2004), pp. 103-122.
Jedrzejewski, J. and W. Modrzycki, A New Approach to Modelling Thermal Behaviour of a Machine Tool under Service Conditions, CIRP Annals - Manufacturing Technology, Vol. 41, No. 1 (1992), pp. 455-458.
Kim, S. K. and D. W. Cho, Real-time estimation of temperature distribution in a ball-screw system, International Journal of Machine Tools and Manufacture, Vol. 37, No. 4 (1997), pp. 451-464.
Kohavi, R., A study of cross-validation and bootstrap for accuracy estimation and model selection, International Joint Conference on Artificial Intelligence (1995), pp. 1137-1143.
Kohavi, R. and G. H. John, Wrappers for feature subset selection, Artificial Intelligence, Vol. 97, No. 1-2 (1997), pp. 273-324.
Li, F., T. Li, H. Wang and Y. Jiang, A Temperature Sensor Clustering Method for Thermal Error Modeling of Heavy Milling Machine Tools, APPLIED SCIENCES-BASEL, Vol. 7, No. 1 (2017), pp. 82.
Li, T., F. Li, Y. Jiang and H. Wang, Thermal error modeling and compensation of a heavy gantry-type machine tool and its verification in machining, International Journal of Advanced Manufacturing Technology, Vol. 92, No. 9-12 (2017), pp. 1-20.
Li, Y., W. Zhao, S. Lan, J. Ni, W. Wu and B. Lu, A review on spindle thermal error compensation in machine tools, International Journal of Machine Tools and Manufacture, Vol. 95, (2015), pp. 20-38.
Li, Y. X., J. G. Yang, T. Gelvis and Y. Y. Li, Optimization of measuring points for machine tool thermal error based on grey system theory, International Journal of Advanced Manufacturing Technology, Vol. 35, No. 7-8 (2008), pp. 745-750.
Liu, Q., J. Yan, D. T. Pham, Z. Zhou, W. Xu, Q. Wei and C. Ji, Identification and optimal selection of temperature-sensitive measuring points of thermal error compensation on a heavy-duty machine tool, International Journal of Advanced Manufacturing Technology, Vol. 85, No. 1-4 (2016), pp. 345-353.
Miao, E. M., Y. Y. Gong, L. C. Dang and J. C. Miao, Temperature-sensitive point selection of thermal error model of CNC machining center, International Journal of Advanced Manufacturing Technology, Vol. 74, No. 5-8 (2014), pp.
Ni, J., CNC Machine Accuracy Enhancement Through Real-Time Error Compensation, Journal of Manufacturing Science and Engineering, Vol. 119, No. 4B (1997), pp. 717-725.

Ramesh, R., M. A. Mannan and A. N. Poo, Error compensation in machine tools — a review : Part II: thermal errors, International Journal of Machine Tools and Manufacture, Vol. 40, No. 9 (2000), pp. 1257-1284.

Sabir, B., B. Touri and M. Moussetad, Correlation between acoustic measures, voice handicap index and GRBAS scales scores among Moroccan students, Current Pediatric Research, Vol. 21, No. 2 (2017), pp. 343-353.

Tseng, P. C. and S. L. Chen, The Neural-fuzzy Thermal Error Compensation Controller on CNC Machining Center, Jsmie International Journal, Vol. 45, No. 2 (2002), pp. 470-478.

Veldhuis, S. C. and M. A. Elbestawi, A Strategy for the Compensation of Errors in Five-Axis Machining, CIRP Annals - Manufacturing Technology, Vol. 44, No. 1 (1995), pp. 373-377.

Wang, K, Thermal Error Modeling of a Machining Center using Grey System Theory and HGA-Trained Neural Network, 2006 IEEE Conference on Cybernetics and Intelligent Systems, (2006), pp. 1-7.

Weck, M., P. Mckeown, R. Bone and U. Herbst, Reduction and Compensation of Thermal Errors in Machine Tools, CIRP Annals - Manufacturing Technology, Vol. 44, No. 2 (1995), pp. 589-598.

White, A. J., S. R. Postlethwaite and D. G. Ford, A general purpose thermal error compensation system for CNC machine tools, Laser Metrology & Machine Performance V, (2001).

Wu, C. H. and Y. T. Kung, Thermal Analysis and Compensation of a Double-Column Machining Centre, Proceedings of the Institution of Mechanical Engineers Part B Journal of Engineering Manufacture, Vol. 220, No. 2 (2006), pp. 109-117.

Yan, J. Y. and J. G. Yang, Application of synthetic grey correlation theory on thermal point optimization for machine tool thermal error compensation, International Journal of Advanced Manufacturing Technology, Vol. 43, No. 11-12 (2009), pp. 1124-1132.

Yang, X. and G. Beddoe, Feature selection for computer-aided polyp detection using MRMR, Proceedings of SPIE - The International Society for Optical Engineering, Vol. 7624, No. 1 (2010), pp. 666-673.

ZHANG, Ting, Wenhua, LIANG, Ruijun, Peihuang, YANG and Xiaolan, Temperature Variable Optimization for Precision Machine Tool Thermal Error Compensation on Optimal Threshold, Chinese Journal of Mechanical Engineering, Vol. 26, No. 1 (2013), pp. 158-165.