

Received: 31 March 2023 / Accepted: 22 June 2023 / Published online: 26 June 2023

*machine tool, thermal error compensation,
machine learning, artificial neural network*

Christian BRECHER¹,
Mathias DEHN^{1*},
Stephan NEUS¹

AN INVESTIGATION OF THE RELATIONSHIP BETWEEN ENCODER DIFFERENCE AND THERMO-ELASTIC MACHINE TOOL DEFORMATION

New approaches, using machine learning to model the thermo-elastic machine tool error, often rely on machine internal data, like axis speed or axis position as input data, which have a delayed relation to the thermo-elastic error. Since there is no direct relation to the thermo-elastic error, this can lead to an increased computation inaccuracy of the model or the need for expensive sensor equipment for additional input data. The encoder difference is easy to obtain and has a direct relationship with the thermo-elastic error and therefore has a high potential to improve the accuracy thermo-elastic error models. This paper first investigates causes of the encoder difference and its relationship with the thermo-elastic error. Afterwards, the model is presented, which uses the encoder difference to compute the thermo-elastic error. Due to the complexity of the relationship, it is necessary, to use a machine learning approach for this. To conclude, the potential of the encoder difference as an input of the model is evaluated.

1. INTRODUCTION

There is an ongoing trend in the machine tool industry for more productive machine tools and higher workpiece quality, while reducing the environmental machine footprint [1]. In order to comply to the increasing demand for high accuracy machine tools, manufacturers are driven to decrease any error influences impacting the machining accuracy. One of the largest errors influencing the machine tool accuracy is the thermo-elastic error. When a machine is in operation, internal and external heat sources lead to a transient heat flux in the machine tool structure, which results in a deformation of the structure and a displacement of the Tool Center Point (TCP) and thus a machining error. To investigate the machine tool behaviour, trials are conducted using heat pads as an external heat source with a controlled heat flux position, axis motion as an internal heat source resulting in a heat flux similar to the real working conditions of a machine tool, and daily changes in ambient temperature.

¹ Chair of Machine Tools, Laboratory for Machine Tools and Production Engineering (WZL) of RWTH Aachen University, Germany

* E-mail: m.dehn@wzl.rwth-aachen.de

<https://doi.org/10.36897/jme/168701>

2. THERMO-ELASTIC ERROR

A lot of research has already been conducted on thermo-elastic TCP errors. An overview of recent research is presented by Gao et al. [2]. The research can be clustered into approaches focussing on compensation and correction of the thermo-elastic error. Compensation approaches target the cause of the error by reducing the heat flux in the machine tool structure. This can be done by cooling the structural components [3, 4]. Compensation approaches show reliable results and have proven to reduce the thermo-elastic error effectively. However, they are energy intensive, which is not compatible with the effort to reduce resource and energy consumption. Error correction methods use models to compute the thermo-elastic error at the TCP and correct it using the machine control, which does not need additional energy. The correction approaches used can be clustered into white box [5], gray box [6] and black box [7–10] models with each having its own advantages and disadvantages in specific use cases. Apart from the model type used, all models need input data. The type of input data needed has a direct effect on the usability of the model. Therefore, the input data needs to be easy to obtain (availability), reliable to measure (reliability) and there should be a strong dependency between the data and the thermo-elastic error (correlation). An overview of commonly used input data and an evaluation of the three given criteria is shown in Table 1.

Table 1. Typical input data for thermo-elastic error models

data	axis speed	axis position	motor current	structural temp.	cooling fluid temp.	ambient temp.
availability	++	++	++	o	+	-
reliability	++	++	++	+	+	o
correlation	o	-	o	+	+	o

Machine internal data like the axis speed or position is easily available and very reliable. However, there is only a delayed dependency to the thermo-elastic error, thus reducing the usability in a model [11, 12]. Data from external sensors like temperature sensors have a good and direct dependency to the thermo-elastic error, but they are prone to errors and have a higher effort regarding installation and maintenance [9]. In the literature, there is no input data used, which is good in all categories. Previous studies have found the encoder difference to constitute such a value with high availability, reliability and correlation to the thermo-elastic error. This encoder difference therefore has the potential to increase the applicability of a model in practical use drastically. The sources of the encoder difference and its dependency to the thermo-elastic TCP error are described in the following section.

3. ENCODER DIFFERENCE

Typically, there are two encoders in a machine tool linear axis: A linear encoder measuring the position directly at the axis and a rotary encoder measuring the rotary position of the axis motor, thus measuring the axis position indirectly. The axis position is therefore

measured redundantly, with both measurements being available in the machine control. The difference between these two measurements is called encoder difference. Current research investigating the encoder difference is very limited and in general not aimed at establishing a dependency to the thermo-elastic error.

The company Heidenhain investigates, how the encoder difference develops over time when the machine is under constant load with an average movement speed of 10 m/min at three machine poses for a fixed / floating ball screw configuration. The encoder difference was only measured during stand still of the axis. They find that the encoder difference increases over time. As expected, the encoder difference increases more for positions further from the fixed bearing than for positions close to the fixed bearing [13].

Wang et al. reach similar results in their investigation [14]. An analysis of the encoder difference depending on the axis position and the time shows a dependence between the encoder difference and the thermo-elastic machine state. Furthermore, a small periodic error can be seen which is repeated for every turn of the ball screw. Xi et al. investigate the behaviour of the encoder difference depending on the process forces on the TCP [15]. The experiments show a linear behaviour between the force measured at the TCP and the encoder difference, with a stiffness of roughly 175 N/ μm .

Figure 1 (left) shows further influences on the encoder difference. Production errors in the ball screw, like for example a pitch error, can lead to a difference between the position measurements. Forces acting on the axis components like process forces, acceleration and friction lead to a mechanical deformation of axis components depending on the stiffness. Lastly, the thermal state of the components close to the machine axis has an influence on the encoder difference.

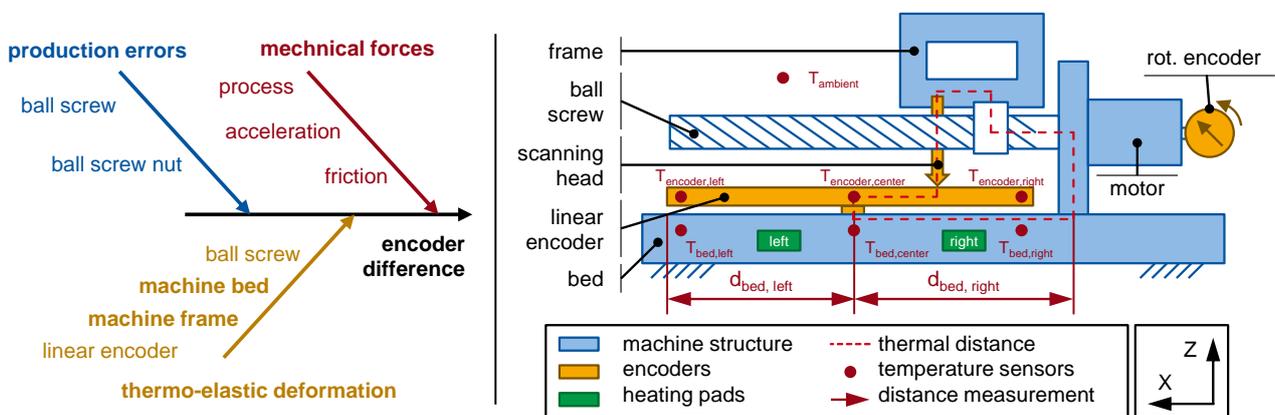


Fig. 1. left: influences on the encoder difference; right: sketch of a machine axis

Figure 1 (right) shows the axis components and the influence of component deformations on the encoder difference. The geometry of the components marked with “thermal distance” defines the encoder difference. Any deformation of these components will influence the encoder difference. Deformations may occur due to a change in component temperature. However, it is evident, that only deformations in the machine bed, machine frame or the linear encoder lead to a TCP displacement. A deformation of the ball screw leads to a change in

motor position but will not affect the position of the frame, since this is defined by the scanning head of the linear encoder.

The aim of this paper is a detailed investigation of the encoder difference and its connection to the component deformation close to the axis. Therefore, a series of loads is put on the machine and the resulting encoder difference and axis deformation is measured. The setup of these experiments is described in the following section.

4. MEASUREMENTS ON THE MACHINE

To investigate the machine tool behaviour close to the axis, the X axis of a stand milling machine with two working areas is equipped with temperature sensors and distance sensors as shown in Fig. 1. The machine axes length are: $X = 2000$ mm, $Y = 840$ mm and $Z = 550$ mm. Each working area has a table with two rotary axes in A and C direction. To measure the temperature, PT100 elements are used. For the displacements, linear variable differential transformers (LVDT) are mounted on the machine using Invar rods which do not deform with changing material temperature. The distances d_{bed} are measured using the measurement principle introduced by Brecher et al. [16].

5. INVESTIGATIONS

5.1. SELECTIVE LOAD USING HEATING PADS

For a first analysis of the machine behaviour, the machine is not moved and the structure and ambient temperature as well as the structure deformation are measured according to the measurement setup presented in Section 4. Additionally, the TCP displacement is measured using an environmental temperature variation error test (ETVE test) [17]. External thermal load is applied using heating pads (comp. Fig. 1). First, the left heating pad is activated for 12 h. After this, it is turned off and after another 12 h period, the right heating pad is activated for 12 h. The measured results are shown in Fig. 2.

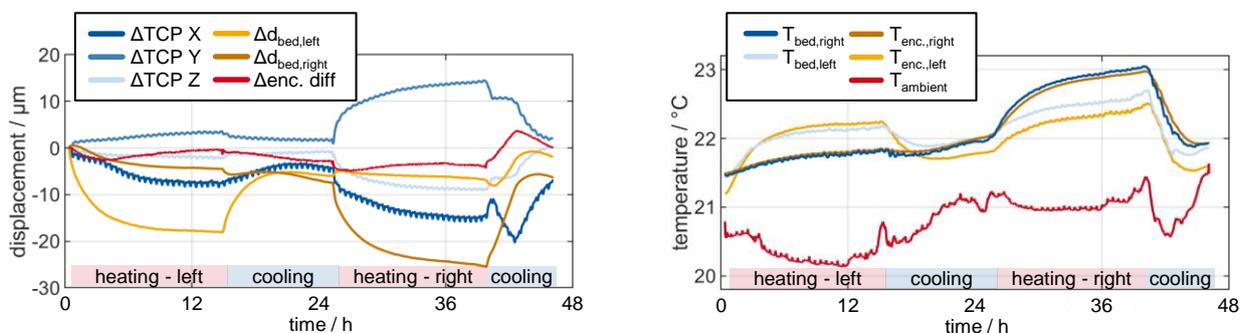


Fig. 2. left: encoder difference and measured displacements at the TCP and machine bed during external load using heating pads; right: machine structure and ambient temperature

During the heat up phases using the pads, the corresponding sections of the machine bed heat up and expand by $18\ \mu\text{m}$ on the left side and $24\ \mu\text{m}$ on the right side. The encoder difference however changes by maximally $5\ \mu\text{m}$. This shows, that when the heat flux is applied on a very specific point at the machine structure, multiple structure components close to the axis heat up resulting in a complex deformation of the axis components. This is reflected in the recorded temperatures on the structure components (comp. Fig. 2 – right). Therefore, in this observed case, the encoder difference does not show a suitable behaviour to detect changes in the TCP position.

5.2. INTERNAL THERMAL LOAD

To investigate the machine behaviour under internal thermal loads, a machine movement according to Fig. 3 is run on the machine. The X axis of the investigated machine is driven between the positions $X = -1935\ \text{mm}$ and $X = -1765\ \text{mm}$ at constant speeds for 12 h with cooling phases in between lasting 12 h as well. The speed is altered in four steps between $v_1 = 175\ \text{mms}^{-1}$ and $v_4 = 106\ \text{mms}^{-1}$.

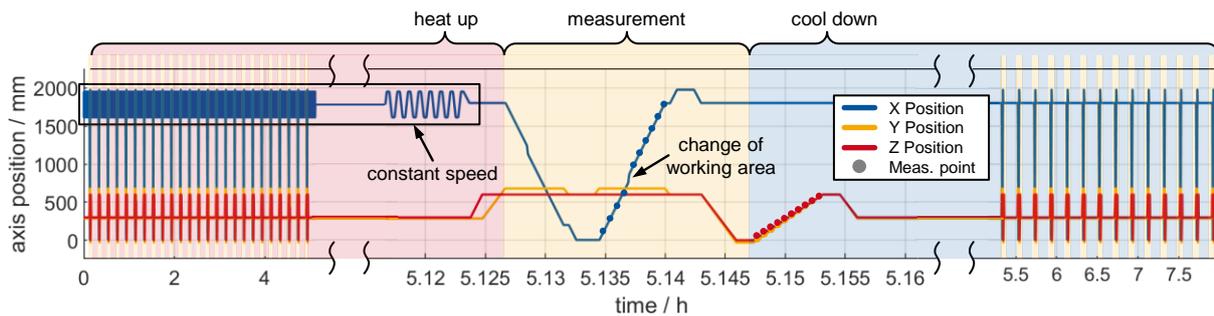


Fig. 3. measurement cycle of the encoder difference

During the whole experiment, every five minutes a measurement cycle is conducted by the machine, where all axes are driven with a constant speed over their complete length. The encoder difference of all axes is evaluated at ten positions. Figure 4 shows the machine structure temperature (left) and the encoder difference of the X axis (right) during the experiment. The data was recorded with the measurement setup presented in Section 4.

Figure 4 shows the expected temperature rise during the heat up phase and temperature fall during the cool down phases. The time constants of the machine components vary as expected. The encoder difference of the X axis follows the trend seen in the temperatures. It can be seen, that the change in encoder difference is smaller for positions closer to the fixed bearing than for positions further from it, which is consistent with the findings from Heidenhain [13]. The reason for this is the large influence of the ball screw. The balls screw contributes the most to the change in encoder difference, since it heats up the most.

Even though only the X axis is moved in the trials, the Y and Z axis also see a change in encoder difference. This is shown in Fig. 5 (left). Since the Z axis is geometrically furthest from the X axis, where the heat flows into the machine structure, the encoder difference of the Z axis is the smallest.

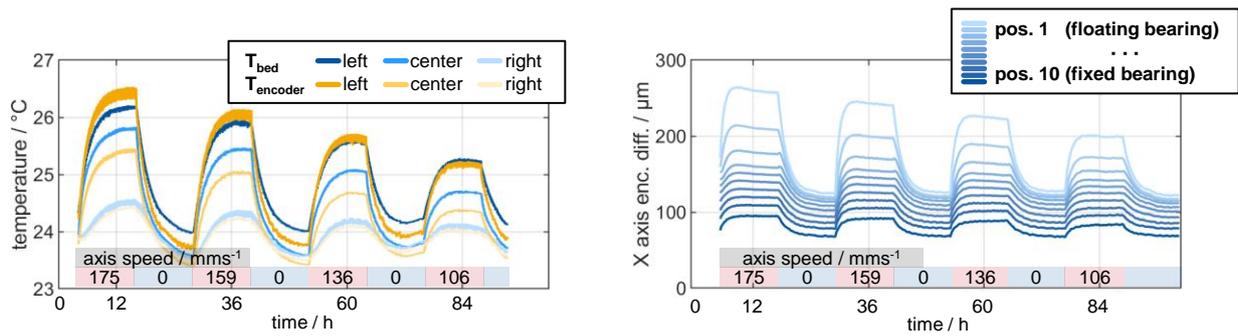


Fig. 4. left: machine structure temperature; right: encoder difference of the X axis

The encoder difference of the Y axis however shows a significant delay element behaviour as well.

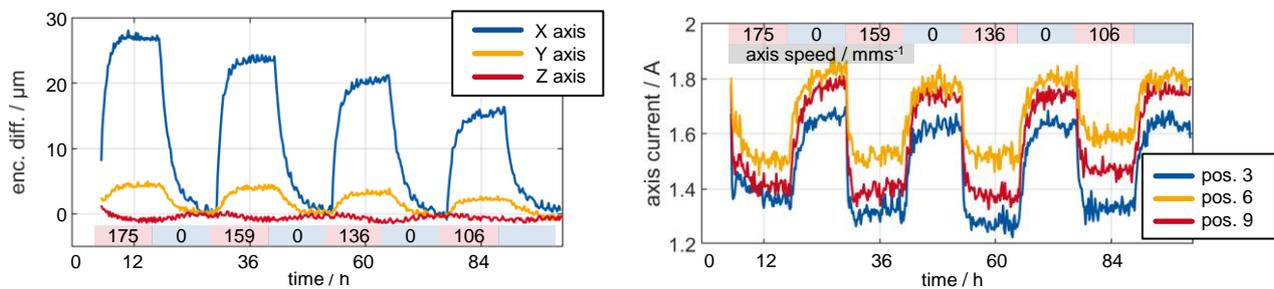


Fig. 5. left: encoder difference for all linear axes close to their fixed bearing; right: axis current evaluated at three positions during constant speed

Furthermore, the change in motor torque for constant movement of the X axis during measurement cycles is evaluated at three positions by assessing the motor current (Fig. 5 – right). It can be assumed that the current is proportional to the torque of the axis [15]. Consistent with the research of Wang et al., the trial in this paper shows a decreasing torque with time during machine load [14]. The reason for this is a change in the fit between the ball screw and the ball screw nut due to a change their changing temperatures and a change of the viscosity of the lubricant. The presented trial shows, that the current follows a low order delay element behaviour. The amplification of the delay elements seems to be independent of the axis speed. Depending on the position of the measurement, the current shows an offset. This may be attributed to production tolerances regarding the fit between the ball screw and the nut. The result shows, that the current measured carries an information about the thermo-elastic machine state and can therefore be used as an input parameter for a model computing the thermo-elastic TCP displacement. In Fig. 6, the amplification factor K of all three axes encoder differences is evaluated over the machine pose for different speeds.

The amplification factor K of the X axis shows significant changes depending on the axis speed. The encoder difference of the X axis has the largest amplification factor for all measurement positions. The amplification factor of the X axis shows a progressive trend for the axis positions upwards of 1500 mm. This trend is explained by the position of the warm up. For the warm up, the axis is run at a constant speed between the positions 1610 mm and

1960 mm (compare Fig. 3). Due to the friction of the ball screw nut the ball screw temperature is higher in this area, resulting in a larger elongation and thus a higher resulting amplification factor of the encoder difference. Since the fixed bearing is at the 0 mm position, this part of the bearing shows a smaller displacement.

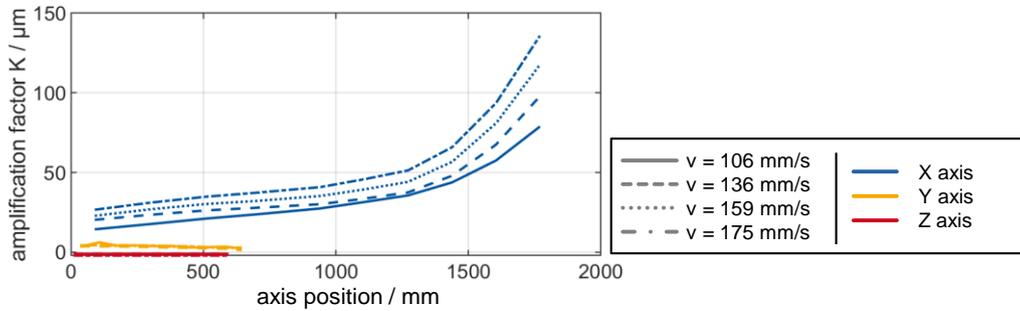


Fig. 6. Evaluation of the amplification factor K of the encoder difference

To generate an overview of the dependency between the position of the heat up, and the encoder difference, the experiment is repeated at the same speeds for varying heat up positions. Table 2 shows the positions between which the machine is driven to put a thermal load on the machine.

Table 2. Minimum and maximum position of the machine warm up for all experiments

experiment number	left working area				right working area			
	1	2	3	4	5	6	7	8
maximum position	385	385	300	215	1935	1935	1850	1765
minimum position	45	215	130	45	1595	1765	1680	1595

all data in mm

The result is shown in Fig. 7. It can be seen, that the position of the heat up cycle influences the encoder difference amplification factor distribution over the measurement positions. This means, that the current encoder difference distribution does contain an information about the historic energy input position in the machine tool, thus rendering the input valuable for a model to determine the TCP position.

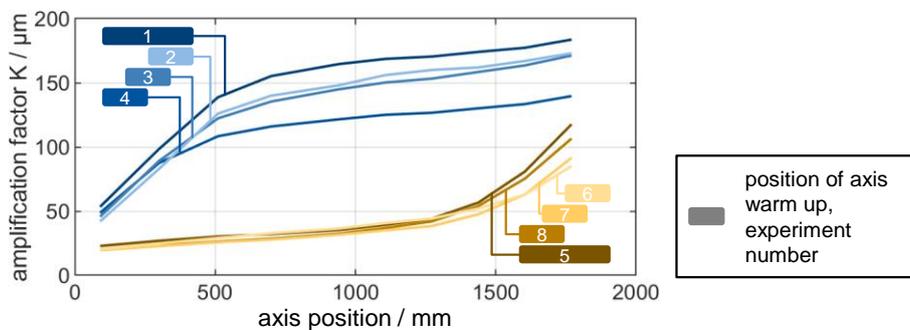


Fig. 7. Amplification factor of the encoder difference for all measurements dependent on the measurement position at maximum speed

6. EXTERNAL LOADS

The previous section shows, that the encoder difference is a good indicator for the current machine state, regarding internal loads. External thermal loads like a changing ambient temperature induce thermal distortions in the machine as well. To investigate the effect of the changing ambient temperature, the machine is not moved for eight days while the ambient temperature fluctuates (compare Fig. 8 – right). The TCP displacement is measured using an ETVE test [17]. The encoder difference, structure and ambient temperature as well as the TCP and structure displacement are recorded. The result of this test is shown in Fig. 8 (left).

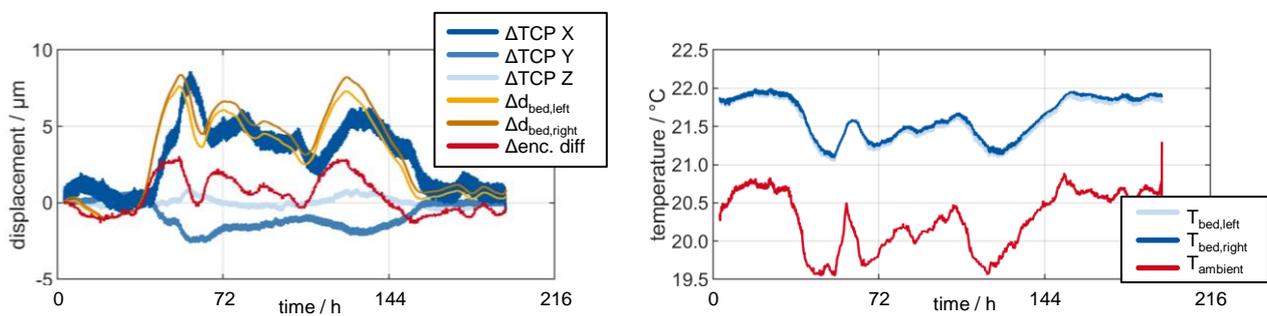


Fig. 8. left: encoder difference and measured displacements at the TCP and machine bed during external load by changing environment temperature; right: machine structure and ambient temperature

During the experiment, the ambient temperature changes by 1.5°C . This results in a change in bed temperature of about 1.0°C and a change in bed deformation of up to $7.5\ \mu\text{m}$. The bed deformation behaves very similarly to the displacement of the TCP position in X direction and is qualitatively reflected by the encoder difference. However, the encoder difference has a maximum deviation of around $4\ \mu\text{m}$. This shows that there is a qualitative relationship between the encoder difference and the thermo-elastic machine state based on ambient temperature.

7. MODELLING THE TCP DISPLACEMENT

To quantify the benefit of using the encoder difference, a model is built that estimates the TCP displacement. Machine learning approaches have proven to be very accurate for complex models while still being fast in computation time. Therefore a machine learning approach is chosen in this paper. Many published approaches model the TCP displacement efficiently using machine learning tools. Most of these approaches use temperatures measured by external sensors to model TCP displacement, as can be seen in the work of Liu et al. [10]. The types of models used vary widely. Backpropagation networks as used by Ma et al. [18], ant colony algorithm based backpropagation networks as used by Guo et al. [9], feedforward networks as used by Chen et al. [19], fuzzy ARTMAP networks as used by Mize et al. [20]

and many other types of networks can be found in the literature to model the TCP displacement of a machine tool. Because the studies are performed on fundamentally different machines, it is not possible to compare the approaches and define a superior model. However, a reoccurring model type in the literature is a Long Short Term Memory (LSTM) network to model TCP displacement. LSTM networks were first introduced by Hochreiter et al. [21] and are the most basic type of network that is able to “memorize” data from previous time steps. This is important for modelling the delayed dependency between inputs and TCP displacement described in Section 2. Therefore, this type of network is used in this paper to model TCP displacement. An occlusion sensitivity analysis is performed to evaluate which input values have the greatest impact on the accuracy of the TCP displacement calculation. The test is replicated using a commonly used backpropagation network. This type of network is not able to store data. If the backpropagation network also delivers good results, it can be deduced, that the input data shows a good and direct correlation to the thermos-elastic machine tool error. Figure 9 shows the model architectures used.

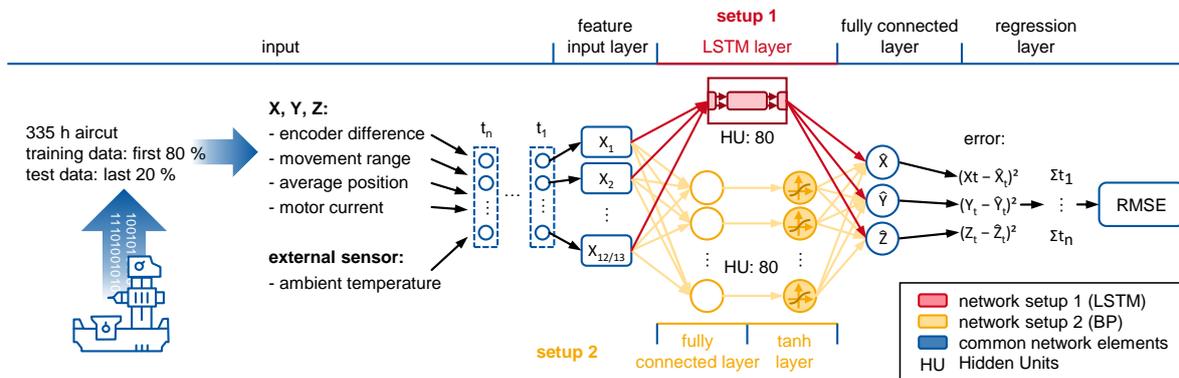


Fig. 9. LSTM and BP network model for computing the TCP displacement

To train the network a new data set is generated. This data set is supposed to resemble the machine movements in an industrial application. To do this the machine is driven for 355 h following randomly generated shapes like holes which are commonly used during machining of real parts. These shapes are shown in Fig. 10 (left).

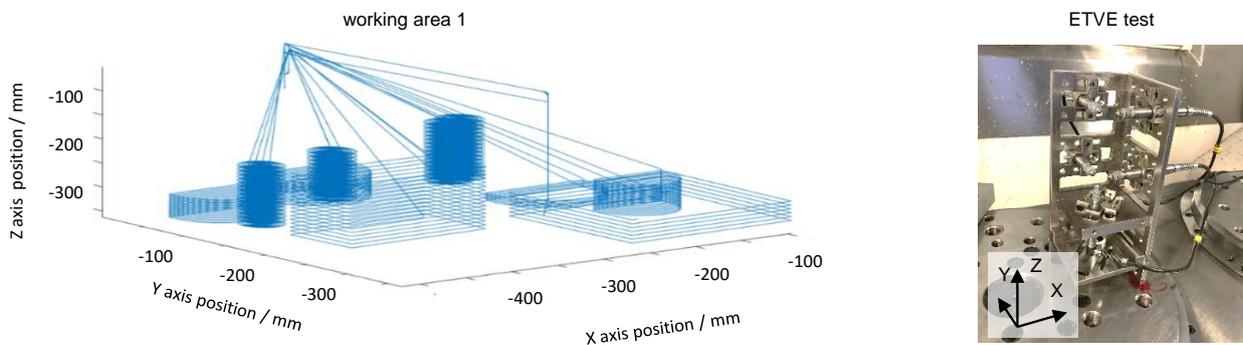


Fig. 10. left: TCP path for the generation of the data set; right: ETVE test

The speed of the machine tool is varied randomly as well as the geometric dimensions and positions of the shapes. The TCP displacement is measured at five minute intervals using an ETVE test [17] (Fig. 10 – right). While the TCP displacement is being measured, the machine is stopped for 5 seconds. During this standstill, the encoder difference is also measured for all axes. For the intervals between measurements, the motion, the average axis position and the motor current are calculated. The range of motion is the difference between the maximum and minimum position in the interval. The motor current is calculated as the average absolute motor current of the traced values.

The first 80% of a 335 h data set is used to train the network and the last 20% are used for validation. Figure 11 shows the result of the LSTM and BP networks using all input data. An occlusion sensitivity analysis is performed. In an occlusion analysis, specific inputs are not used in the training of the networks. The resulting change in the RSME of the network is compared with the RSME of the network using all data. The larger the increase in comparison to the network using all data is, the more important is the occluded value. For each network, the result of the calculated TCP displacement in the Z direction using all data is shown. The Z direction is the direction with the largest total TCP displacement.

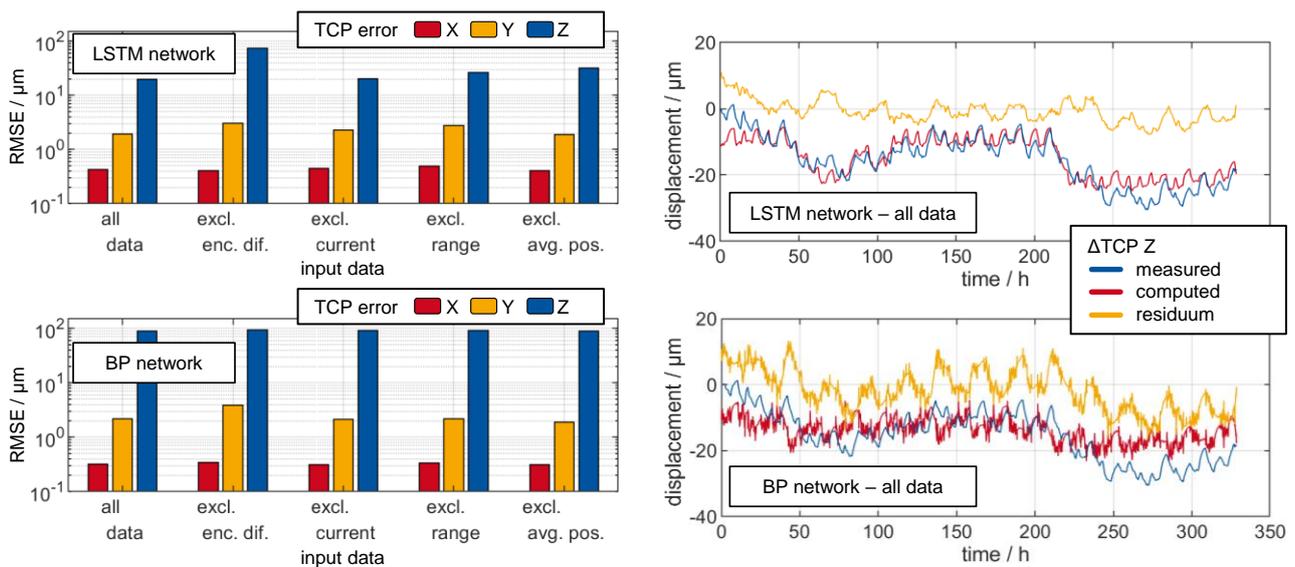


Fig. 11. Results for computing the TCP displacement using the LSTM and the BP networks

Figure 11 shows that the TCP displacement can be accurately calculated using the input data collected from the machine. The predicted displacement and the measured displacement are very similar and the peaks are accurately calculated. The occlusion sensitivity analysis demonstrates how the RMSE changes when no input data is used. The use of all input data forms the baseline. When occluding input data, the RMSE is expected to be higher because information about the machine state remains unused. The initial RMSE when using all data is 20 μm in the Z direction. When the encoder difference is excluded, the RMSE is 75 μm. This implies that the encoder difference has a large effect on the result. The other input dimensions display a smaller increase in RMSE, indicating that they contribute less to the

overall result of TCP displacement. However, the RMSE increases with the occlusion of any input data, showing that all input data contribute to the overall result. A similar result can be seen for the other TCP displacement directions.

The BP network, however, produces different results. The overall error generated by the BP network is significantly larger. The occlusion sensitivity analysis does not demonstrate a dominant shift in the RMSE as a function of the input variables. This implies that the network is unable to accurately predict TCP displacement. It can be concluded that the encoder difference does not have a sufficient direct relationship to the current thermal machine state to model the TCP displacement without considering the delay of the TCP displacement. However, the machine-internal data contains sufficient information to model the TCP displacement considering the delayed relationship.

8. SUMMARY AND OUTLOOK

Several experiments were conducted to investigate the relationship between the thermo-elastic deformation of the machine structure and the encoder difference on a upright milling machine with two working spaces. The trials were conducted using heat pads as an external heat source with a controlled heat flux position, axis motion as an internal heat source resulting in a heat flux similar to the real working conditions of a machine tool, and daily changes in ambient temperature.

The experiments show a dependency between the encoder difference and the TCP displacement under various operating conditions. The dependency appears to be complex, which means that it cannot be used in a white box model to calculate the thermo-elastic TCP error. However, the encoder difference can be used as an additional input in black box models with the potential to significantly improve the accuracy of the thermo-elastic TCP error calculation.

In addition, simple black-box models that compute the TCP displacement have been developed. The LSTM model shows the best prediction accuracy. By performing an occlusion analysis, it is shown that the encoder difference is the input in the neural network that has the greatest influence on the prediction accuracy.

The shown results reflect the high potential of the encoder difference as an input dimension for models calculating the thermo-elastic TCP displacement of machine tools. However, a drawback of the model is the large training data set required. Further research should focus on the relationship between the size of the training data set and the prediction accuracy, as well as approaches to reduce the size of the training data set, such as transfer learning.

ACKNOWLEDGEMENTS

The presented findings are funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 174223256.

REFERENCES

- [1] HACKSTEINER M., DUER F., AYATOLLAHI I., BLEICHER F., 2017, *Automatic Assessment of Machine Tool Energy Efficiency and Productivity*, Procedia CIRP, 62, 317–22.
- [2] GAO W., IBARAKI S., DONMEZ M.A., KONO D., MAYER J., CHEN Y.-L., SZIPKA K., ARCHENTI A., LINARES J.-M., SUZUKI N., 2023, *Machine Tool Calibration: Measurement, Modeling, and Compensation of Machine Tool Errors*, International Journal of Machine Tools and Manufacture, 187, <https://doi.org/10.1016/j.ijmachtools.2023.104017>.
- [3] DONMEZ M.A., HAHN M.H., SOONS J.A., 2007, *A Novel Cooling System to Reduce Thermally-Induced Errors of Machine Tools*, CIRP Annals – Manufacturing Technology, 56/1, 521–524, <https://doi.org/10.1016/j.cirp.2007.05.124>.
- [4] HOREJS O., MARES M., FIALA S., HAVLIK L., STRITESKY P., 2020, *Effects of Cooling Systems on the Thermal Behaviour of Machine Tools and Thermal Error Models*, Journal of Machine Engineering, 20/4, 5–27, <https://doi.org/10.36897/jme/128144>.
- [5] BRECHER C., NEUS S., DEHN M., *EFFICIENT FE.*, 2020, *Efficient FE-Modelling of the Transient Thermo-Elastic Machine Behaviour of 5-Axes Machine Tools*, Euspen, Thermal Issues, Northampton, 148–149.
- [6] BRECHER C., FEY M., WENNEMER M., 2016, *Volumetric Measurement of the Transient Thermo-Elastic Machine Tool Behavior*, Prod. Eng. Res. Devel., 10/3, 345–350.
- [7] NAUMANN C., GLÄNZEL J., DIX M., IHLENFELDT S., KLIMANT P., 2022, *Optimization of Characteristic Diagram Based Thermal Error Compensation via Load Case Dependent Model Updates*, Journal of Machine Engineering, 22/2, 43–56, <https://doi.org/10.36897/jme/148181>.
- [8] NAUMANN C., GLÄNZEL J., PUTZ M., 2020, *Comparison of Basis Functions for Thermal Error Compensation Based on Regression Analysis – a Simulation Based Case Study*, Journal of Machine Engineering, 20/4, 28–40, <https://doi.org/10.36897/jme/128629>.
- [9] GUO Q., YANG J., WU H., 2010, *Application of ACO-BPN to Thermal Error Modeling of NC Machine Tool*, International Journal of Advanced Manufacturing Technology, 50/5–8, 667–675.
- [10] LIU P.-L., DU Z.-C., LI H.-M., DENG M., FENG X.-B., YANG J.-G., 2021, *Thermal Error Modeling Based on BiLSTM Deep Learning for CNC Machine Tool*, Advances in Manufacturing, 9/2, 235–249.
- [11] WENNEMER M., 2018, *Methode zur messtechnischen Analyse und Charakterisierung volumetrischer thermo-elastischer Verlagerungen von Werkzeugmaschinen*, Diss. Aachen: Apprimus.
- [12] CZWARTOSZ R., JEDRZEJEWSKI J., 2022, *Application of Machine Learning in the Precise and Cost-Effective Self-Compensation of the Thermal Errors of CNC Machine Tools – A Review*, Journal of Machine Engineering, 22/3, 59–77, <https://doi.org/10.36897/jme/152246>.
- [13] Dr. JOHANNES HEIDENHAIN GmbH, 2006, *Accuracy of Feed Axes: Technical Information*, Traunreuth.
- [14] WANG H., LI F., CAI Y., LIU Y., YANG Y., 2020, *Experimental and Theoretical Analysis of Ball Screw Under Thermal Effect*, Tribology International, 152, 106503, <https://doi.org/10.1016/j.triboint.2020.106503>.
- [15] XI T., BENINCÁ I.M., KEHNE S., FEY M., BRECHER C., 2021, *Tool Wear Monitoring in Roughing and Finishing Processes Based on Machine Internal Data*, International Journal of Advanced Manufacturing Technology, 113/11–12, 3543–3554.
- [16] BRECHER C., LEE TH., TZANETOS F., ZONTAR D., 2019, *Hybrid Modeling of Thermo-Elastic Behavior of a Three-Axis Machining Center Using Integral Deformation Sensors*, Procedia CIRP, 8, 1301–1306, <https://doi.org/10.1016/j.procir.2019.04.017>.
- [17] International Organization for Standardization, 2007, *Test Code for Machine Tools: Determination of Thermal Effects; (230 Teil 3)*.
- [18] MA C., ZHAO L., MEI X., SHI H., YANG J., 2017, *Thermal Error Compensation of High-Speed Spindle System Based on a Modified BP Neural Network*, International Journal of Advanced Manufacturing Technology, 89/9–12, 3071–3085.
- [19] CHEN Y., CHEN J., XU G., 2021, *A Data-Driven Model for Thermal Error Prediction Considering Thermoelasticity with Gated Recurrent Unit Attention*, Measurement, 184, 109891, <https://doi.org/10.1016/j.measurement.2021.109891>.
- [20] MIZE C.D., ZIEGERT J.C., 1999, *Neural Network Thermal Error Compensation of a Machining Center*, Journal of the International Societies for Precision Engineering and Nanotechnology.
- [21] HOCHREITER S., SCHMIDHUBER J., 1997, *Long short-term memory*, Neural Computation, 9/8, 1735–1780, <https://doi.org/10.1162/neco.1997.9.8.1735>.