Characterising the thermal state of machine tools using modal analysis and clustering

N. Ariaga, A. Longstaff, S. Fletcher
University of Huddersfield, UK

Abstract

Compensating thermal errors using predictions from temperature-based empirical thermal error models is a widely used and convenient approach for mitigating the effects of the errors. However, thermal error models are often trained on small datasets that are not representative of all situations that the machine, and so model, may encounter during online use. Various adaptive strategies have been proposed to help models overcome this problem, such as intermittent model update strategies. One challenge that remains open is how such strategies can be automated.

This work proposes the use of modal analysis through Proper Orthogonal Decomposition (POD) and Hidden Markov Models with Gaussian Mixture Model emissions (GMM-HMM) in characterizing the thermal state of a machine tool. Performing a windowed POD analysis on the temperature data results in POD modes that define the heat cycles experienced during the windowed period. GMM-HMM are then used to cluster the POD modes. The approach correctly classifies various temperature data from all but one of the test datasets with True Positive Rate (TPR) value of over 61%. The presented approach can achieve higher accuracies through implementing the discussed improvements and be incorporated into the thermal error modelling strategy to inform adaptive modelling strategies.

1 Introduction

Reducing the effects of thermal errors on the precision of machine tools is an important aspect in ensuring that machining can achieve design tolerances and avoid losses incurred due to part rejection. It is estimated that as much as 50% of waste production can be attributed to thermal errors [1]. Compensating these errors using predictions from empirical thermal error models is a widely used and convenient approach for mitigating the effects of temperature change. These models are primarily trained using temperature measurement data obtained from key points of the machine tools as inputs. The models learn a mapping function from these inputs to the axis-specific thermal errors. However, obtaining training data is expensive in time and cost since machine tools are required to be taken out of production to measure the thermal error from various heating experiments. Thus, thermal error models are often trained on small datasets that are not representative of all situations the machine, and so model, may encounter during online use [2-4]. Various adaptive strategies have been proposed to help models overcome this problem such as intermittent model update strategies.

Ramesh et al [3] used a Support Vector Machine (SVM) model to classify features extracted from temperature measurements into one of three groups representing different sets of machining conditions (parameters). An SVM thermal error model was also trained for each of the classes. The features used were obtained through study of one particular machine tool. The authors then used the appropriate thermal error model based on the classification results. Ramesh et al [2] used a Bayesian Network (BN) model and a rule-based system to classify temperature measurements into one of three classes representing different sets of machining conditions. Compensation under each of the conditions was achieved using an SVM thermal error model for that condition. Creation of the BN model and rule-based system required a good understanding of the machine tool and the relationship of temperature sensor locations on it. Blaser et al [5] used periodic thermal error measurements to determine if the residuals from a thermal error model were higher than a defined tolerance band and trigger a cycle to update the thermal error model. However, the periodic thermal error measurements result in a drop in productivity since they interrupt the machining process. Zimmermann et al [4] used SVM models to detect anomalous temperature measurements and trigger update cycles needed by the thermal error model to adapt to changes in machining conditions. An SVM model was developed for each temperature sensor that had been selected as an input to the thermal error model. Each SVM model had two inputs: the difference in the instantaneous measurement from a reference measurement at the start of the experiment, and the rate of change of the measurement. The SVM models were then used to score temperature measurements as being in or out of class. The approach resulted in a 78% reduction in the number of triggered updates when compared to a periodic model update approach. This translates to an increase in productivity.

Though updating the thermal error model is beneficial in adapting to changing machining conditions, there is still a need for retaining models which are adapted for certain machining conditions, since those conditions could be experienced again depending on the nature of production. This could increase productivity by reducing the need for retraining or relearning mappings for such conditions. One approach of achieving this could be to discretise the possible range of machining parameter values (spindle speeds, feed rates, dept of cuts) and train a model for each combination of discrete values. However, the number of combinations can lead to a large number of thermal error models which would reduce the efficiency by increasing the complexity of training and maintaining the models. Brecher and Wissmann [6] discretised the spindle speed at six points and the effective power of the main spindle at four points. Two thermal error models were then developed for each combination of points (one for the heating cycle and one for the cooling cycle). The thermal error compensation was performed using predictions from one of the resulting forty-eight models based on the value of spindle speed and effective spindle power obtained from the numerical controller. One way of avoiding the combinatorial explosion of models is establishing a method to determine how similar any two sets of machining conditions are, based on temperature measurements from key points of the machine tool. This would be beneficial in implementing adaptive and automated thermal error modelling strategies.

Temperature measurement data from machine tools is usually characterised by high dimensionality and correlations that vary in time depending on the prevailing heating cycles. Ariaga et al [7] used Proper Orthogonal Decomposition (POD) analysis on a sliding window of data to obtain POD modes. POD modes in this application were synonymous to vectors in the space spurned by the temperature measurements, “temperature space” and pointed in directions that depended on the heating cycles within the sliding window. K-means clustering was then used to determine centroid POD modes (descriptor POD modes) which could represent a range of similar POD modes. Consequently, the temperature data being evaluated was characterised by a set of descriptor POD modes. The cosine distance metric was then used to determine how close these descriptor POD modes were to any POD modes extracted from a different temperature dataset. This paper builds on the approach presented in [7] by first using modal analysis to characterise the thermal state of machine tool. Hidden Markov Models with Gaussian Mixture Model emissions (GMM-HMM) are then used to model the sequential nature of the POD modes and obtain a probabilistic similarity measure.

2 Methodology

The process of characterising a dataset of temperature data and comparing whether two temperature datasets were generated under the same machining conditions is summarised in Figure. This section introduces the POD and GMM-HMM modelling as used in this paper.

2.1 Proper Orthogonal Decomposition (POD)

A window of fixed width, m sequential samples, is slid across the dataset of temperature data obtained from n temperature sensors with respect to time and POD is performed on the data within the window. This approach is used because it is adaptable to online use and only factors most recent temperature measurements at any evaluation instant. POD is performed by obtaining the Singular Value Decomposition (SVD) of the mean subtracted data as summarised in the following equation

$$\mathbf{X} = \mathbf{U}\,\mathbf{\Sigma}\,\mathbf{V}^{\mathbf{T}} \qquad (1)$$

Where $\mathbf{X} \in \mathbf{R}^{\mathbf{n\,x\,m}}$ is the mean subtracted data from n temperature sensors at m sampling instances. $\mathbf{U}$, $\mathbf{\Sigma}$, and $\mathbf{V}$ are the outputs from SVD where $\mathbf{U} \in \mathbf{R}^{\mathbf{n\,x\,n}}$ are left singular vectors, $\mathbf{\Sigma} \in \mathbf{R}^{\mathbf{n\,x\,n}}$ is a diagonal matrix of singular values, $\mathbf{V} \in \mathbf{R}^{\mathbf{n\,x\,m}}$ are the right singular vectors. Both the left and right singular vectors are orthogonal column vectors. The left singular values represent the correlations of the temperature sensors and form the POD modes of the data while the right singular vectors represent the time specific weightings needed to linearly combine the POD modes. The singular values are ordered in decreasing absolute values where each singular value represents how much dispersion of the data occurs along the direction of the corresponding POD mode. Most of the dispersion in the data is captured by the leading POD modes which forms a basis for the use of POD in data dimensionality reduction applications [8].

In this paper, only the first POD mode is used for clustering application and is shown to account for most of the dispersion. Using the first mode also reduces the complexity of modelling the thermal state. The first POD mode at each evaluation is also scaled by its singular value to enable differentiation of similar heating patterns that result in different rates of heating. POD modes from such data will point in nearly the same direction in temperature space but their singular values will have significant differences.

2.2 Hidden Markov Models with Gaussian Mixture Model emissions (GMM-HMM)

Certain dynamic systems possess a hidden state which governs their observable behaviour. An example of this is the health status of an animal which cannot be measured, except through measuring observable symptoms such as behaviour patterns (eating, sleeping, and playing) or body temperature (cold, normal, and hot). The hidden and observable states of such systems can take on both discrete values or continuous values such as temperature measurements. Such a system is said to obey the Markov property if the current value of its hidden state is a probability function of its value at the previous time instant and independent of all other previous values and if the current value of the observable state is a probability function of the current value of the hidden state and independent of all previous values of both the hidden and observable states [9]. Such systems are also called Markov processes. Hidden Markov Models (HMM) are used to model Markov processes whose hidden and observable states take on discrete values such as DNA sequence data which is made up of a sequence of discrete protein members [10].

Figure 1 shows the health status of a person visualised as a Markov process. The hidden state takes on two discrete values (HEALTHY or SICK) while the observable state is the measured body temperature which is a continuous value. The range of body temperature values is represented by three discrete state values (COLD, NORMAL, and HOT) using three Gaussian models. This enables the Markov process to be represented by a HMM model. The parameters of the HMM model include the probability of the hidden state transitioning to a given value from its current value. For the example in Figure 1, a person who is currently “HEALTHY” will remain “HEALTHY” with a probability of 0.6 or will become “SICK” with a probability of 0.4 in the next time instant. The HMM is also parametrised by the probability of observing a certain state value for a given hidden state value. For example in Figure 1, a person who is currently “SICK” will be observed to have a “NORMAL” body temperature with a probability of 0.2.

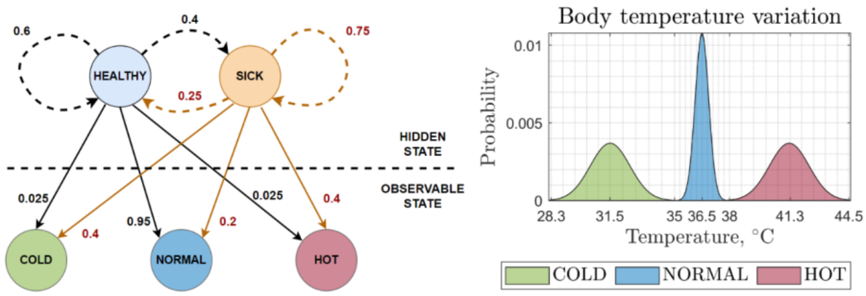

Figure 1: Example of a GMM-HMM model of the health status of a person

The thermal state of the machine tool can also be represented as a Markov process since the temperature change patterns are governed by the (hidden) machining operation being carried out. POD modes indicate the direction of the temperature change in a continuous temperature space and form the observable state values of the system. Gaussian Mixture Models (GMM) can be used to model continuous observable state values with multiple dimensions such as POD modes. Therefore, the thermal state of a machine tool can be modelled using Hidden Markov Models with Gaussian Mixture Model emissions (GMM-HMM) where the hidden state, machining conditions, take on discrete values while the observable states are continuous. As seen in the example in Figure 1, a range of temperature values were associated with a discrete observable state such as NORMAL (body temperature). Similarly, a range of POD mode values modelled by a GMM is associated with a discrete observable state. The number of discrete values which the hidden and observable states can take need to be specified when training HMM models. The parameters of the HMM model can then be estimated from a set of observations using algorithms such as the Baum-Welch algorithm [9]. The probability that a given sequence of observations was generated by a Markov process represented by the trained GMM-HMM model can also be obtained using algorithms such as the “forward algorithm” and the “backward algorithm” whose details can be found in [9]. All computations presented in this paper were implemented using MATLAB R2021a as well as using functions contained in an open source HMM library [11].

The true positive rate (TPR) metric is used to determine the accuracy of GMM-HMM models in classifying temperature data [12]. Each GMM-HMM model was considered as a binary classifier that predicted whether a series of data was generated by the Markov process represented by the model. Thus, a positive prediction (**P+**) indicated that the data belonged to the class while a negative prediction (**P-**) indicated that the data did not belong to the class. The TPR value, presented as a percentage, was obtained by the following equation

$$\mathbf{TPR} = \frac{(\mathbf{P}+)}{(\mathbf{P}\,+) + (\mathbf{P}-)} \qquad (2)$$

3 Experiment setup

The proposed approach for characterising the temperature data of a machine tool was tested using experiment data obtained from a three-axis vertical machining centre (VMC). The thermal state of the VMC was monitored using thirty-eight temperature sensors positioned to measure the temperature and the temperature gradient of key points of the VMC. This was done at a sampling frequency of 10 Hz. The positioning of the sensors was based on engineering knowledge and research studies [13], [14]. In each run of the experiment, the VMC was repeatedly subjected to the following three different heating cycles using a different spindle speed and axis feed rate:

- Cycle 01. Air cutting machining cycle which consisted of motion of the spindle and all axes
- Cycle 02. X-axis heating cycle
- Cycle 03. Spindle heating cycle

All runs of the experiment included Cycle 01 while Cycle 02 and Cycle 03 featured in some of the runs. However, only data from Cycle 01 is analysed in this paper. Thermally-induced errors were also measured using probing spheres though this data was not used in this presentation.

3 Results and discussion

Nine experiment runs (labelled A to I) were performed on different days covering a span of two months. Cycle 01 was performed using a feed rate of 15,000 mm/min and varying spindle speeds (4,500 rpm, 5,000 rpm, 6,000 rpm, 8,000 rpm, and 9,000 rpm). Cycle 02 and Cycle 03 were also performed using different machining parameters but the analysis in this paper on data from Cycle 01. The temperature data obtained from dataset H is shown in Figure 2.

Data from Cycle 01 of datasets A, F, and H were selected as training data for the GMM-HMM models. POD analysis was performed using a sliding window equivalent to 3.3 minutes (20 samples of data). A one-class GMM-HMM model (containing one hidden state) was then trained on the data for each of the training datasets. All the models were structured to have four observable states.

Classification was performed by passing a sequence of POD modes contained in a sliding window to the GMM-HMM models which evaluated the probability that the sequence was generated by the Markov process represented by the models. The data was then assigned the class of the model with the highest probability. The TPR metric was then used to evaluate the accuracy of the classification results.

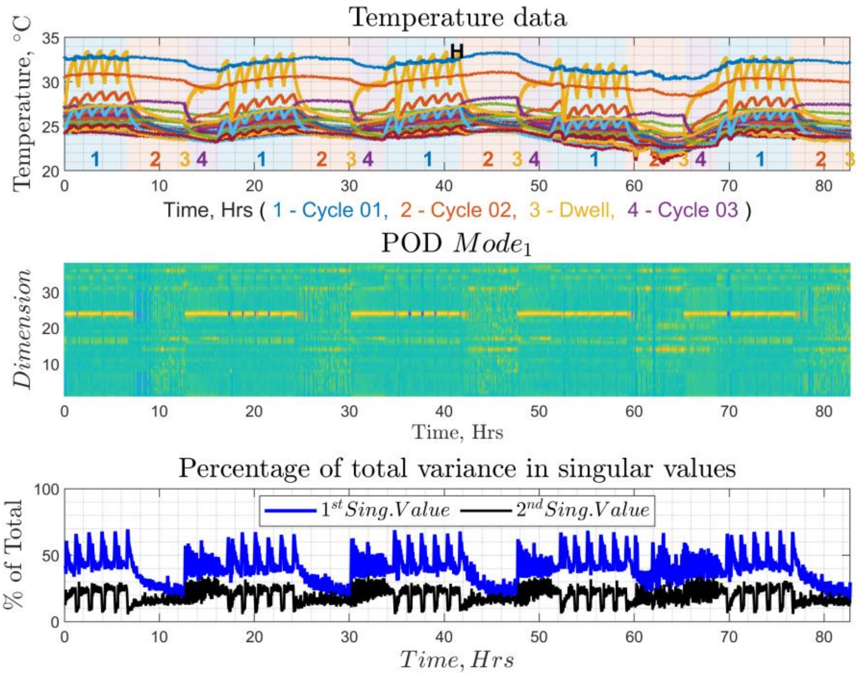

Figure 2: Temperature data and POD analysis results of dataset H

The results of performing the POD analysis on dataset H are shown in Figure 2. The POD modes are displayed as a surface plot where the Z-axis represents the range of values each dimension of the POD mode can have (from -1 to 1). Sustained heating or cooling cycles appear as homogeneous sections in the POD modes' plot. The singular values for first and second POD mode are also shown in Figure 2 as a percentage of the total sum of singular values for each time step. It is observed that the singular value for the first POD mode accounts for over 40% of the dispersion in the data which is slightly over double that of the second singular value. The exception to this is Cycle 02, during which the total dispersion in the temperature data is relatively low. The rest of the singular values accounted for less dispersion than the second singular values and are not shown.

The results obtained from using the trained GMM-HMM models to classify Cycle 01 data from test datasets are summarised in Table 1. Sections of data from the test datasets are shown in Figure 3. The GMM-HMM models trained on datasets A, F, and H were selected to identify one of three ranges of spindle speeds: low spindle speeds (4,500 rpm), medium spindle speed (5,000 to 6,000 rpm), and high spindle speed (8,000 to 9,000 rpm). The results in Table 1 show that the models correctly classified the test datasets with a TPR values of over 61% except for dataset C. The mean value of the Gaussian mixture values learnt by the GMM-HMM models are shown in Figure 4.

Table 1: Summary of results from classifying data using GMM-HMM models

	Model								
	gmm-hmm A1			gmm-hmm F1			gmm-hmm H1		
Dataset	P+	P-	TPR	P+	P-	TPR	P+	P-	TPR
A	5,864	501	92	494	5,871	8	7	6,358	0
B	3,596	1,123	76	1,102	3,617	23	21	4,698	0
C	2,896	1,825	61	1,664	3,057	35	161	4,560	3
D	1,677	3,039	36	2,911	1,805	62	128	4,588	3
E	2,898	4,890	37	4,774	3,014	61	116	7,672	1
F	1,007	6,080	14	5,848	1,239	83	232	6,855	3
G	305	4,428	6	585	4,148	12	3,843	890	81
H	248	11,592	2	556	11,284	5	11,036	804	93
I	412	9,052	4	528	8,936	6	8,524	940	90

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

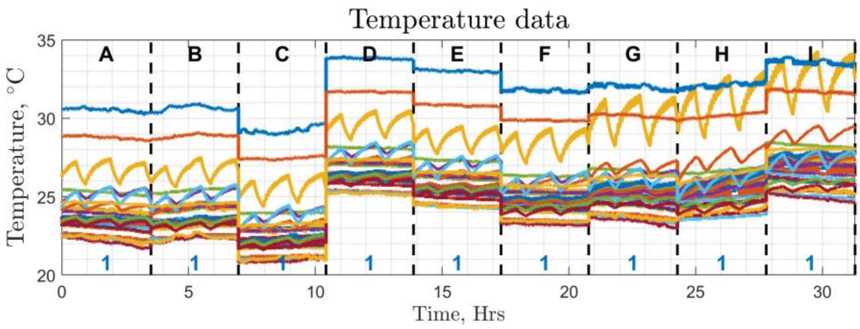

Figure 3: Section of data from Cycle 01 of the test datasets

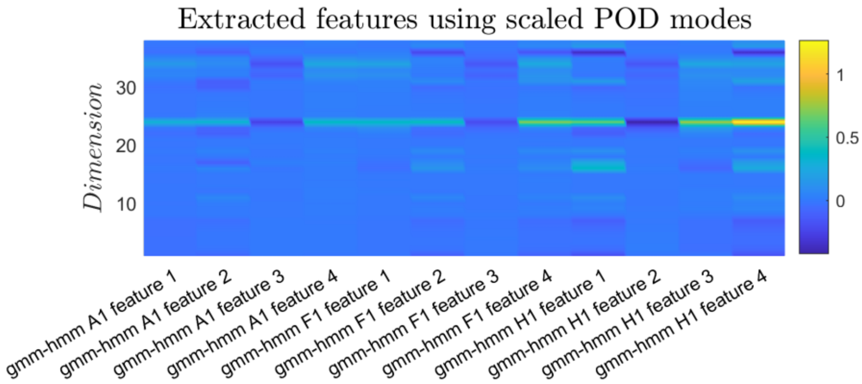

Figure 4: Features learnt by the GMM-HMM models

Some of the features from different models appear indistinguishable. For example, "gmm-hmm A1 feature 3" and "gmm-hmm F1 feature 3" which represent cooling cycles are similar in appearance. The exact value of how close these features are can be obtained using distance metrics such as variants of the cosine distance and the Euclidean distance. Thus, the results could be improved further by incorporating application specific modifications. In this case, it can be argued that features associated with heating cycles are more informative than those associated with cooling cycles and should be weighted appropriately when predicting the class.

4 Conclusion

A method of characterising the thermal state of a machine tool using modal analysis and Hidden Markov Models with Gaussian Mixture Model emissions (GMM-HMM) was presented. Temperature datasets measured at key points of a VMC machine tool were recorded under heating cycles with varying spindle speeds and feed rates. Proper Orthogonal Decomposition (POD) was used to extract features from the temperature data. Three GMM-HMM models were trained to learn the sequence of these features and enable classification. The models represented classes of machining conditions where the feed rate was held constant, but the spindle speed was discretised into three values: low (4,500 rpm), medium (6,000 rpm), and high (9,000 rpm) spindle speeds.

The GMM-HMM models had True Positive Rate (TPR) values of over 60% in accurately predicting the class of all but one of the test datasets. It was observed that some learnt features were relatively less informative in the class assignment since they were similar in appearance those in other models. An example of this were features associated with cooling cycles. Further research is being carried out to improve the accuracy and incorporate the presented methods in the thermal error modelling workflow.

Acknowledgement

The authors gratefully acknowledge the UK's Engineering and Physical Sciences Research Council (EPSRC) funding of the Future Metrology Hub (Grant Ref: EP/P006930/1) and UKRI-funded Advanced Machinery and Productivity Initiative (84646).

References

[1] M. Putz, C. Richter, J. Regel, and M. Bräunig, "Industrial consideration of thermal issues in machine tools," Prod. Eng., vol. 12, no. 6, pp. 723–736, 2018

[2] R. Ramesh, M. A. Mannan, A. N. Poo, and S. S. Keerthi, "Thermal error measurement and modelling in machine tools. Part II. Hybrid Bayesian Network - Support vector machine model," Int. J. Mach. Tools Manuf., vol. 43, no. 4, pp. 405–419, 2003

[3] R. Ramesh, M. A. Mannan, and A. N. Poo, "Support vector machines model for classification of thermal error in machine tools," Int. J. Adv. Manuf. Technol., vol. 20, no. 2, pp. 114–120, 2002

[4] N. Zimmermann, M. Breu, J. Mayr, and K. Wegener, "Autonomously triggered model updates for self-learning thermal error compensation," CIRP Ann., vol. 70, no. 1, pp. 431–434

[5] P. Blaser, F. Pavliček, K. Mori, J. Mayr, S. Weikert, and K. Wegener, "Adaptive learning control for thermal error compensation of 5-axis machine tools," J. Manuf. Syst., vol. 44, pp. 302–309, 2017

[6] C. Brecher and A. Wissmann, "Compensation of thermo-dependent machine tool deformations due to spindle load: Investigation of the optimal transfer function in consideration of rough machining," Prod. Eng., vol. 5, no. 5, pp. 565–574, 2011

[7] 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, 2021.

[8] I. T. Jollife and J. Cadima, "Principal component analysis: A review and recent developments," Philos. Trans. R. Soc. A Math. Phys. Eng. Sci., vol. 374, no. 2065, 2016

[9] J. P. (Mathematician) Coelho, T. M. Pinho, and J. Boaventura-Cunha, "Hidden Markov models: theory and implementation using Matlab." CRC Press, Boca Raton, FL, pp. 29–37, 2019

[10] L. E. . Baum and T. Petrie, "Statistical Inference for Probabilistic Functions of Finite State Markov Chains," Ann. Math. Stat., vol. 37, no. 6, pp. 1554–1563, 1966

[11] Q. Kong, "Hidden Markov Model Toolbox." GitHub.com, 2015. [Online]. Available: https://github.com/qiuqiangkong/matlab-hmm

[12] D. M. W. Powers, "Evaluation : From Precision , Recall and F-Factor to ROC , Informedness , Markedness & Correlation," no. December, 2007

[13] A. M. Abdulshahed, A. P. Longstaff, S. Fletcher, and A. Myers, "Thermal error modelling of machine tools based on ANFIS with fuzzy c-means clustering using a thermal imaging camera," Appl. Math. Model., vol. 39, no. 7, pp. 1837–1852, 2015

[14] S. Fletcher, "Computer aided system for intelligent implementation of machine tool error reduction methodologies," University of Huddersfield, 2001.