Special Interest Group meeting
Thermal Issues
Eindhoven, The Netherlands, March 2024
www.euspen.eu

Machine tool thermal state classification using LSTM models

N. Ariaga[1], A. Longstaff[1], S. Fletcher[1]

[1]*Centre for Precision Technologies, University of Huddersfield, Huddersfield, UK*

N.Ariaga2@hud.ac.uk

Abstract
This paper presents an approach for classifying the machining operation conditions of a machine tool using temperature measurements from key points of the machine tool structure. Classifying data and measuring the similarity of datasets can improve the compensation of thermal error by enabling empirical models to adapt to changes in machining conditions. This would extend the long-term use of the models beyond the conditions used to train the models. In this presentation, Proper Orthogonal Decomposition is used to extract features from the temperature data. The features are used by a long short-term memory model to classify the data into three classes according to the spindle speed used when recording the data. The presented approach achieves high prediction accuracy and can be used to measure similarities in datasets.

Thermal error; Neural network

1. Introduction

Thermal errors are a major source of errors in machine tools with estimates attributing 50% of waste production in machined components [1] to thermal errors. Data driven models be conveniently used in the compensation of thermal errors. Such models learn to predict the thermal error from inputs which are mostly temperature measurements of key points on the machine tool. However, the accuracy of the predictions drops when the machining conditions (e.g. spindle speeds, feed rates, and depth of cut used) differ from those used to obtain the training data. Moreover, obtaining training data is time consuming and expensive. Empirical thermal error models are thus constrained to learn from a limited number of examples of temperature measurements and the corresponding thermal error at the tool centre point (TCP). Therefore, long term use of empirical thermal error models requires strategies for handling conditions which were not part of the training data.

One such strategy is performing periodic model updates [2] to continuously update the model regardless of whether the machining conditions have changed or not. Another strategy involves intermittent model updates where a method of determining changes in the machining conditions is used to trigger model updates [3]. Another strategy involves training multiple models for different conditions and switching between the models depending on the prevailing conditions [4]. These methods lack a means of quantifying the amount of change or similarity in the machining conditions from the data. Quantifying this change would avoid unnecessary updates in the first strategy. It would also enable reuse of previously seen data that may is similar to the prevailing conditions when retraining models in the second strategy. This would improve efficiency by reducing process intermittent probing needed to obtain new training data. It could also guide the determination of the number of multiple models to train in the model switching strategy instead of using arbitrary number of models.

This presentation builds upon previous studies [5, 6] on finding a measure for the changes in the machining conditions using temperature data from key points of the machine tool. The studies used Proper Orthogonal Decomposition (POD) to extract features (POD modes) representing the direction of change in the machine tool's thermal state over a period. These features could then be classified using K-Means clustering and a distance metric such as the cosine distance [5]. Hidden Markov Models with Gaussian Mixture Model emissions (GMM-HMM) can also be used to measure the similarity of the extracted features [6]. GMM-HMM models rely on the Markov assumption which states that the current hidden state is only affected by the immediate past hidden state. The hidden state in this case is the prevailing machining condition. Challenges such as the presence of noise in the extracted features may require the inference window to extend back in time beyond the Markov assumption. This present work seeks to extend the inference window using the Long Short-Term Memory (LSTM) recurrent neural network model. The next section introduces LSTM models. This is followed by a presentation of the methodology used in this work. A discussion of the results obtained using the approach conditions is then presented. Finally, concluding remarks from the study are presented.

2. Long Short-Term Memory (LSTM) models

Hochreiter and Schmidhuber [7] introduced LSTMs to overcome the challenge of using gradient based learning with recurrent neural networks. LSTM models prevent the back propagated error from shrinking or exploding by using a series of three gates to control the flow of information through the network as shown in ***Figure 1***. The first gate controls how much past information is retained in the hidden state. The second gate controls how much new information from the instantaneous input ($\boldsymbol{x_t}$) is added to the hidden state. While the third gate controls what output ($\boldsymbol{h_t}$) is obtained from the hidden state. Sigmoid functions (σ) and hyperbolic tangent (tanh) functions are used at various points as activation functions.

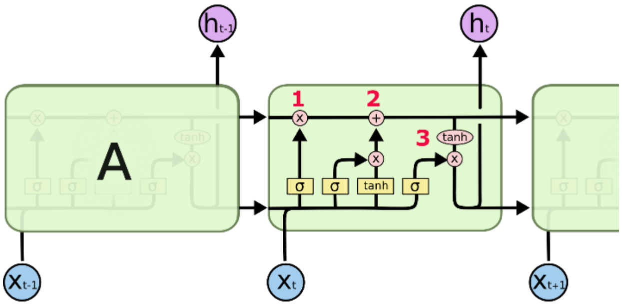

Figure 1. Structure of a LSTM neural network model

3. Methodology

Nine air cutting experiments were performed on a three-axis vertical machine tool. Data from these experiments was used to determine how effective an LSTM model could classify temperature data into three classes of machining conditions: low spindle speed (4,500 rpm), medium spindle speed (6,000 rpm), and high spindle speed (8,000 rpm). A fixed feed rate of 15,000 mm/min was used but the spindle speed for each experiment was selected from a list of spindle speeds (4,500 rpm, 5,000 rpm, 6,000 rpm, 8,000 rpm, and 9,000 rpm). The temperature of the machine tool was measured at 10 Hz resulting in datasets labelled A to H. Training data was obtained from datasets A, F, and H to represent the low, medium, and high spindle speeds respectively. POD analysis was performed on the temperature data by sliding a window containing 20 timestamps to obtain features (POD modes) for the data within the window [6]. The LSTM model took a sequence of 50 past features and predicted the probability of conditions represented by each of the three classes occurring at the end of the 50 features. The LSTM model used in this study had a sequence of eight layers: a sequence input layer, two pairs of an LSTM layer followed by a dropout layer, a fully connected layer, a softmax layer, and a classification output layer. The next section presents the results from testing the model on datasets A through H.

4. Results and discussion

A summary of the model's predictions is shown in **Table *1***.

Table 1. LSTM predictions of the number of positive predictions (P+), negative predictions (P-) and percentage of positive predictions (P+ %).

	Predicted class								
	Low speed			Medium speed			High speed		
Dataset	P+	P-	P+ %	P+	P-	P+ %	P+	P-	P+ %
A	6323	42	99	42	6323	1	0	6365	0
B	3155	1564	67	1497	3222	32	67	4652	1
C	2207	2514	47	2424	2297	51	90	4631	2
D	1217	3499	26	3479	1237	74	20	4696	0
E	1251	6537	16	6371	1417	82	166	7622	2
F	111	6976	2	6873	214	97	103	6984	1
G	58	4675	1	86	4647	2	4589	144	97
H	26	11814	0	0	11840	0	11814	26	100
I	1	9463	0	39	9425	0	9424	40	100

Spindle speed (rpm)				
4,500	5,000	6,000	8,000	9,000
A, B	C, D, E	F	G	H, I

The model correctly classified over 74% of the samples in datasets D and E as medium spindle speed data. Over 97% of the samples from datasets G and I were also correctly classified as high spindle speed data. The model learnt patterns in the training data that enabled these predictions. The accuracy of the model was high for classes separated by a significant spindle speed such as the medium spindle speed and high spindle speed. The accuracy of the model decreased when classes had close spindle speed values such as the low and medium spindle speed classes as seen in the results for datasets B and C whose spindle speeds differ by 500 rpm. This observation can be ideal when determining if data from different machining conditions is similar. In this case data obtained from close spindle speeds has high similarity resulting in the low prediction accuracy.

The probability scores from the LSTM model for dataset D are shown **Figure 2**. The model had difficulties predicting samples regions when the machine tool is cooling down such as the period before the eighth hour. Such periods can be detected using the singular values from POD analysis [6] and appropriate action taken to improve the model's accuracy.

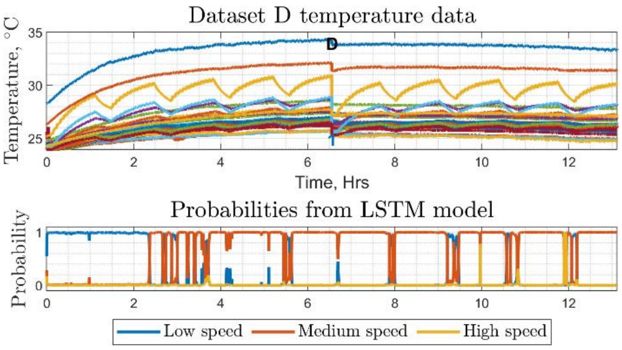

Figure 2. Probability scores for classifying dataset H using LSTM model.

5. Conclusion

A method for classifying the machining conditions from temperature data using POD analysis and LSTM modelling was presented. The LSTM model gave a probability score of whether temperature data belonged to a set number of classes. These probability scores can act as similarity measures and incorporated in the modelling of thermal errors in machine tools. Further work is being carried out to improve the model's accuracy and incorporate the model's output in the modelling of thermal errors in machine tools.

References

[1] M. Putz, C. Richter, and J. Regel, "Industrial relevance and causes of thermal issues in machine tools," *Proc Conf Therm Issues Mach tools*, no. June, pp. 127–139, 2018.

[2] H. Yang and J. Ni, "Adaptive model estimation of machine-tool thermal errors based on recursive dynamic modeling strategy," *Int J Mach Tools Manuf*, vol. **45**, no. 1, pp. 1–11, Jan. 2005, doi: 10.1016/j.ijmachtools.2004.06.023.

[3] N. Zimmermann, S. Lang, P. Blaser, and J. Mayr, "Adaptive input selection for thermal error compensation models," *CIRP Ann*, vol. **69**, no. 1, pp. 485–488, 2020, doi: 10.1016/j.cirp.2020.03.017.

[4] Y. Liu, E. Miao, H. Liu, and Y. Chen, "Robust machine tool thermal error compensation modelling based on temperature-sensitive interval segmentation modelling technology," *Int J Adv Manuf Technol*, vol. **106**, no. 1–2, pp. 655–669, 2020, doi: 10.1007/s00170-019-04482-8.

[5] N. Ariaga, A. P. Longstaff, S. Fletcher, and W. Pan, "Machine tool thermal state representation using modal analysis," in *14th International Conference and Exhibition on Laser Metrology, Coordinate Measuring Machine and Machine Tool Performance*, Dec. 2021, pp. 25–34.

[6] N. Ariaga, A. Longstaff, and S. Fletcher, "Characterising the thermal state of machine tools using modal analysis and clustering," *Laser Metrol Mach Perform XV - 15th Int Conf Exhib Laser Metrol Mach Tool, C Robot Performance, LAMDAMAP 2023*, pp. 71–80, 2023.

[7] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*, vol. **9**, no. 8, pp. 1735–1780, Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.