euspen's 21st International Conference & Exhibition, Copenhagen, DK, June 2021
www.euspen.eu

Thermal error modeling for CNC machine tools using a random forest machine learning algorithm

Mengrui Zhu[1]*, Yun Yang[1], Xiaobing Feng[1], Zhengchun Du[1], Jianguo Yang[1]

[1] *State Key Laboratory of Mechanical System and Vibration, School of Mechanical Engineering, Shanghai Jiao Tong University, Shanghai, China*

zhumengrui@sjtu.edu.cn *(M. Zhu).*

Abstract
Thermal error modeling is essential to improving the machining accuracy as thermal error accounts for a majority of the total errors in CNC machine tools. Current modeling methods, such as physics-based modeling and conventional machine learning techniques (e.g. artificial neural networks) are limited in terms of accuracy and robustness. This paper presents a novel error modeling method for the spindle thermal errors based on the Random Forest algorithm. Spindle thermal errors in a three-axis machine tool as well as the temperature key points are both measured to train the proposed model. The model parameters of the regression trees are optimized by integrating the grid search method with five-fold cross-validation to derive the optimization tree model and prevent from overfitting. Comparing with the actual measurement results of spindle thermal errors, the model prediction accuracy is over 95%, which validates the proposed model. Compared to conventional modeling methods based on artificial neural networks, the proposed model requires less training data, achieves higher prediction accuracy, and is more robust against measurement noise.

Keywords: Thermal Error, Machine Tools, Machine Learning, Random Forest, Regression Tree

1. Introduction

Thermal error accounts for over 70% among the total errors of the machine tools. It is of great significance to model the thermal error accurately. Recently, various methods have been presented. For the physics-based method, the relevant researches focus on the mechanism analysis and generally require the precise physical assumption and mathematical derivation, which cannot be guaranteed in the actual machining process. For the data-driven method, such as the conventional machine learning technique (e.g. artificial neural network [1] and support vector machine), it aims to explore the data characteristics and establish the relationship between the temperature data and thermal error data without consideration of the intrinsic physical process. However, there are still exist some distinct disadvantages in terms of the model accuracy, robustness, and parameter tuning.

This paper presents a novel modeling method for the spindle thermal errors based on random forest (RF) to improve the model accuracy and robustness with the faster parameter tuning. The modeling processes are detailed. Furthermore, an experiment is conducted in a three-axis machine tool to validate the effectiveness of the proposed model.

2. Random forest

The random forest algorithm [2] is a tree-based ensemble learning method and has been widely applied to the classification and regression. It requires little data preparation, is simple to interpret, and less likely to overfit a dataset.

2.1. Model structure

Figure 1 illustrates the model structure of random forest which consists of a forest of decision trees from bootstrapped samples.

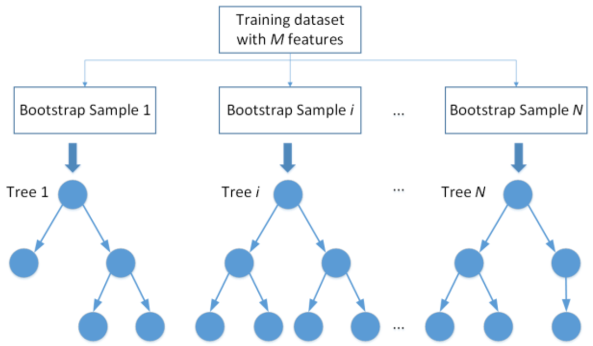


Figure 1. Model structure of random forest

Each decision tree has the branches and nodes. Each node represents a test on a certain input feature and each branch represents the output result of the test. The leaf node represents a class label for classification or a response for regression. The binary tree is usually for the regression where the response is continuous. In this paper, the thermal error prediction is a regression problem. Each decision tree is a weak learner. Multiple decision trees are grown in parallel to reduce the bias and variance of the model.

2.2. Model construction process

The modelling process using RF is divided as follows :

1) N bootstrapped sample sets of size n are randomly drawn from the original dataset.
2) Each bootstrapped sample set is divided to be the training dataset and testing dataset to grow a regression tree.
3) M features are randomly selected without replacement from all available features to be taken as split candidates.
4) At each node, choose the best split (i.e., the splitting feature and split point) based on the regression criteria that the residual sum of squares would be minimized after the split.
5) The above steps are repeated until N such trees are grown under the predefined stop criterion. The final response is predicted by averaging the predicted values of N trees.

2.3. Hyper-parametric tuning

Model's hyper-parameters affect its predictive performance, robustness and generalization capability. In the random forest algorithm, there are three important hyper-parameters, i.e. the number of trees, the maximum depth of the tree, the number of randomly selected features. The grid searching method is implemented to find the optimal values of the hyper-parameters. Five-fold cross-validation is integrated in this paper to prevent over-fitting problems and evaluate the model' performance.

3. Experimental Setup

A spindle thermal error experiment is conducted in a three-axis vertical machining center to collect the temperature and thermal error data. Four temperature sensors, T0, T1, T2, and T3 are used to measure the temperatures of the environment, front bearing, rear bearing, and spindle box respectively. An eddy current sensor is used to measure the axial thermal error△L of the spindle. The experimental setup is shown in figure 2.

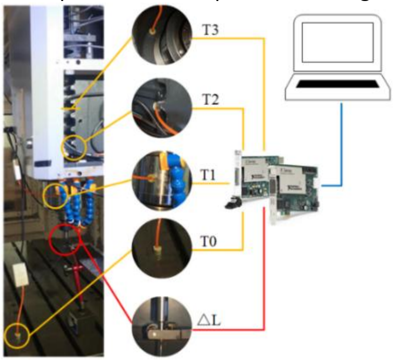
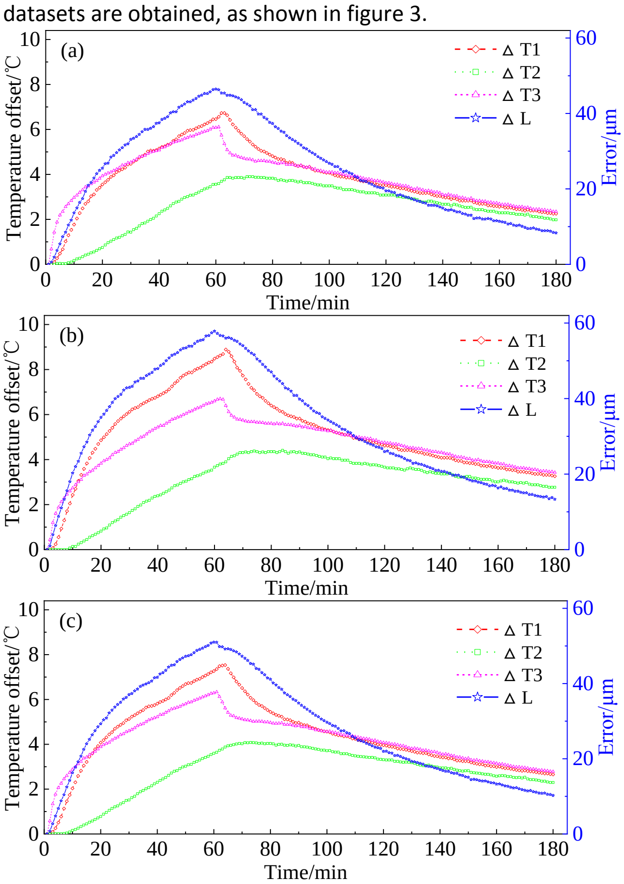


Figure 2. Experimental setup for the thermal error experiment

In this test, the spindle rotates without load at the speed of 6000r/min for 1h and then remains stop for 2h. The temperature offset to the ambient temperature and thermal error data are recorded every 1 min. The test is repeated four times, and four datasets are obtained, as shown in figure 3.



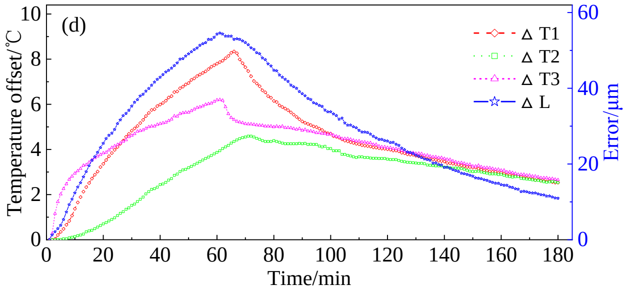


Figure 3. Measured datasets of the temperature and thermal error

4. Results and Discussion

Three measured datasets are randomly selected as the training dataset (e.g. figure 3a, 3b, and 3c) and the remaining one (e.g. figure 3d) is taken as the testing dataset. To compare with the conventional machine learning techniques, the thermal error model is also built using back-propagation neural network (BPNN) with a single hidden layer of 100 neurons. The predicted and observed values using RF and BPNN are shown in figure 4.

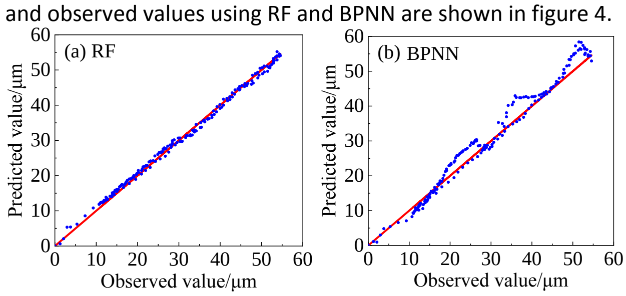


Figure 4. Observed and predicted values of errors using RF and BPNN

The model accuracy using RF and BPNN are measured through the coefficient of determination R^2, mean absolute error (MAE) and mean square error (MSE) in statistics, as listed in table 1.

Table 1 Comparison of the model accuracy using RF and BPNN

Model	R^2	MAE/μm	MSE/μm^2
RF	0.996	0.817	0.843
BPNN	0.982	1.991	6.803

It can notably be seen from Table 1, RF can lead to higher prediction accuracy than BPNN. Furthermore, for the different rotate speed and stop time, the proposed model can still maintain the prediction accuracy of over 90%. In addition, aggregating the prediction of all these diverse trees can significantly eliminate the influence of the noise and reduce the overall variance. Therefore, RF has stronger robustness.

5. Conclusions

To overcome the limitation of the accuracy and robustness of the current modeling methods, this paper presents a novel thermal error modeling method using random forest, which can achieve higher accuracy and has stronger robustness. In future work, random forest will also be applied to other errors of the machine tools such as cutting-tool wear-induced error.

Acknowledgements

This work is financially supported by the National Key R&D Program of China (Grant No. 2018YFB1701204).

References

[1] Yin Q, Tan F, Chen H and Yin G 2018 Spindle thermal error modeling based on selective ensemble BP neural networks *Int. J. Adv. Manuf. Technol.* **101** 1699-713

[2] Wu D, Jennings C, Terpenny J, Kumara S and Gao R X 2018 Cloud-Based Parallel Machine Learning for Tool Wear Prediction *ASME J. Manuf. Sci. Eng.* **140** 4