Int J Adv Manuf Technol (2011) 54:669–676
DOI 10.1007/s00170-010-2961-3

ORIGINAL ARTICLE

Modified Elman network for thermal deformation compensation modeling in machine tools

Zhiyong Yang · Minglu Sun · Weiqian Li · Wenyong Liang

Received: 11 June 2010 / Accepted: 20 September 2010 / Published online: 7 October 2010

Abstract Thermal deformation is one of the most significant causes of machining errors in machine tools. One effective method is to build a compensation system to offset the thermal errors. Therefore, an accurate model is the key part of the compensation system. This study proposed a modified Elman network (EN) to improve the prediction accuracy of the compensation model in machine tools. And the improved EN can be regarded as a feed-forward neural network with feedback from hidden layer and output layer as an additional set of inputs. The structure of this network determines its dynamic characteristic with memory function. On the other hand, thermal deformation of the spindle contributes the largest part of total thermal errors in precision machining. Then a precise finite element model of machine tool spindle was established. And a new method for calculating the heat transfer convection coefficient on the surface of the spindle was proposed in this paper. The improved EN was used to map the nonlinear relationship between temperature field and thermal errors of the spindle. At last, a verification experiment was implemented on a CNC center and some satisfying results were achieved.

Keywords Thermal deformation · Elman network · Finite element method · Simulation · Machine tool

Z. Yang (✉) · M. Sun
School of Mechanical Engineering, Tianjin University,
Tianjin 300072, China
e-mail: meyang@tju.edu.cn

W. Li · W. Liang
Spark Machine Tools Ltd,
Tianshui 741024, China

1 Introduction

In recent years, with the development of manufacture industry, high speed machining technology has been widely used. Because of high speed rotation of the spindle, a large quantity of friction heat is generated at the front bearing, rear bearing, and other moving parts. Different parts of machine tools have different thermal characteristic. Therefore, the thermal deformation occurs when the friction heat transmits to the spindle head and tool holder, etc. Because mechanical and geometric errors have been well studied and greatly minimized, thermal errors are the most significant factor that influences the accuracy of machine tools, which account for as much as 40% to 70% of total errors [1]. Although thermal errors could be decreased by optimization design of machine tool structure in some sense, error compensation is another more effective and economic way to improve the accuracy, especially to some less accurate machine tools which cannot be replaced immediately [2].

The commonly used error compensation technology is to build the off-line empirical model between temperatures of a few points and the thermal deformation [3]. By measuring temperatures of key points during machining, the compensation value of coordinate system could be estimated by the off-line model, and then feed the value to the controller for real-time compensation. Temperature distribution usually changes widely under different machining conditions. An accurate and robust thermal error estimation model is the most important part of the compensation system. Linear regression model has been widely used to predict the thermal deformation [4, 5]. And autoregressive moving average has also been used in a horizontal machining center [6]. With simple structure, the linear regression and autoregressive moving average models need less computer source. But they cannot take time delay into consideration or accurately map the nonlinear relationship between the temperature field and deformation. Grey system theory was used to develop the thermal deformation modeling methodology and optimize measuring points [7, 8]. Choi and Lee applied genetic algorithm to build the thermal error model [9]. And in recent years, different types of artificial neural network have been widely used in thermal error modeling. Hattori [10], Veldhuis [11], and Yuan Kang [12] proposed feed-forward neural network to build the compensation model, and the results showed that feed-forward neural network with back-propagation algorithm has high level of accuracy in data fitting, but is not so well in data forecast.

In section 2, this study proposed a modified EN with feedback not only from the hidden layer, but also from the output layer. This type of structure makes it sensitive to the history of input data. It could be very useful for dynamic system modeling. The modified EN can get more information from limited sampling spots, and has higher level of generalization ability than normal EN. Numerical solutions can approximate the analytical solutions very well as long as the spindle structure is correctly and finely meshed and the power of heat sources and heat transfer coefficients are well defined [13–15]. An accurate finite element model of machine tool spindle was established in section 3. And the modified EN was used to map the relationship between temperatures of key points and thermal errors of the spindle. In section 4, a verification experiment was implemented on a CNC center and the results showed that modified EN has high level of prediction accuracy and generalization ability. And some conclusions were drawn in section 5.

2 Modified EN for thermal deformation modeling

There are two main types of neural network for thermal deformation modeling: feed-forward neural network and recurrent neural network. In feed-forward network, the information can only transmit in one direction: from input layer to output layer. But in recurrent network, information can transmit in two directions: from input layer to output layer and the other way round. Feed-forward network is a spatial model to solve temporal modeling problem, which requires a large number of neurons and a long computation time, but has poor generalization ability. Meanwhile, the recurrent network allows feeding back previous outputs of one front layer to input layer. This makes the recurrent network have the characteristic of memory function. And EN is one of the most typical recurrent networks, as shown in Fig. 1.

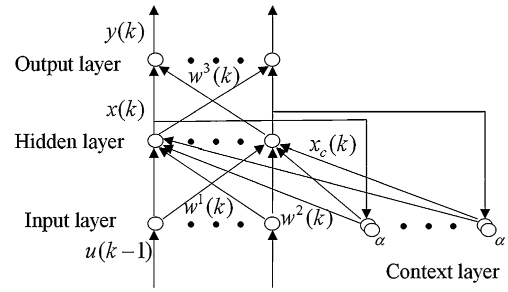


Fig. 1 EN structure

2.1 The structure of EN

It has an r-dimensional external input vector $u(k-1) = [u_1(k-1), u_2(k-1), \cdots, u_r(k-1)]^T$ and an m-dimensional output vector $y(k) = [y_1(k), y_2(k), \cdots, y_m(k)]^T$. The output layer weight matrix is $w^3(k) \in R^{m\times n}$, and the hidden layer weight matrix is $w(k) = [w^1(k), w^2(k)] \in R^{n\times(r+n)}$, where $w^1(k) \in R^{n\times r}$ is the input signal sub-weight matrix, $w^2(k) \in R^{n\times n}$ is the context sub-weight matrix, and n is the number of hidden layer neurons. For the input vector $u(k-1)$, the corresponding output $y(k)$ can be expressed as:

$$y(k) = w^3(k)x(k) \tag{1}$$

$$x(k) = f\left(w^1(k)u(k-1) + w^2(k)x_c(k)\right) \tag{2}$$

$$x_c(k) = x(k-1) + \alpha x_c(k-1) \quad 0 \leq \alpha 1 \tag{3}$$

where $f(\cdot)$ is the activation function, which is usually defined as $f(x) = \frac{1}{1+e^{-x}}$, α is the feedback gain of self-connection, and $x(k) \in R^n$ is the output of context layer.

The desired output is $y_d(k)$. Therefore, the error function can be defined as:

$$E(k) = \frac{1}{2}(y_d(k) - y(k))^T(y_d(k) - y(k)) \tag{4}$$

The EN's training objective is to minimize $E(k)$ by recursively updating the weights of each layer. The training process can be expressed as follows:

$$\Delta w_{ij}^3 = \beta_3 \delta_i^0 x_j(k) \qquad i = 1, 2, \ldots, m; j = 1, 2, \ldots, n \tag{5}$$

$$\Delta w_{jl}^2 = \beta_2 \sum_{i=1}^{m} \left(\delta_i^0 w_{ij}^3\right) \frac{\partial x_j(k)}{\partial w_{jl}^2} \; j = 1, 2, \ldots, n; l = 1, 2, \ldots n \tag{6}$$

$$\Delta w_{jq}^1 = \beta_1 \delta_j^h u_q(k-1) \; j = 1, 2, \ldots, n; q = 1, 2, \ldots, r \tag{7}$$

where $\beta1$, $\beta2$, and $\beta3$ are the learning steps of $w^1(k)$, $w^2(k)$, and $w^3(k)$, respectively, and

$$\delta_i^0 = y_{d,i}(k) - y_i(k) \tag{8}$$

$$\frac{\partial x_j(k)}{\partial w_{jl}^2} = f_j^{'}(\cdot)x_l(k-1) + \alpha\frac{\partial x_j(k-1)}{\partial w_{jl}^2} \tag{9}$$

$$\delta_j^h = \sum_{i-1}^{m}\left(\delta_i^0 w_{ij}^3\right)f_j^{'}(\cdot) \tag{10}$$

2.2 The modified EN

The standard EN only takes the feedback from hidden layer into consideration and has been widely used in nonlinear system modeling. To improve the characteristic of the EN, a new improved EN is proposed in this paper, as shown in Fig. 2.

The hidden layer weight matrix becomes $w(k) = [w^1(k), w^2(k), w^4(k)] \in R^{n\times(r+2n)}$ and the output $y(k)$ can be expressed as:

$$y(k) = w^3(k)x(k) \tag{11}$$

$$x(k) = f\left(w^1(k)u(k-1) + w^2(k)x_c(k) + w^4(k)y_c(k)\right) \tag{12}$$

$$x_c(k) = \alpha x_c(k-1) + x(k-1) \qquad 0 \le \alpha 1 \tag{13}$$

$$y_c(k) = \gamma y_c(k-1) + y(k-1) \qquad 0 \le \gamma 1 \tag{14}$$

where $y_c(k) \in R^n$ is the output of context layer 2 and γ is the feedback gain of output layer.

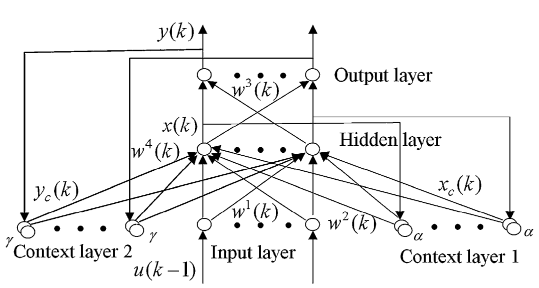


Fig. 2 The structure of modified EN

The training process of modified EN can be expressed as Eqs. 5, 6, and 7 and follows:

$$\Delta w_{js}^4 = \beta_4\sum_{i=1}^{m}\left(\delta_i^0 w_{ij}^3\right)\frac{\partial x_j(k)}{\partial w_{js}^4} \qquad j = 1,2,\ldots,n; s = 1,2,\ldots,m \tag{15}$$

$$\frac{\partial x_j(k)}{\partial w_{js}^4} = f_j^{'}(\cdot)y_c(k-1) + \gamma\frac{\partial x_j(k-1)}{\partial w_{js}^4} \; j = 1,2,\ldots,n; s = 1,2,\ldots,m \tag{16}$$

where β_4 is the learning step of $w^4(k)$.

The proposed modified EN is more sensitive to the history of input data, which would be very useful for dynamic system modeling and could get more information from limited sampling spots, and has higher level of generalization ability than standard EN.

3 Numerical simulation of spindle thermal deformation

Because of different sizes, shapes and thermal characteristic of each part and the uncertainty of thermal contact resistance at the interface of different parts of the spindle system, it is very difficult to get the analytical solutions of spindle thermal deformation. Finite element method is a practical way to get the numerical solutions of the thermal deformation. Numerical solutions can reach high level of accuracy as long as the boundary conditions are well defined.

3.1 Computation of heat generation

The friction of bearings is the main heat source of spindle system that causes thermal deformation. It can be computed as follows:

$$H_f = 1.047 \times 10^{-4}\mathrm{nM} \tag{17}$$

where H_f is the heat generation power (*W*), n is the rotating speed of the spindle (*rpm*), and M is the total frictional torque which consists of two components:

$$M = M_1 + M_2 \tag{18}$$

$$M_1 = f_1 p_1 d_m \tag{19}$$

$$M_2 = 10^{-7} f_0 (v_0 n)^{2/3} d_m^3 \qquad \text{if } v_0 n \ge 2,000 \tag{20}$$

$$M_2 = 160 \times 10^{-7} f_0 d_m^3 \qquad \text{if } v_0 n < 2,000 \tag{21}$$

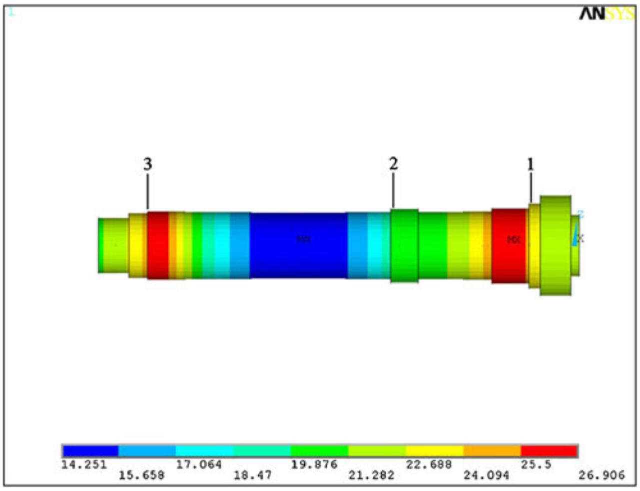


Fig. 3 Temperature distribution of the spindle at thermal balance state in work cycle 8

where M_1 is caused by the applied load, M_2 is caused by the viscosity of lubricant, f_1 is a factor related to the bearing type and load, P_1 is the bearing preload (N), d_m is the mean diameter of the bearing (mm), f_0 is a factor related to bearing type and lubrication method, and v_0 is the kinematic viscosity of the lubricant (mm^2/s). As the temperature rises, the viscosity of the lubricant decreases and then the heat generation power decreases, too.

3.2 Computation of heat transfer coefficients

The spindle rotating in the air is very similar to the situation in which the air flows along a flat plate. Both of them are forced convection. The boundary layer thickness of the air grows steadily and the heat transfer coefficient attenuates against time rather than a constant, no matter whether it is in the condition of laminar flow or turbulent flow.

3.2.1 The condition of laminar flow

The temperature of the spindle changes against time. When the air around the spindle is in the condition of laminar flow, the coefficient of heat convection can be computed as follows:

$$h_x = \frac{Nu_x \cdot \lambda}{x} \tag{22}$$

$$Nu_x = 1.61 \times 0.332\mathrm{Re}_x^{1/2}\mathrm{Pr}^{1/3} \tag{23}$$

$$\mathrm{Re}_x = \frac{u_x \cdot x}{v} \tag{24}$$

$$x = u_x t \tag{25}$$

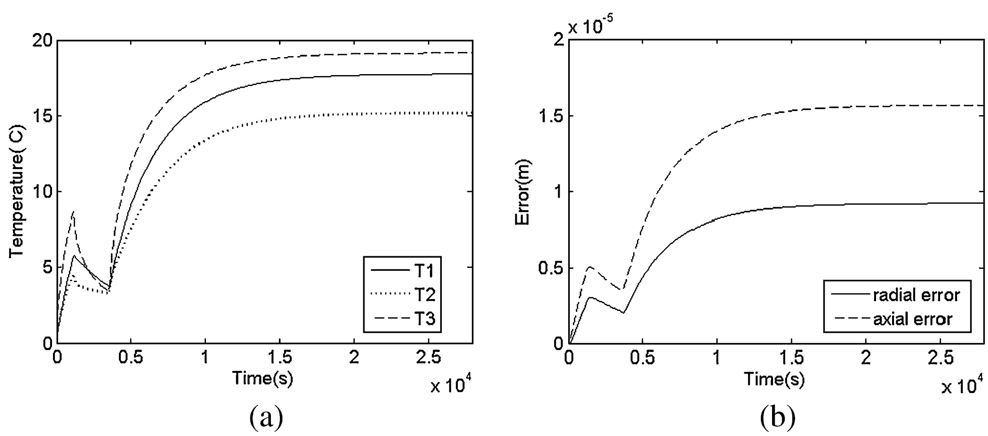
$$u_x = \frac{\pi d n}{60} \tag{26}$$

where h_x is the coefficient of heat convection $(W/(m^2 \cdot {}^\circ C))$. Nu, Re, Pr, and λ are the Nusselt number, Reynolds number, Prandtl number, and heat conductivity factor of the air, respectively. x is the moving trajectory length of a surface point of the rotating spindle. u_x is the tangential velocity of the spindle's surface (m/s). v is the viscosity of air (m^2/s). t is the spindle rotating time, and d is the diameter of the spindle. These equations are valid when $\mathrm{Re}_x 5 \times 10^5$, $0.6 \leq \mathrm{Pr} \leq 15$.

Substituting Eqs. 23, 24, and 25 into Eq. 22 gives the following equation:

$$h_x = 0.5345\frac{\mathrm{Pr}^{1/3}\lambda}{v^{1/2}t^{1/2}} \tag{27}$$

Fig. 4 Work cycle 1. **a** Temperature of the three points. **b** Radial error and axial error



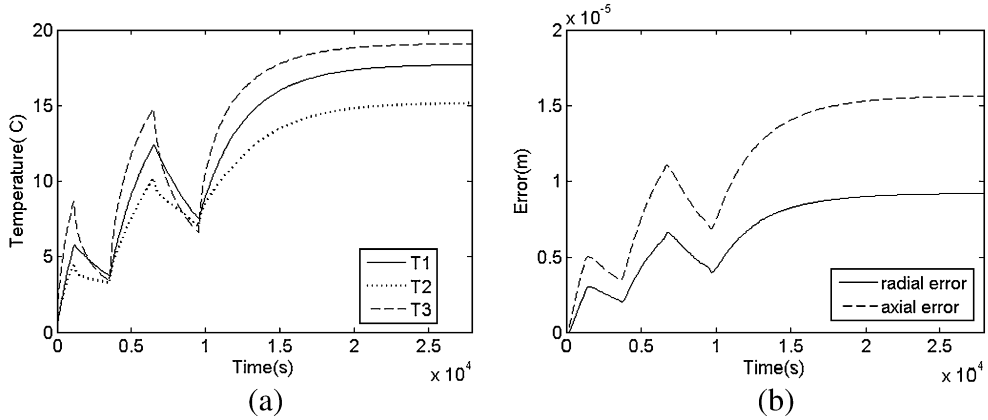


Fig. 5 Work cycle 4. **a** Temperature of the three points. **b** Radial error and axial error

3.2.2 The condition of turbulent flow

The critical point changing from laminar flow to turbulent flow can be generally computed as:

$$x_{\text{critical}} = \frac{\text{Re}_{\text{critical}} \nu}{u_x} \tag{28}$$

According to practical experience, the critical Reynolds number $\text{Re}_{\text{critical}}$ equals to 5×10^5. When the air around the spindle is in the condition of turbulent flow, the coefficient of heat convection can be computed as:

$$\text{Nu}_x = 0.0296 \text{Re}_x^{4/5} \text{Pr}^{1/3} \tag{29}$$

Substituting Eqs. 24, 25, and 29 into Eq. 22 gives the following equation:

$$h_x = 0.0296 \frac{u_x^{3/5} \text{Pr}^{1/3} \lambda}{\nu^{4/5} t^{1/5}} \tag{30}$$

Equations 27 and 30 show that the coefficient of convection heat transfer changes against time. It does not sustain decreasing because of the limitation of space, but approaches to a steady value. Therefore, the ambient space of the spindle is a crucial factor that influences the steady value.

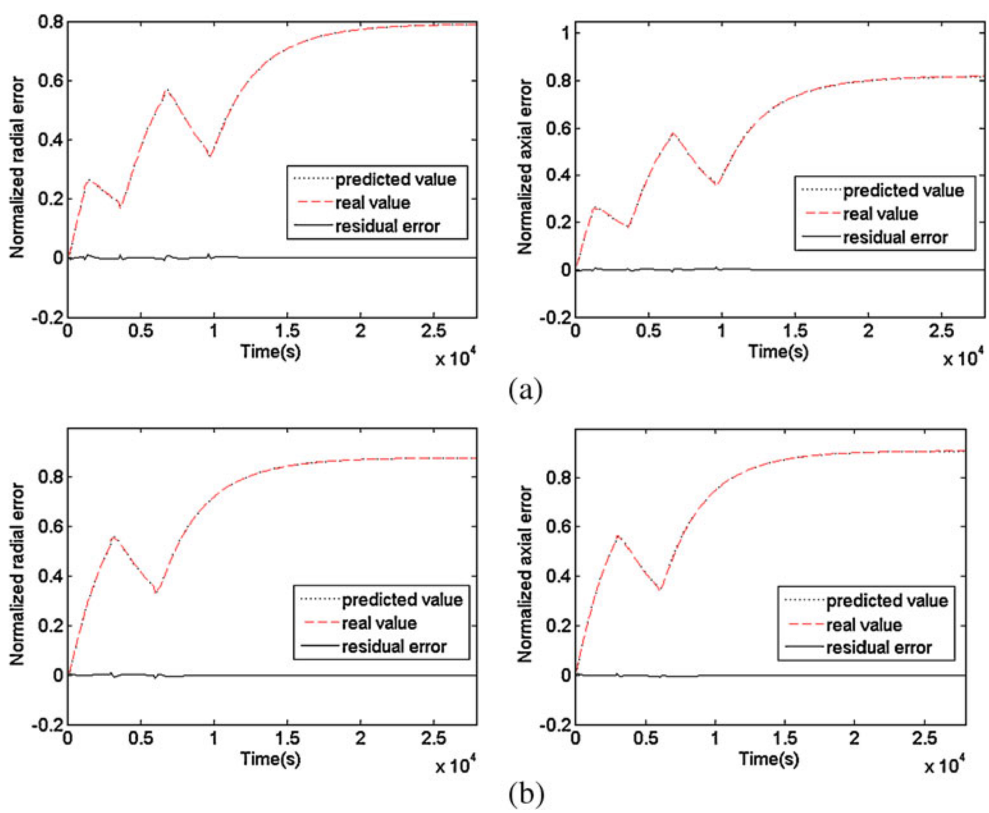


Fig. 6 The predicted value of the modified EN. **a** Work cycle 7. **b** Work cycle 8

Fig. 7 The CNC center and sensors distribution

3.2.3 The condition of natural convection

When the spindle stops rotating, the convection form becomes natural convection, and the coefficient of convection heat transfer can be computed as follows:

$$\mathrm{Nu} = \left\{ 0.60 + \frac{0.387\mathrm{R}\alpha^{1/6}}{\left[1 + (0.559/\,\mathrm{Pr})^{9/16}\right]^{8/27}} \right\}^2 \tag{31}$$

$$\mathrm{R}\alpha = \mathrm{Gr}\,\mathrm{Pr} \tag{32}$$

$$\mathrm{Gr} = \frac{g\alpha_v \Delta t d^3}{v^2} \tag{33}$$

where α_v is the temperature expansion coefficient of air, Δt is temperature difference between air and the surface of spindle, and g is the acceleration of gravity.

3.3 Simulation of machine tool spindle by finite element method

A machine tool spindle is studied here. The thermal loads and coefficients of convection heat transfer can be computed by equations mentioned above. The spindle is sustained by front bearing 1, front bearing 2 which can increase the rigidity of the spindle, and the rear bearing. Only the front bearing 1 does not allow the spindle to move freely in axial direction in order to reduce the axial error. Therefore, axial thermal error is caused only by the part between front bearing 1 and spindle nose. And the material density is 7.8×103 kg/m^3. Thermal conductivity is $42.8\mathrm{W}/(\mathrm{m}^2\times{}^\circ\mathrm{C})$. Thermal capacity is $480\mathrm{J}/(\mathrm{kg}\times{}^\circ\mathrm{C})$. Thermal expansion coefficient is $1.15\times10^{-5}1/{}^\circ\mathrm{C}$. The modulus of elasticity is 206 GPa, and Poisson ratio is 0.3. Machine tool's work cycles 1, 2, and 3 are as follows: starting to run for 1,200 s, and then stopping for 2,400 s, and then restarting to run for 24,400 s at the speed of 1,500, 2,000, and 2,500 rpm, respectively. Work cycles 4, 5, 6, and 7 are as follows: starting to run for 1,200 s, and then stopping for 2,400 s, and then restarting to run for 3,000 s, and then stopping for 3,000 s, and then restarting to run for 18,400 s at the speed of 1,500, 2,000, 2,500, and 1,800 rpm, respectively. And work cycle 8 is as follows: starting to run for 3,000 s, and then stopping for 3,000 s, and then restarting to run for 22,000 s at the speed of 2,000 rpm. At the end of each cycle, the thermal balance state was got, as shown in Fig. 3. The dynamic characteristic of the spindle in work cycle 2 and 3 has the similar trend but different amplitudes with work cycle 1, as shown in Fig. 4. And the dynamic characteristic in work cycles 5, 6, and 7 has the similar trend but different amplitudes with work cycle 4, as shown in Fig. 5.

Thermal key points are the best points used to model thermal errors by correlating their temperatures to the thermal errors. According to the method mentioned in [13], points 1, 2, and 3 are selected to be the key points, as shown in Fig. 3. Due to the advantage for nonlinear dynamic system identification, the modified EN is used in this study to model the thermo-elastic system. The input layer has three nodes (r=3): T1, T2, and T3. And the output layer has two nodes (m=2): radial error and axial error. The number of the hidden nodes is an important factor to the success of the network. It is difficult to get convergent if the number is too small. And the network overworks and has poor antidisturbance ability if the number is too large. According to the Kolmogorov theorem [16] and practical experience, the number of hidden nodes is defined to be 8. Use the data from work cycles 1, 2, 3, 4, 5, and 6 to train the modified EN, and

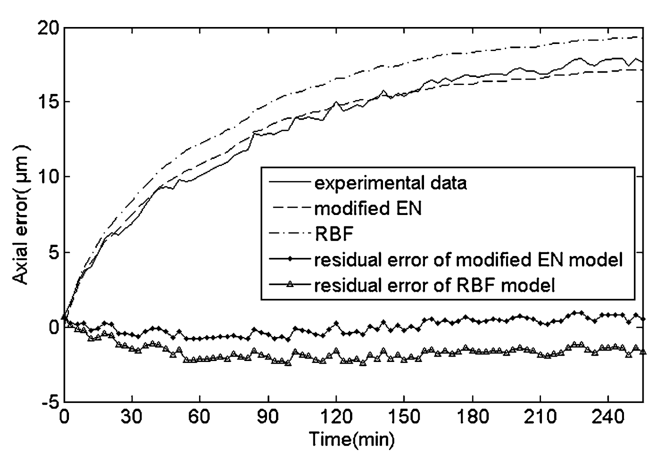


Fig. 8 The comparison between modified EN model and RBF model

data from work cycles 7 and 8 to test the prediction accuracy of the model. The training data and testing data for the network have been normalized within the range [0, 1].

The comparison between the real error from finite element analysis and the predicted error from the modified EN model is represented in Fig. 6. It can be observed that the model predicts the error very well, and the residual error is quite small, depicting that an accurate thermal model has been obtained.

4 Experiments

To further confirm the validity of the modified EN model, a verification experiment was implemented on a CNC center, as shown in Fig. 7. The machining center used in this experiment is produced by VICTOR Taichung Machinery Works Co., Ltd., and its model number is V-65. The reference cylinder is the standard spindle test rod. Three temperature sensors were placed on three key points, and another one was used to record the ambience temperature. The type of the temperature sensors is PT100 platinum resistance thermometer, and they offer an accuracy of ±0.2°C. Five noncontact displacement sensors were placed to measure thermal errors of the machine tool spindle. The type of the displacement sensors is Eddy current sensor KD2306 with the resolution of 0.1 μm, and their accuracy is more than 99%, upper limit of sampling frequency is 50 KHz. The sampling time of this experiment is set to be 1 min, and the measurement method is five-point method (shown in Chinese Standard T17421.3-2009). The radial error resulted from vibration of the cylinder and thermal deformation of the whole system and so on. In this experiment, the radial error was quite small, just several micrometers. Meanwhile, the axial error was a little larger, about 18 μm, and was affected by the vibration very little. Therefore, only the axial error was studied which was more suitable for testing the accuracy of the modified EN. Use the data from the first five experiments to train the modified EN and the sixth experimental data to test the prediction accuracy of the modified EN.

It took some time to build the compensation model. But as long as the compensation model was established, it only took a few seconds to work out the predicted error based on the temperatures which were gotten by temperature sensors. As shown in Fig. 8, it can be observed that the modified EN model predicted the thermal error much better than the RBF network model, and more than 90% of thermal error could be compensated by this model. In all, the modeling method presented in this paper is applicable and effective for practical processing. It can notably improve the precision of the thermal error model.

5 Conclusions

In this paper, a robust thermal error modeling method based on the modified EN was developed to estimate the thermal errors of machine tools. The modified EN is a recurrent network with feedback from hidden layer and output layer as additional input series. The finite element model of a CNC center spindle was built, and a new method for calculating the heat transfer convection coefficient on the surface of the spindle was proposed in this paper. The results demonstrated that the modified EN model has high level of robustness and accuracy. And then an experiment was implemented on a CNC machine tool. It can be found from the experimental results that the modified EN model could effectively predict the thermal error very well in practice and has higher level of prediction accuracy than RBF network model.

Acknowledgment This work is supported by the NC Key Foundation of China Government (No. 2009ZX04001-054-02). The authors are very appreciative of the help of key laboratory of School of Mechanical Engineering of Tianjin University.

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Krulewich DA (1998) Temperature integration model and measurement point selection for thermally induced machine tool errors. Mechatronics 8:395–412
2. Chen JS, Yuan JX, Ni J (1996) Thermal error modeling for real-time error compensation. Int J Adv Manuf Technol 12(4):266–275
3. Yan JY, Yang JG (2009) Application of synthetic grey correlation theory on thermal point optimization for machine tool thermal error compensation. Int J Adv Manuf Technol 43:1124–1132
4. Donmez MA, Blomquist RJ, Hocken RJ, Liu CR, Barash MM (1986) A general methodology for machine tool accuracy enhancement by error compensation. Precis Eng 8(4):87–195
5. Chen JS, Yuan JX, Ni J, Wu SM (1993) Real-time compensation for time-variant volumetric errors on a machining center. J Eng Ind 115:472–479
6. Yang H, Ni J (2003) Dynamic modeling for machine tool thermal error compensation. J Manuf Sci Eng 125:245–254
7. Wang Y, Zhang G, Moon KS, Sutherland JW (1998) Compensation for the thermal error of a multi-axis machining center. J Mater Process Technol 75:45–53
8. Li YX, Yang JG, Gelvis T, Li YY (2008) Optimization of measuring points for machine tool thermal error based on grey system theory. Int J Adv Manuf Technol 35:745–750
9. Choi H-S, Lee S-K (2002) Machining error compensation of external cylindrical grinding using thermally actuated rest. Mechatronics 12:643–656
10. Hattori M, Noguchi H, Ito S, Suto T, Inoue H (1996) Estimation of thermal deformation in machine tools using neural network technique. J Mater Process Technol 56:765–772
11. Veldhuis SC, Elbestawi MA (1995) A strategy for the compensation of errors in five-axis machining. CIRP Ann-Manuf Techn 44(1): 373–377
12. Kang Y, Chang CW, Huang Y, Hsu CL, Nieh I-F (2007) Modification of a neural network utilizing hybrid filters for the

compensation of thermal deformation in machine tools. Int J Mach Tools Manuf 47:376–387
13. Zhao HT, Yang JG, Shen JH (2007) Simulation of thermal behavior of a CNC machine tool spindle. Int J Mach Tools Manuf 47:1003–1010
14. Xu M, Jiang SY, Cai Y (2007) An improved thermal model for machine tool bearings. Int J Mach Tools Manuf 47:53–62
15. Pahk HJ, Lee SW (2002) Thermal error measurement and real time compensation system for the CNC machine tools incorporating the spindle thermal error and the feed axis thermal error. Int J Adv Manuf Technol 20:487–494
16. Hecht-Nielsen R (1987) Kolmogorov's mapping neural network existence theorem. In IEEE First International Conference on Neural Networks, vol. 3. New York, pp 11–14