

Research Article

A Study on the Generalized Approximation Modeling Method Based on Fitting Sensitivity for Prediction of Engine Performance

Lin Lin, Fang Wang, and Shisheng Zhong

School of Mechatronics Engineering, Harbin Institute of Technology, Harbin, China

Correspondence should be addressed to Lin Lin; waiwaiyl@163.com

Received 9 November 2016; Revised 13 January 2017; Accepted 5 February 2017; Published 6 March 2017

Academic Editor: Zhan Zhou

Copyright © 2017 Lin Lin et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Prediction technology for aeroengine performance is significantly important in operational maintenance and safety engineering. In the prediction of engine performance, to address overfitting and underfitting problems with the approximation modeling technique, we derived a generalized approximation model that could be used to adjust fitting precision. Approximation precision was combined with fitting sensitivity to allow the model to obtain excellent fitting accuracy and generalization performance. Taking the Grey model (GM) as an example, we discussed the modeling approach of the novel GM based on fitting sensitivity, analyzed the setting methods and optimization range of model parameters, and solved the model by using a genetic algorithm. By investigating the effect of every model parameter on the prediction precision in experiments, we summarized the change regularities of the root-mean-square errors (RMSEs) varying with the model parameters in novel GM. Also, by analyzing the novel ANN and ANN with Bayesian regularization, it is concluded that the generalized approximation model based on fitting sensitivity can achieve a reasonable fitting degree and generalization ability.

1. Introduction

Prediction technology for aeroengine performance is significantly important in operational maintenance and safety engineering. In January 2008, Australia's Qantas Airlines Boeing 747-400 power system malfunctioned in mid-flight, and four engines were all unusable [1]. In America, transport airplane accidents due to mechanical malfunctions were investigated, which covered 7,571 registered airplanes from 1980 to 2001. It is concluded that landing gear and turbine engine are most likely broken. Besides that, in commercial aircraft industry, \$31B is spent worldwide on aircraft maintenance in 2007, but 31% of it is on engine maintenance [2]. So, prediction technology for aeroengine performance can decrease the possibility of plane crash and save the cost of maintenance.

Aeroengine is complicated, nonlinear equipment, because it includes a large number of parts and there is a high degree of coupling. Also, performance parameters of aeroengine consist of many great discrete and nonlinear points. Consequently, this complexity brings about great difficulty to the development of prediction technology.

There are many methods to predict aeroengine performance. Approximate mathematical models based on data-driven approaches are acceptable substitutes for accurate physical models that are unlikely obtained in fields of multi-disciplinary design optimization, prediction, and so on. Many approximation models with random and nonlinear features have been constructed, and they include the response surface model (RS) [3, 4], polynomial regression model [5, 6], autoregressive moving average (ARMA) [7, 8], artificial neural network model (ANN) [9, 10], support vector machines (SVM) [11, 12], hidden Markov model (HMM) [13, 14], and Grey model (GM) [15, 16]. These approximation models are widely used in nonlinear simulation, classification, regression, and other domains. In the training phase, overfitting reduces the fitting errors but it causes the generalization performance of the model to decline. By contrast, underfitting causes models to veer away from true models with a relatively low fitting accuracy.

The overfitting problem has been studied by many scholars from different perspectives. To measure the fitting degree,

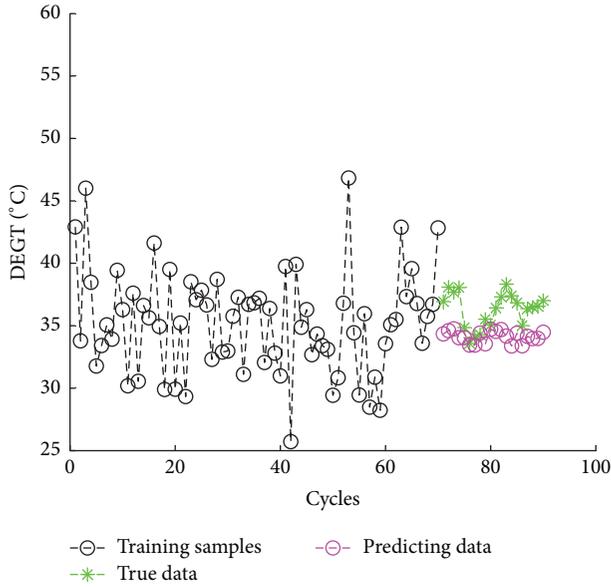


FIGURE 1: Prediction results of ANN with regularization.

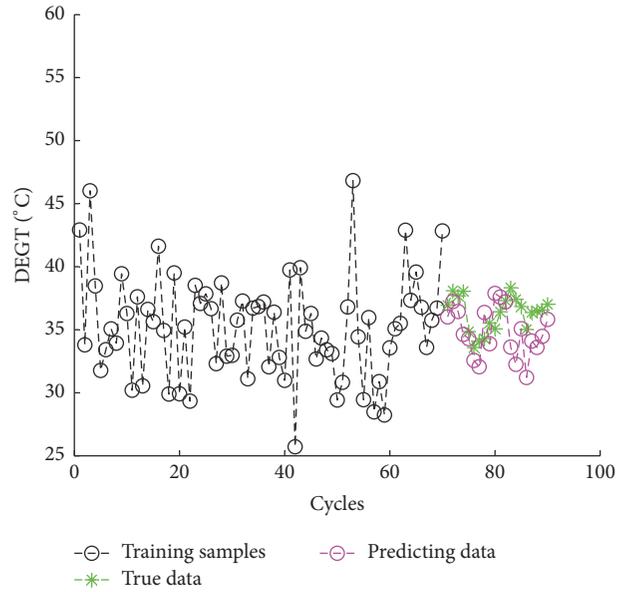


FIGURE 2: Prediction results of SVM with cross validation.

Akaike proposed an entropy-based information criterion (i.e., Akaike information criterion or AIC) with consideration of the complexity and fitting precision of models. A small AIC is indicative of good model performance and has thus been widely studied by many authors [17, 18]. For the sake of decreasing the high complexity of models, the Bayesian information criterion is studied on the basis of the AIC [19, 20]. Utilizing existing regularization methods [21, 22] is another method to avoid overfitting. With this approach, optimal parameters can be obtained to simplify models by adjusting either the L1 norm or the L2 norm of the weight coefficients [23]. Also, adjusting algorithm parameters, dimensionality reduction, and cross validation experiments [24] are effective means to avoid the overfitting problem.

Some examples in prediction domain are shown with the same EGT data (Exhaust Gas Temperature) as follows. Prediction results with an ANN with Bayesian regularization algorithm are seen in Figure 1, prediction results with a SVM with cross validation are seen in Figure 2, and prediction results with Grey model (GM) are seen in Figure 3.

Although regularization method is used in ANN and SVM, error precision ought to be better. GM is usually applied to predict parameters with noise, but error precision should be improved too.

To obtain high prediction precision, overfitting and underfitting problems must be addressed better. In addressing fitting problems, we set up a fitting sensitivity model to make the approximation model insensitive to samples that are far from the core of sample clusters and sensitive to samples that are close to the core of sample clusters. The fitting sensitivity model is achieved by controlling the fitting sensitivity to adjust the fitting accuracy of the approximation model. In using the fitting sensitivity model, the approximation model approaches the center of the sample clusters instead of the samples far from the center of the sample clusters. Hence, the

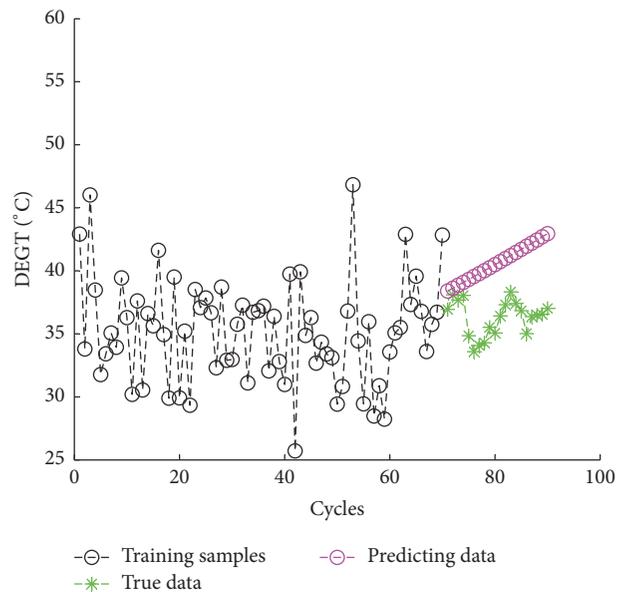


FIGURE 3: Prediction results of GM.

robustness and adaptability of the model to new samples are improved.

In Section 2, an approximation modeling method based on fitting sensitivity is introduced. With this method, a generalized approximation model is built by analyzing the fitting sensitivity and its correlation with the approximating precision. In Section 3, a GM is taken as an example and the approximation modeling method based on the fitting sensitivity is employed to gain a novel GM that could avoid overfitting and underfitting problems better. This novel GM exhibits such capability because the main tendency of its training samples can be obtained through fitting, which can reduce the fluctuations in prediction results. In Section 4,

values of parameters for the specific model proposed in Section 3 are set, and the optimization model solved with a genetic algorithm is established with a reasonable fitting degree. In Section 5, experimental verification with a single variable is performed to analyze the impacts of the model parameters on the accuracy of novel GM. The contrast experiment between ANN with Bayesian regularization and the novel ANN showed that the approximation model based on fitting sensitivity can yield better prediction results in comparison with traditional models.

2. Modeling Methods of the Generalized Approximation Model Based on Fitting Sensitivity

Firstly, fitting sensitivity is introduced and then the relationship between fitting degree and fitting sensitivity is illustrated. Secondly, generalized approximation model is constructed based on fitting sensitivity. Also, fitting error analysis for the model is derived. Lastly, a novel GM based on fitting sensitivity is studied.

2.1. Fitting Sensitivity Analyses. Training samples are represented as $X = [x_1, x_2, \dots, x_n]$ and fitting values are represented as $Y = [y_1, y_2, \dots, y_n]$. The fitting degree can be expressed by the fitting sensitivity model dY/dX , when there is a same initial value in both Y and X ; that is, $x_1 = y_1$, and $dY/dX > 0$.

The corresponding analysis is presented as follows:

- (1) When $dY/dX \rightarrow 1$, there is overfitting of y_k to x_k ; that is, $\Delta Y \approx \Delta X$. More specifically, the changing tendency of the fitting value y_k is in accordance with that of x_k , as shown in Figure 4(a).
- (2) When $dY/dX > 1$, there is underfitting of y_k to x_k with $\Delta Y > \Delta X$, as shown in Figure 4(b). Moreover, the fitting value y_k enlarges the trend of the training sample x_k . In this case, y_k is unstable and is fluctuating along with x_k . Consequently, inaccurate prediction results are got. This condition is called "excessive underfitting."
- (3) When $0 < dY/dX < 1$, there is underfitting of y_k to x_k with $\Delta Y < \Delta X$, as shown in Figure 4(c). Moreover, the fitting value y_k compresses the trend of the training sample x_k . In this case, y_k is close to the main trend of x_k . Consequently, accurate prediction results are got. This condition is called "reasonable underfitting."

In conclusion, when the initial value of X is the same as that of Y , the different levels of overfitting and underfitting correspond to the different values of the fitting sensitivity dY/dX . In particular, the reasonable fitting degree is obtained by setting dY/dX in the interval $(0, 1)$, which can make the model avoid overfitting and excessive underfitting in some degree.

2.2. Generalized Approximation Model Based on Fitting Sensitivity. Setting dY/dX in the interval $(0, 1)$, a fitting sensitivity model is built as follows when $x_1 = y_1$:

$$\frac{dy_k}{dx_k} = \frac{1}{a + b \cdot (x_k - y_k)^2}, \quad (1)$$

where a is coefficient of whole compressibility to x_k ; $a > 1$. b is coefficient of compressibility of $|x_k - y_k|$ to x_k .

The description of (1) is as follows.

(1) $|x_k - y_k|$ becomes large, which means the fitting value y_k is far from x_k . Because the fitting value y_k is required to represent the average level of the training samples, strong noise and violent fluctuations are included in x_k . To obtain a gentle main trend in y_k , the sensitivity of y_k to x_k , dy_k/dx_k , should be decreased.

(2) $|x_k - y_k|$ becomes small, which means the fitting value y_k is close to x_k . Because the fitting value y_k is required to represent the average level of the training samples, the changing trends of x_k are gentle. To maintain this gentle main trend, the changing trend of y_k should follow that of x_k . In this way, the sensitivity of y_k to x_k , dy_k/dx_k , should be increased.

Equation (1) is transformed into the following integral equation:

$$y_k = \int \frac{1}{a + b \cdot (x_k - y_k)^2} dx_k. \quad (2)$$

The integral variable is transformed into

$$y_k = \frac{1}{\sqrt{a \cdot b}} \int \frac{1}{1 + (b/a) \cdot (x_k - y_k)^2} d \left(\sqrt{\frac{b}{a}} \cdot (x_k - y_k) \right) \quad (3)$$

$$y_k = \frac{1}{\sqrt{a \cdot b}} \cdot \left[\arctan \left(\sqrt{\frac{b}{a}} \cdot (x_k - y_k) \right) + c_1 \right], \quad (4)$$

$k = 1, \dots, n,$

where c_1 is integral offset.

Equation (4) is the implicit expression of the approximation model based on fitting sensitivity. The generalized approximation model is obtained by transforming (4) as follows:

$$x_k = y_k + \sqrt{\frac{a}{b}} \cdot \tan \left(y_k \cdot \sqrt{a \cdot b} - c_1 \right). \quad (5)$$

The first item in (5) is the fitting value y_k of the approximation model to the training sample x_k . The second item is $\sqrt{a/b} \cdot \tan(y_k \cdot \sqrt{a \cdot b} - c_1)$, which denotes the adjustment of the traditional prediction model.

The approximation model mentioned above can avoid overfitting. However, the constraint equation needs to be added to avoid "excessive underfitting," as shown in Figure 5.

If $|x_k - y_k| \rightarrow +\infty$, then $dy_k/dx_k \rightarrow 0$, which inevitably leads to a large underfitting degree of y_k to x_k , "excessive

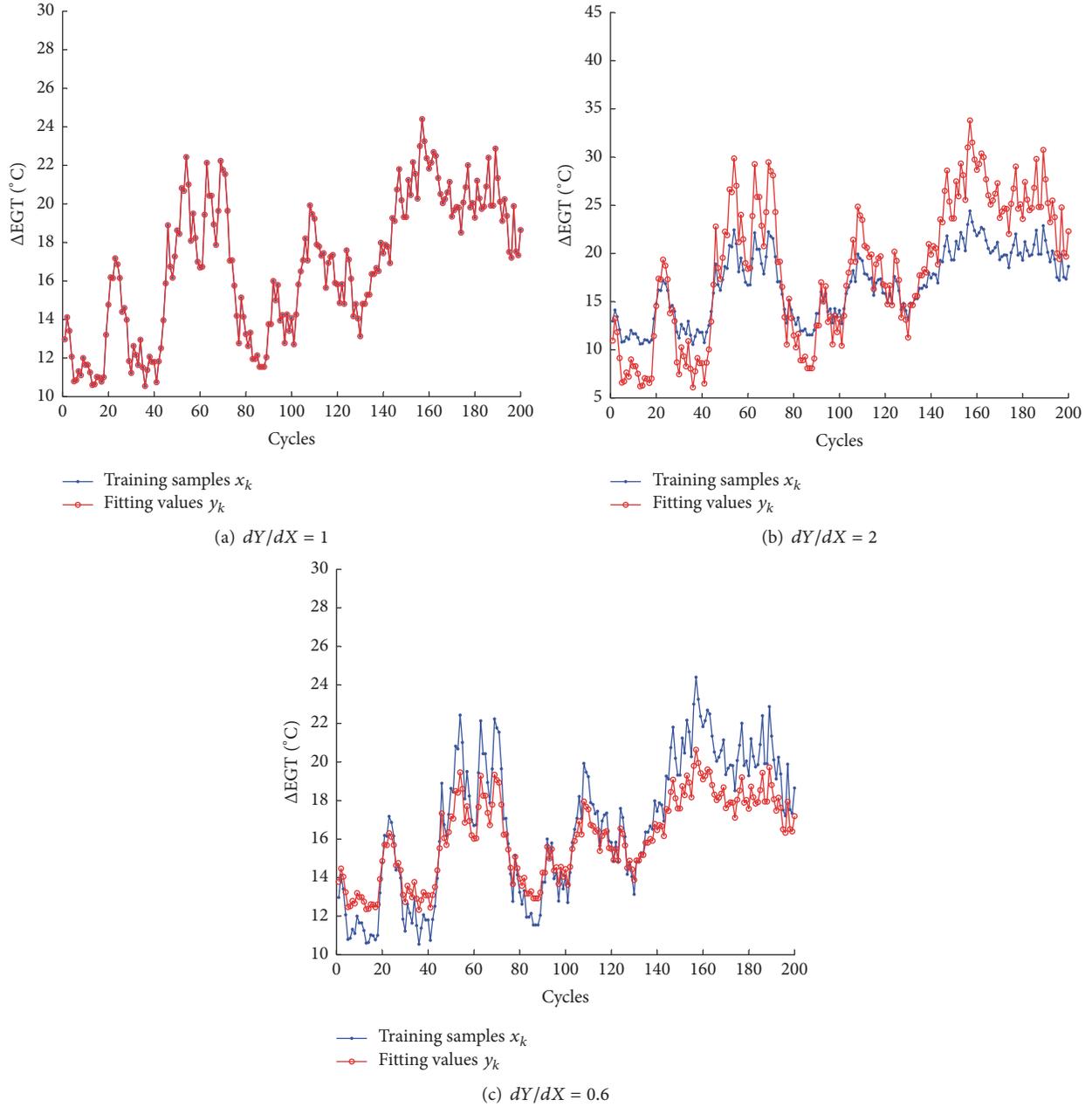


FIGURE 4: Position relations between fitting value y_k and training sample x_k under different fitting degrees.

underfitting.” Consequently, the lower bounds for dy_k/dx_k should be set up to avoid excessive underfitting as follows:

$$\frac{1}{a + b \cdot (x_k - y_k)^2} > \frac{1}{a} - \varepsilon, \quad k = 1, \dots, n, \quad (6)$$

where ε is adjusting coefficient and $\varepsilon \in (0, 1/a)$.

When $\varepsilon \rightarrow 0$, $|x_k - y_k| \rightarrow 0$; when $\varepsilon \rightarrow 1/a$, $|x_k - y_k| \rightarrow +\infty$. Thus, the coefficient ε should be changed to control $|x_k - y_k|$ in a range. The definition domain of (6) is whole time domain; that is, $k = 1, \dots, n$. However, when the number of training samples is too large, it is difficult to maintain $|x_k - y_k|$

less than a small number Δ at any point. As samples which are close to the prediction moment play an important role in improving forecast precision, the definition domain of (6) is controlled into the last p points of the training samples. That is, (6) is effective when $k = n - p + 1, n - p + 2, \dots, n$ (see the following equation):

$$\frac{1}{a + b \cdot (x_k - y_k)^2} > \frac{1}{a} - \varepsilon, \quad (7)$$

$$k = n - p + 1, n - p + 2, \dots, n,$$

where ε is adjusting coefficient and $\varepsilon \in (0, 1/a)$.

Equation (5) is substituted into (7) to finally obtain the generalized approximation model based on fitting sensitivity as follows:

$$x_k = y_k + \sqrt{\frac{a}{b}} \cdot \tan(y_k \cdot \sqrt{a \cdot b} - c_1) \quad (8)$$

$$\text{s.t. } \frac{1}{a + a \cdot \tan^2(y_k \cdot \sqrt{a \cdot b} - c_1)} > \frac{1}{a} - \varepsilon, \quad (9)$$

$$k = n - p + 1, n - p + 2, \dots, n,$$

where y_k is fitting value at the point k .

2.3. Fitting Error Analysis for the Generalized Approximation Model Based on Fitting Sensitivity. The fitting error D is to be analyzed between the fitting values and training samples.

$$D = \sum_{k=1}^n (x_k - y_k)^2. \quad (10)$$

The following is obtained from (8):

$$D = \sum_{k=1}^n \left(\sqrt{\frac{a}{b}} \cdot \tan(y_k \cdot \sqrt{a \cdot b} - c_1) \right)^2 \quad (11)$$

$$= \sum_{k=1}^n \left| \frac{a}{b} \right| \cdot \tan^2(y_k \cdot \sqrt{a \cdot b} - c_1).$$

As shown in (11), when $y_1 = y_2 = \dots = y_n = c_1/\sqrt{a \cdot b}$, $D = 0$. That is, when the fitting value y_k is equal to a constant $c_1/\sqrt{a \cdot b}$, the fitting error is zero. However, y_k actually varies with x_k . Thus, D must be a numeric value greater than a small number δ ; that is, $D > \delta$.

The following is obtained from (9):

$$a + a \cdot \tan^2(y_k \cdot \sqrt{a \cdot b} - c_1) < \frac{1}{1/a - \varepsilon}$$

$$\tan^2(y_k \cdot \sqrt{a \cdot b} - c_1) < \frac{1}{1 - a \cdot \varepsilon} - 1$$

$$\tan^2(y_k \cdot \sqrt{a \cdot b} - c_1) < \frac{a \cdot \varepsilon}{1 - a \cdot \varepsilon} \quad (12)$$

$$D = \sum_{k=1}^n \left| \frac{a}{b} \right| \cdot \tan^2(y_k \cdot \sqrt{a \cdot b} - c_1)$$

$$< \sum_{k=1}^n \left| \frac{a}{b} \right| \cdot \frac{a \cdot \varepsilon}{1 - a \cdot \varepsilon}.$$

The lower and upper bounds of the fitting error D are obtained as follows:

$$\delta < D < \sum_{k=1}^n \left| \frac{a}{b} \right| \cdot \frac{a \cdot \varepsilon}{1 - a \cdot \varepsilon}. \quad (13)$$

The model controls the fitting error D in a certain range, as shown in (13). Then, the model can avoid the overfitting and underfitting problems to some extent.

2.4. Modeling Methods of a Novel GM Based on Fitting Sensitivity. The GM is suitable for predicting the time series of performance parameters with great randomness. The GM can reduce accumulated errors and the fluctuations of prediction results by accumulated summation. Thus, we introduce a modeling technique for the approximation model based on fitting sensitivity by taking the GM as an example.

Overfitting and underfitting problems also exist in the traditional GM during the training phase, similar to other approximation models. As shown in (14), the developed coefficient \hat{a} and Grey-controlled variable \hat{u} in the GM are derived from the least squares method.

$$x_k = \left(x_1 - \frac{\hat{u}}{\hat{a}} \right) \cdot (e^{-\hat{a}} - 1) \cdot e^{-\hat{a} \cdot (k-2)}, \quad (14)$$

where \hat{a} is developed coefficient, \hat{u} is Grey-controlled variable, and x_1 is initial value of the training samples.

To address the fitting problem, model parameters \hat{a} and \hat{u} are evaluated with the approximation model based on fitting sensitivity instead of the least squares method to effectively avoid overfitting and excessive underfitting in the training phase and ultimately gain the precise estimation parameters \hat{a} and \hat{u} . We can then construct the novel GM based on fitting sensitivity by using \hat{a} and \hat{u} , as mentioned in Section 2.2.

The novel GM based on fitting sensitivity is written as follows:

$$x_k = \left(x_1 - \frac{\hat{u}}{\hat{a}} \right) \cdot (e^{-\hat{a}} - 1) \cdot e^{-\hat{a} \cdot (k-2)} + \sqrt{\frac{a}{b}} \cdot \tan(y_k \cdot \sqrt{a \cdot b} - c_1) \quad (15)$$

$$\text{s.t. } \frac{1}{a + a \cdot \tan^2(y_k \cdot \sqrt{a \cdot b} - c_1)} > \frac{1}{a} - \varepsilon, \quad (16)$$

$$k = n - p + 1, n - p + 2, \dots, n.$$

3. Solution of Novel GM Based on Fitting Sensitivity

The model parameters in (15) and (16) include the adjusting coefficients a, b, c_1 that can avoid the overfitting problem, the adjusting coefficients ε, p that can avoid the excessive underfitting problem, and the shape parameters \hat{a} and \hat{u} . Then, a, b, ε, p are set according to their physical meaning, and c_1, \hat{a}, \hat{u} are set with the constraint condition of avoiding overfitting and excessive underfitting. Using these parameters, the optimization model based on fitting degree is solved with the genetic algorithm.

3.1. Parameter Setting of the Novel GM. a, b, ε, p are set to a reasonable range according to their physical meaning in the novel GM. More specifically, the last p training samples are used to satisfy the constraint condition to improve the prediction accuracy, ε is employed to avoid excessive underfitting in the training phase, and a, b are adjusting coefficients to avoid overfitting in the model.

3.1.1. Setting of Parameter p . Excessive underfitting is challenging to be avoided during the whole training phase because of the presence of a large number of training samples. Thus, only the last p training samples are used to satisfy the constraint condition to improve the prediction accuracy, as mentioned in Section 2.2. The value of p ranges from 1 to n , that is, the length of the training samples; that is, $p \in [1, n]$, and $p \in N$.

When $p \rightarrow 1$, a few training samples are constrained to avoid excessive underfitting, but these training samples do not contain the trend of x_k . Inaccurate prediction results are thus obtained.

When $p \rightarrow n$, it is difficult to make all training samples be constrained to avoid excessive underfitting. Even with the avoidance of excessive underfitting, inaccurate prediction results are still obtained because of the unrelated historical information on the training samples.

As a result, p is set up to reasonable range from 1 to n . The last p training samples are then selected to be constrained, including the development tendency of the training samples instead of the unrelated historical information.

3.1.2. Setting of Parameter ε . The parameter ε is employed to avoid excessive underfitting in the training phase when $k = n - p + 1, n - p + 2, \dots, n$, making the fitting sensitivity $dy_k/dx_k > 1/a - \varepsilon$, $\varepsilon \in (0, 1/a)$ at the last p training samples.

When $\varepsilon \rightarrow 0$, $dy_k/dx_k \rightarrow 1/a$, as shown in (16), which results in high fitting precision and the high sensitivity of the fitting value y_k to x_k when $k = n - p + 1, n - p + 2, \dots, n$; when $\varepsilon \rightarrow 1/a$, $dy_k/dx_k \rightarrow 0$, which results in the low sensitivity of the fitting value y_k to x_k when $k = n - p + 1, n - p + 2, \dots, n$ and the inaccurate fitting values at the last p training samples.

As a result, ε is set up to reasonable range from 0 to $1/a$ on the basis of the sensitivity of the fitting value y_k to x_k at the last p training samples; that is, $k = n - p + 1, n - p + 2, \dots, n$.

3.1.3. Setting of Parameters a, b . The adjusting coefficients to avoid overfitting are a , which makes the fitting sensitivity $dy_k/dx_k < 1$, and b , which is the compressibility of x_k affected by $|x_k - y_k|$ to make y_k approach the main trend of x_k and reduce the fluctuation of prediction results, as shown in (16).

The adjustment of a can avoid the overfitting of y_k to x_k . Overfitting means $\max |y_k - x_k| \rightarrow 0$. At the same time, $dy_k/dx_k = 1/a + b \cdot (x_k - y_k)^2 \rightarrow 1/a$, as mentioned in

Section 2.1. That is, a fitting sensitivity that is less than 1 can avoid the overfitting problem. So, $1/a < 1$; that is, $a > 1$. However, when $a \rightarrow +\infty$, $dy_k/dx_k = 1/a + b \cdot (x_k - y_k)^2 \rightarrow 0$, and y_k is unaffected by x_k , thus leading to a meaningless fitting process. Hence, the study controls a into $(1, 1/\varepsilon)$ with consideration of $1/a - \varepsilon > 0$.

The adjustment of b can make y_k approach the main trend of x_k . As mentioned in Section 2.2, the fitting value y_k represents the average level of the training samples. Thus, when $|x_k - y_k| < \Delta$ (Δ is a threshold), y_k exhibits high sensitivity to x_k ; when $|x_k - y_k| > \Delta$, y_k exhibits low sensitivity to x_k . In other words, y_k can obtain the main trend of x_k with consideration of the two points above.

The lowest fitting sensitivity is $1/a - \varepsilon$, which is obtained from the constraint in (15). When the distance between y_k and x_k reaches the maximum, that is, $|x_k - y_k| = \Delta$, the minimum fitting sensitivity dy_k/dx_k can be obtained as follows:

$$\frac{1}{a + b \cdot \Delta^2} = \frac{1}{a} - \varepsilon. \quad (17)$$

We obtain b from (17).

$$b = \frac{a \cdot \varepsilon}{(1/a - \varepsilon) \cdot \Delta^2}. \quad (18)$$

The parameter b can be well defined after setting the threshold Δ by taking the maximum difference between the neighboring points, that is, $M = \max_k |x_k - x_{k-1}|$, in the training samples as the unit. In engineering practice, when $|x_k - y_k| > s \cdot M$ (s is a constant), the fitting value y_k is far from x_k , with the fitting being ineffective and with dy_k/dx_k exhibiting low sensitivity. The study sets $s = 5$, which means that when the distance between y_k and x_k is over five times M , it is a meaningless fitting, that is, when $\Delta = 5 \max_k |x_k - x_{k-1}|$.

3.2. Adjustment of Parameters Based on the Genetic Algorithm.

After the setting of a, b, ε, p based on their physical meanings, other parameters, c_1, \hat{a}, \hat{u} , are solved with the genetic algorithm with consideration of the constraint for overfitting and underfitting.

3.2.1. Establishment of the Optimization Model. The first step in the design of the genetic algorithm is to establish the optimization model, and the key to construct the model is to build the adaptive function and nonlinear constraint.

Step 1 (build an adaptive function). The adaptive function is the objective function of the optimization model which is used to decide the seeking direction of the group. As the novel GM based on fitting sensitivity can avoid the overfitting problem, the adaptive function is transferred from (15) as follows:

$$\min \sum_{k=1}^n \left[x_k - \left(x_1 - \frac{\hat{u}}{\hat{a}} \right) \cdot (e^{-\hat{a}} - 1) \cdot e^{-\hat{a}(k-2)} - \sqrt{\frac{a}{b}} \cdot \tan \left(y_k \cdot \sqrt{a \cdot b} - c_1 \right) \right]^2. \quad (19)$$

Step 2 (construct the constraint condition). The adaptive function considers the overfitting; it is to be supplemented with the constraint for the excessive underfitting problem.

Finally, the optimization model solved with the genetic algorithm is obtained as follows.

Find \hat{a}, \hat{u}, c_1 :

$$\min \sum_{k=1}^n \left[x_k - \left(x_1 - \frac{\hat{u}}{\hat{a}} \right) \cdot (e^{-\hat{a}} - 1) \cdot e^{-\hat{a}(k-2)} - \sqrt{\frac{\hat{a}}{\hat{b}}} \cdot \tan(y_k \cdot \sqrt{\hat{a} \cdot \hat{b}} - c_1) \right]^2 \quad (20)$$

$$\text{s.t. } \frac{1}{a + a \cdot \tan^2(y_k \cdot \sqrt{a \cdot b} - c_1)^2} > \frac{1}{a} - \varepsilon, \quad k = n - p + 1, n - p + 2, \dots, n. \quad (21)$$

3.2.2. Design of the Initial Value of the Genetic Variable. The design of the genetic operators is very mature and is not repeated here. On the basis of the requirements of this study, we need to set up the initial population of \hat{a}, \hat{u}, c_1 as the optimization variables, as shown in (20).

Because the methodology proposed in this study modifies the traditional GM, the initial values of the GM parameters \hat{a}, \hat{u} are set to the values obtained through the least squares method as shown in the following equation:

$$\hat{a} = \frac{(1/(n-1)) \cdot \sum_{k=2}^n x(k) \cdot \sum_{k=2}^n z(k) - \sum_{k=2}^n x(k) \cdot z(k)}{\sum_{k=2}^n (z(k))^2 - (1/(n-1)) \cdot (\sum_{k=2}^n z(k))^2} \quad (22)$$

$$\hat{u} = \frac{1}{n-1} \cdot \left[\sum_{k=2}^n x(k) + a \cdot \sum_{k=2}^n z(k) \right],$$

where $z(k) = (x^{(1)}(k) + x^{(1)}(k-1))/2$, $k = 2, 3, \dots, n$; $x^{(1)}(k) = \sum_{i=1}^k x(i)$, $k = 1, 2, \dots, n$.

The design of offset c_1 employed in the integration process is to be discussed. The fitting accuracy at the last p training samples should be improved during the training phase to achieve expected prediction precision. Thus, $\forall k \in [n-p+1, n]$ and $k \in N$, make c_{1k} satisfy (4); that is,

$$y_k = \frac{1}{\sqrt{a \cdot b}} \cdot \left[\arctan \left(\sqrt{\frac{b}{a}} \cdot (x_k - y_k) \right) + c_{1k} \right], \quad (23)$$

$$k = n - p + 1, n - p + 2, \dots, n.$$

The fitting value y_k of the GM is substituted into (23). Then, c_{1k} is obtained as follows:

$$c_{1k} = \sqrt{a \cdot b} \cdot \left(\left(x_1 - \frac{\hat{u}}{\hat{a}} \right) \cdot (e^{-\hat{a}} - 1) \cdot e^{-\hat{a}(k-2)} - \arctan \left(\sqrt{\frac{b}{a}} \cdot (x_k - y_k) \right) \right), \quad (24)$$

where $k = n - p + 1, n - p + 2, \dots, n$.

The initial value of c_1 is set to the mean of c_{1k} , $k = n - p + 1, n - p + 2, \dots, n$, as follows:

$$c_1 = \frac{1}{p} \cdot \sum_{k=n-p+1}^n c_{1k}. \quad (25)$$

4. Experimental Validation

Firstly, the experimental data is described. And then the different prediction results under different values of parameters a, n, p, ε in novel GM are analyzed. Lastly, the prediction precision of ANN with regularization method is compared with that of approximation ANN based on fitting sensitivity.

4.1. Experimental Data. The time series of performance parameter DEGT (the difference between the monitored Exhaust Gas Temperature and the benchmark) of an aeroengine over 200 cycles is used in the experiment. It is an important performance parameter of aeroengine. But significant randomness and serious fluctuation are observed in it as shown in Figure 6.

Randomicity and fluctuation in observed parameters bring about difficulties to predict DEGT. There are mainly four reasons which are given as follows. Firstly, random factors such as the actual working condition of equipment and human operation result in the fluctuation of the time series data, and thus the chosen approximation model cannot easily approach the original nonlinearization. Secondly, for the accumulation of errors caused by the iteration method for prediction, a long prediction phase equates to a large forecast deviation. Thirdly, when nonlinear models are aimed at high precision, overfitting occurs in the training phase with forecast deviation. Lastly, in the case involving many training samples during the training phase, the model parameters for approximating such samples are difficult to determine and model accuracy is adversely affected by the underfitting problem.

Real signals are difficult to find because of the presence of noise signals. So, the smoothed data approved by the aeroengine manufacturer are regarded as true values in evaluation of prediction accuracy in this work as shown in Figure 7.

4.2. Experiment Analysis of Parameters a, n, p, ε in Novel GM. To analyze the effect of parameters a, n, p, ε on novel GM precision, a single variable is maintained in turn to observe the effect of every variable on the prediction precision in novel Grey model. Values of a, n, p, ε are set to their certain range. And then prediction errors are compared and analyzed varying with a, n, p, ε in turn. Besides, 20 points are to be predicted in all experiments.

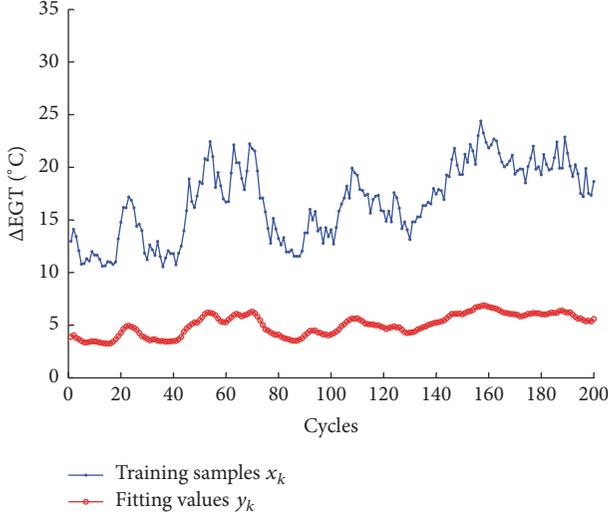


FIGURE 5: Position relation between fitting value y_k and training sample x_k under excessive underfitting.

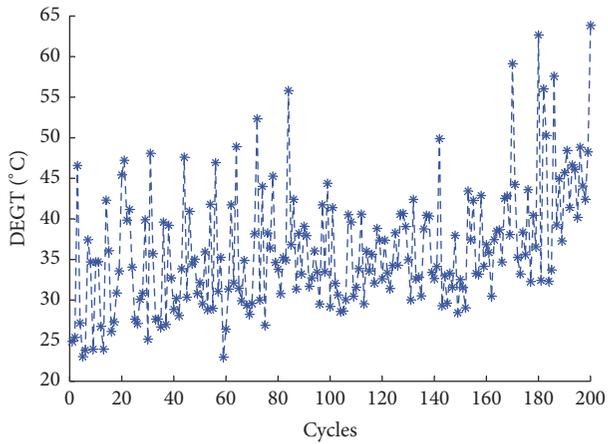


FIGURE 6: DEGT of an aeroengine over 200 cycles.

4.2.1. Effect of Parameter a on Model Precision. To study the effect of the single variable a on the prediction precision, we set $n = 70$, $p = 10$, and $\varepsilon = 0.1$. Value range of a is $(1, 1/\varepsilon) = (1, 10)$, and a is set to 1.25, 3, 5, 7, and 9 in five groups in novel GM. 60 experiments are done in every a . The box plot of RMSEs at different values of a with the novel GM is shown in Figure 8. For comparison, RMSEs of traditional GM are drawn in Figure 8 too. There are seven indicators of RMSEs in box plot: max, maximum value; Q3, 75th percentile; median, the median of RMSEs; Q1, 25th percentile; min, minimum value; outlier, the number of outliers; and DQQ, distance between Q1 and Q3.

Values of indicators in box plot of RMSEs at different values of a with the novel GM and GM are shown in Table 1.

As shown in Table 1, the least value of RMSE median is 3.22, and the least DQQ is 2.93 with $a = 3$ in novel GM. That is, the prediction precision of novel GM gathers round 3.22, and the dispersion of errors is small when $a = 3$.

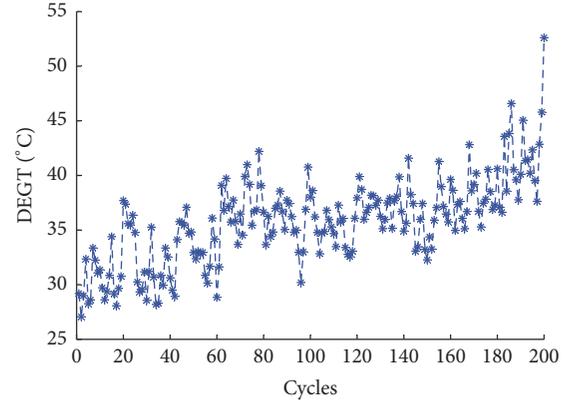


FIGURE 7: True values in evaluation of DEGT prediction accuracy.

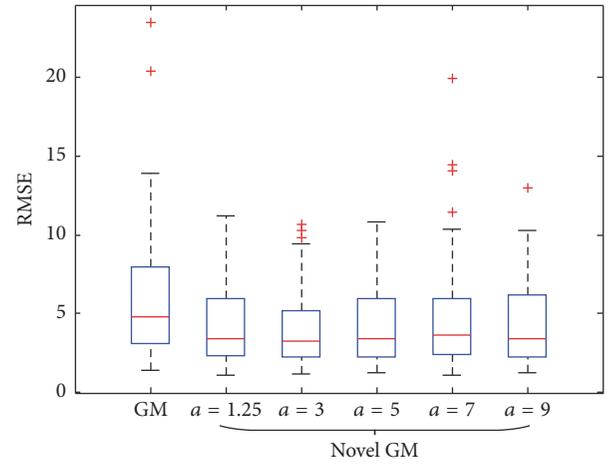


FIGURE 8: Box plot of RMSEs at different values of a in the novel GM and GM.

In the prediction model proposed in the study, the fitting degree in the training phase is partly decided by a . When $a \rightarrow 1$, overfitting is the possible problem, while when $a \rightarrow +\infty$, underfitting is the possible problem. So, there must be an a at the highest prediction precision. Here, $a = 3$.

Prediction results of the novel GM are better than the results of the traditional GM, because there are smaller median and DQQ, corresponding to higher precision and lower dispersion.

4.2.2. Effect of Parameter n on Model Precision. To study the effect of the single variable n on the prediction precision, we set $a = 3$, $p = 10$, and $\varepsilon = 0.1$. And n is set to 30, 40, 50, 60, 70, and 80 in six groups in novel GM. 60 experiments are done in every n . The box plot of RMSEs at different values of n is shown in Figure 9.

Values of indicators in box plot of RMSEs at different values of n with the novel GM and GM are shown in Table 2.

As shown in Table 2, the least value of RMSE median is 3.22, and the least DQQ is 2.93 with $n = 70$ in novel GM. That is, the prediction precision of novel GM gathers round 3.22, and the dispersion of errors is small when $n = 70$.

TABLE 1: Values of indicators in box plot of RMSEs at different values of a in the novel GM and GM.

Indicator	GM	Novel GM				
		$a = 1.25$	$a = 3$	$a = 5$	$a = 7$	$a = 9$
Max	13.93	11.20	9.42	10.8	10.37	10.24
Q3	7.96	5.94	5.18	5.91	5.95	6.17
Median	4.81	3.38	3.22	3.37	3.60	3.36
Q1	3.10	2.27	2.25	2.24	2.37	2.23
Min	1.37	1.09	1.10	1.19	1.05	1.18
Outlier	2	0	3	0	4	1
DQQ	4.86	3.67	2.93	3.67	3.58	3.94

TABLE 2: Values of indicators in box plot of RMSEs at different values of n in the novel GM and GM.

Indicator	GM	Novel GM					
		$n = 30$	$n = 40$	$n = 50$	$n = 60$	$n = 70$	$n = 80$
Max	13.93	10.83	11.11	10.95	11.91	9.42	11.61
Q3	7.96	6.73	6.22	6.19	6.21	5.18	6.26
Median	4.81	4.26	3.53	3.57	3.27	3.22	3.38
Q1	3.10	2.91	2.48	2.39	2.40	2.25	2.32
Min	1.37	1.17	1.08	1.10	1.12	1.10	1.40
Outlier	2	2	1	0	0	3	1
DQQ	4.86	3.82	3.74	3.8	3.81	2.93	3.94

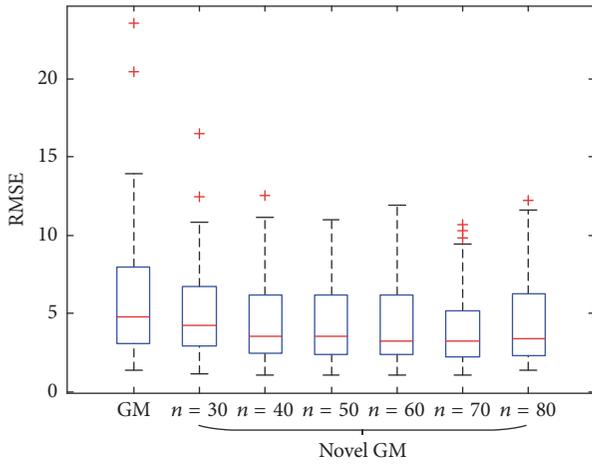


FIGURE 9: Box plot of RMSEs at different values of n in the novel GM and GM.

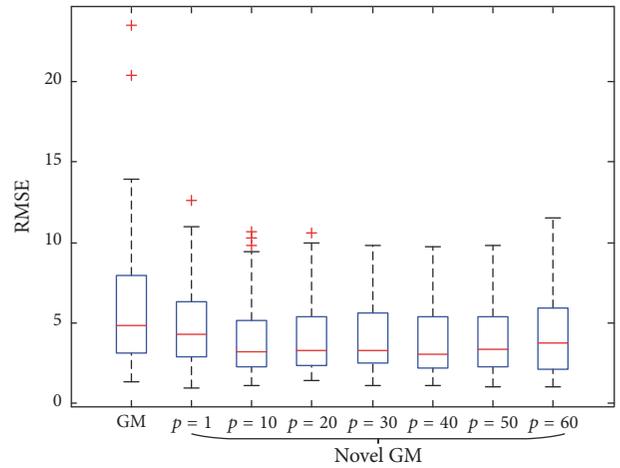


FIGURE 10: Box plot of RMSEs at different values of p in the novel GM and GM.

Excessively large n introduces too much unrelated information about the trend of the training samples, while small n does not include sufficient information. So, there must be an n at the highest prediction precision. Here, $n = 70$.

4.2.3. Effect of Parameter p on Model Precision. To study the effect of the single variable p on the prediction precision, we set $a = 3$, $n = 70$, and $\varepsilon = 0.1$. p is set to 1, 10, 20, 30, 40, 50, and 60 in seven groups with novel GM. 60 experiments are done in every p . The box plot of RMSEs at different values of p is shown in Figure 10.

Values of indicators in box plot of RMSEs at different values of p with the novel GM and GM are shown in Table 3.

As shown in Table 3, the least RMSE median is 3.05 with $p = 40$ and the least DQQ is 2.93 with $p = 10$ in novel GM. Prediction results of the novel GM are better than results of the traditional GM, because there are smaller median and DQQ, corresponding to higher precision and lower dispersion.

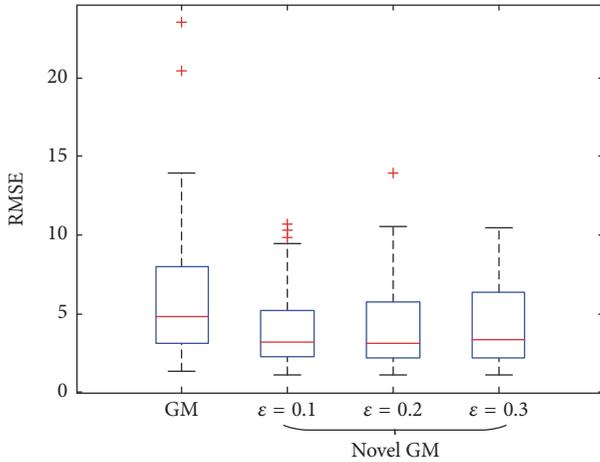
When $p \rightarrow 1$, a few training samples which do not contain the trend of y_k are constrained to avoid excessive underfitting. And when $p \rightarrow n$, all training samples which include the unrelated historical information are constrained to avoid excessive underfitting. So, there must be a p at the highest prediction precision.

TABLE 3: Values of indicators in box plot of RMSEs at different values of p in the novel GM and GM.

Indicator	GM	Novel GM						
		$p = 1$	$p = 10$	$p = 20$	$p = 30$	$p = 40$	$p = 50$	$p = 60$
Max	13.93	11.00	9.42	9.99	9.83	9.76	9.81	11.56
Q3	7.96	6.33	5.18	5.40	5.62	5.40	5.36	5.92
Median	4.81	4.32	3.22	3.26	3.28	3.05	3.39	3.71
Q1	3.10	2.86	2.25	2.34	2.53	2.20	2.30	2.14
Min	1.37	0.94	1.10	1.38	1.09	1.08	1.01	1.03
Outlier	2	1	3	1	0	0	0	0
DQQ	4.86	3.47	2.93	3.06	3.09	3.2	3.06	3.78

TABLE 4: Values of indicators in box plot of RMSEs at different values of ε in the novel GM and GM.

Indicator	GM	Novel GM		
		$\varepsilon = 0.1$	$\varepsilon = 0.2$	$\varepsilon = 0.3$
Max	13.93	9.42	10.54	10.46
Q3	7.96	5.18	5.71	6.35
Median	4.81	3.22	3.14	3.35
Q1	3.10	2.25	2.20	2.15
Min	1.37	1.10	1.12	1.14
Outlier	2	3	1	0
DQQ	4.86	2.93	3.51	4.2

FIGURE 11: Box plot of RMSEs at different values of ε in the novel GM and GM.

4.2.4. Effect of Parameter ε on Model Precision. To study the effect of the single variable ε on the prediction precision, we set $a = 3$, $n = 70$, and $p = 10$. And ε is set to 0.1, 0.2, and 0.3 in three groups in novel GM. 60 experiments are done in every ε . The box plot of RMSEs at different values of ε is shown in Figure 11.

Values of indicators in box plot of RMSEs at different values of ε with the novel GM and GM are shown in Table 4.

As shown in Table 4, the least value of RMSE median is 3.22, and the least DQQ is 2.93 with $\varepsilon = 0.1$ in novel GM.

Prediction results of the novel GM are better than the results of the traditional GM, because there are smaller median and DQQ, corresponding to higher precision and lower dispersion.

When $\varepsilon \rightarrow 0$, there is a high fitting degree in novel GM, and when $\varepsilon \rightarrow 1/a = 0.33$, the low sensitivity of the fitting value y_k to x_k at the last p training samples also could affect the prediction precision. So, there must be an ε at the highest prediction precision. Here, $\varepsilon = 0.1$.

4.2.5. Approximation Model of ANN. Experimental data in Section 4.1 is also used to be predicted with the approximation model of ANN with Bayesian regularization. For comparison, the ANN with Bayesian regularization algorithm is employed in the experiment.

Eight experiments are designed as shown in Table 5.

Prediction errors in novel ANN and ANN with regularization are shown in Figure 12 and Table 6.

By analyzing the data, we conclude that there is a smaller median in novel ANN than in ANN with regularization, but DQQ in novel ANN is not the smallest. So, higher prediction precision is obtained, but there is little big dispersion in novel ANN than in ANN with Bayesian regularization.

5. Conclusion

This study established a generalized approximation model based on fitting sensitivity to solve the overfitting and underfitting problems. Taking GM as an example, a novel GM based on fitting sensitivity was proposed. Then, the novel

TABLE 5: Eight experiments are designed.

Number	Model	a	n	p	ε
1	ANN with regularization	—	—	—	—
2	Novel ANN	3	70	10	0.1
3	Novel ANN	1.25	70	10	0.1
4	Novel ANN	2	70	10	0.1
5	Novel ANN	2	70	20	0.2
6	Novel ANN	2	70	30	0.3
7	Novel ANN	2	70	15	0.3
8	Novel ANN	2	70	25	0.3

TABLE 6: Values of indicators in box plot of RMSEs in novel ANN and ANN with Bayesian regularization.

Indicator	1	2	3	4	5	6	7	8
Max	9.39	9.86	10.87	10.38	8.69	9.21	10.43	9.37
Q3	5.63	6.35	5.93	5.92	5.37	5.98	6.30	5.46
Median	3.83	3.59	3.43	3.77	3.29	3.55	3.80	3.26
Q1	3.10	2.41	2.34	2.43	2.31	2.60	2.59	2.47
Min	1.06	1.16	1.15	1.11	1.30	1.25	1.22	1.26
Outlier	4	1	0	0	3	0	0	1
DQQ	2.53	3.94	3.59	3.49	3.06	3.38	3.71	2.99

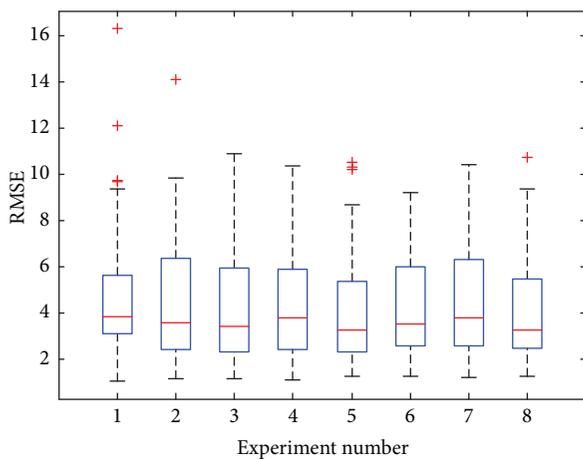


FIGURE 12: Box plot of RMSEs in the novel ANN and ANN with Bayesian regularization.

GM was solved with the genetic algorithm by establishing an optimization model restricted by a reasonable fitting degree. We used RMSE as criteria and compared the effects of the different model parameters of the novel model on the prediction precision in the experiments. It showed that novel GM and novel ANN can get higher prediction precision than traditional GM and ANN with regularization. Therefore, the novel model based on fitting sensitivity proposed in this work could avoid overfitting and underfitting and yield accurate prediction results in accordance with the theoretical analysis.

Competing Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgments

The authors are thankful for the support from the Key National Natural Science Foundation of China, China (no. U1533202), the support from the Shandong Independent Innovation and Achievements Transformation Fund, China (no. 2014CGZH1101), and the support from National Science-Technology Support Plan Project “the Application Paradigm of Full Lifecycle Information Closed-Loop Management for Construction Machinery” (no. 2015BAF32B01-4).

References

- [1] M. Pecht and K. Rui, *Diagnostics, Prognostics and System's Health Management*, PHM Centre, City University of Hong Kong, 2010.
- [2] C. R. Mercer, D. L. Simon, G. W. Hunter et al., “Fundamental technology development for gas-turbine engine health management,” in *Proceedings of the AIAA Infotech @ Aerospace Conference*, Rohnert Park, Calif, USA, May 2007.
- [3] D. Bař and İ. H. Boyacı, “Modeling and optimization I: usability of response surface methodology,” *Journal of Food Engineering*, vol. 78, no. 3, pp. 836–845, 2007.
- [4] V. Gunaraj and N. Murugan, “Application of response surface methodology for predicting weld bead quality in submerged arc welding of pipes,” *Journal of Materials Processing Technology*, vol. 88, no. 1, pp. 266–275, 1999.
- [5] T. Zhang, Q. Zhang, and Q. Wang, “Model detection for functional polynomial regression,” *Computational Statistics and Data Analysis*, vol. 70, pp. 183–197, 2014.
- [6] Y. Han, W. Liu, F. Bretz, F. Wan, and P. Yang, “Statistical calibration and exact one-sided simultaneous tolerance intervals for polynomial regression,” *Journal of Statistical Planning and Inference*, vol. 168, pp. 90–96, 2016.

- [7] R. H. Jones, "Maximum likelihood fitting of ARMA models to time series with missing observations," *Technometrics*, vol. 22, no. 3, pp. 389–395, 1980.
- [8] Y. Grenier, "Time-dependent ARMA modeling of nonstationary signals," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 31, no. 4, pp. 899–911, 1983.
- [9] Q. Zhang, X. Wei, and J. Xu, "On global exponential stability of discrete-time Hopfield neural networks with variable delays," *Discrete Dynamics in Nature and Society*, vol. 2007, Article ID 67675, 9 pages, 2007.
- [10] S. Agatonovic-Kustrin and R. Beresford, "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research," *Journal of Pharmaceutical and Biomedical Analysis*, vol. 22, no. 5, pp. 717–727, 2000.
- [11] Y. Bao, T. Xiong, and Z. Hu, "Forecasting air passenger traffic by support vector machines with ensemble empirical mode decomposition and slope-based method," *Discrete Dynamics in Nature and Society*, vol. 2012, Article ID 431512, 12 pages, 2012.
- [12] Z. Zhang and Q. Zhao, "The application of SVMs method on exchange rates fluctuation," *Discrete Dynamics in Nature and Society*, vol. 2009, Article ID 250206, 8 pages, 2009.
- [13] F. Ye and Y. Wang, "A novel method for decoding any high-order hidden Markov model," *Discrete Dynamics in Nature and Society*, vol. 2014, Article ID 231704, 6 pages, 2014.
- [14] S.-g. Yue, P. Jiao, Y.-b. Zha, and Q.-j. Yin, "A logical hierarchical hidden semi-Markov model for team intention recognition," *Discrete Dynamics in Nature and Society*, vol. 2015, Article ID 975951, 19 pages, 2015.
- [15] Y. Wang, Y. Dang, Y. Li, and S. Liu, "An approach to increase prediction precision of GM(1,1) model based on optimization of the initial condition," *Expert Systems with Applications*, vol. 37, no. 8, pp. 5640–5644, 2010.
- [16] N. Xie and S. Liu, "Discrete GM (1, 1) and mechanism of grey forecasting model," *Systems Engineering-Theory & Practice*, vol. 1, article 014, 2005.
- [17] H. Bozdogan, "Model selection and Akaike's information criterion (AIC): the general theory and its analytical extensions," *Psychometrika*, vol. 52, no. 3, pp. 345–370, 1987.
- [18] D. R. Anderson, K. P. Burnham, and G. C. White, "AIC model selection in overdispersed capture-recapture data," *Ecology*, vol. 75, no. 6, pp. 1780–1793, 1994.
- [19] C. T. Volinsky and A. E. Raftery, "Bayesian information criterion for censored survival models," *Biometrics*, vol. 56, no. 1, pp. 256–262, 2000.
- [20] M. Bogdan, J. K. Ghosh, and R. W. Doerge, "Modifying the Schwarz Bayesian information criterion to locate multiple interacting quantitative trait loci," *Genetics*, vol. 167, no. 2, pp. 989–999, 2004.
- [21] C. Flynn, C. M. Hurvich, and J. S. Simonoff, "Efficiency and consistency for regularization parameter selection in penalized regression: asymptotics and finite-sample corrections," NYU Working Paper no. 2451/31317, 2011.
- [22] J. Diebolt, M. Garrido, and C. Trottier, "Improving extremal fit: a Bayesian regularization procedure," *Reliability Engineering & System Safety*, vol. 82, no. 1, pp. 21–31, 2003.
- [23] J. Zhang, L. Peng, X. Zhao, and E. E. Kuruoglu, "Robust data clustering by learning multi-metric Lq-norm distances," *Expert Systems with Applications*, vol. 39, no. 1, pp. 335–349, 2012.
- [24] Z. Shao and M. J. Er, "Efficient leave-one-out cross-validation-based regularized extreme learning machine," *Neurocomputing*, vol. 194, pp. 260–270, 2016.



Hindawi

Submit your manuscripts at
<https://www.hindawi.com>

