Research Article

Performance Prediction of the Centrifugal Compressor Based on a Limited Number of Sample Data

Hongsheng Jiang [ID],[1] **Sujun Dong** [ID],[1] **Zheng Liu**,[2] **Yue He**,[1] **and Fengming Ai**[3]

[1]*School of Aeronautic Science of Engineering, Beihang University, Beijing 100191, China*
[2]*Shanghai Aircraft Design and Research Institute, COMAC, Shanghai 201210, China*
[3]*Shenyang Aircraft Design and Research Institute, Shenyang 110035, China*

Correspondence should be addressed to Sujun Dong; dsj@buaa.edu.cn

Academic Editor: Juan C. Jauregui-Correa

Centrifugal compressor is widely used in various engineering domains, and predicting the performance of a centrifugal compressor is an essential task for its conceptual design, optimization, and system simulation. For years, researchers seek to implement this mission through various kinds of methods, including interpolation, curve fitting, neural network, and other statistics-based algorithms. However, these methods usually need a large amount of data, and obtaining data may cost considerable computing or experimental resources. This paper focuses on constructing the performance maps of pressure ratio and isentropic efficiency using a limited number of sample data while maintaining accuracy. Firstly, sample data are generated from simulation using Vista CCD. Then, corrected flow rate and corrected rotational speed are used as independent variables, and the regression expressions with physical meaning of pressure ratio and isentropic efficiency are derived and simplified through thermodynamic analysis and loss analysis of centrifugal compressor, resulting in two loss-analysis-based models. Meanwhile, kriging models based on a second-order polynomial and neural network models are built. Results show that, when predicting inside data boundary, the loss-analysis-based model and the kriging model produce higher accuracy prediction even in a small data set, and the predicting result is stable, while the neural network model provides better results only in a more extensive data set with more speed lines. For the prediction outside the data boundary, the loss-analysis-based model can provide relatively accurate results. Besides, it takes less time to train and utilize a loss-analysis-based model than other models.

1. Introduction

Centrifugal compressor is a common type of turbomachine, widely applied in various engineering domains. To design, simulate, or optimize a centrifugal compressor, performance on off-design point is often required, including pressure ratio, efficiency, work, and other thermodynamic parameters. Therefore, predicting the performance of the centrifugal compressor is one of the critical steps in the calculation.

The desired performance parameters in various working status can be measured through a test bench. The data is then drawn into curves, called characteristic map or performance map, and can be used to predict the performance of compressors [1]. It is obvious that more test points lead to a more reliable characteristic map, but also consume more time and

resource. Computational fluid dynamics (CFD) method is another common way to obtain the performance of compressors [2]. However, in many cases, due to the limitation of time and computational cost, the result points are often too loose to cover the whole working area densely. Consequently, a suitable approach is needed to reconstruct the characteristic map, using a limited number of sample data, to provide high precision data at low cost efficiently.

The most intuitive way of predicting the performance of a compressor is to estimate by interpolation and fitting, but the precision may suffer from insufficient sample points. To improve the predicting quality, Kurzke and Riegler [3] introduced reference points and lines to scale the compressor maps, generating the most typical map topology. For fitting approaches, there are also various ways to improve, such as

data smoothing, rescaling, normalizing, and axes transforming. Further, El-Gammal [4] provided a transformation matrix of axes to preserve internal behaviour and cross-coupling. Also, the regression structure, such as third-order polynomial [5], logarithm [6], rotated elliptical curve [7, 8], and Chebyshev polynomial [9] will affect the result of fitting, which has been reported by many authors. However, these structures mainly depend on the mathematical analysis and the physical property in maps and can hardly be reflected. Therefore, it provides poor estimation when the desired point is far away from sample data.

Surrogate models, which mimic the behaviour of actual objects while being computationally cheaper to evaluate, are also used in forecasting the outcome of interest. Commonly used surrogate models include responsive surface, moving least squares, artificial neural network (ANN), and kriging model. Recently, with the growth of computing resource and development of theory, neural networks have been more and more widely used. Lazzaretto and Toffolo [10] trained the neural network using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) and Levenberg–Marquardt algorithm, with different transfer functions, yielding small errors on the thermodynamic parameters of compressors. Trained networks can be used for off-design performance prediction inside the training data boundary [11]. However, it often requires a large amount of sample data. Neural networks can also be applied in extrapolation and multilayer perceptron network can improve the quality of prediction [12]. Tian et al. [13] introduced the hybrid ANN-partial least squared (PLS) model into compressor performance prediction, producing better results than other structures of ANNs when enough data are provided [14]. However, reducing the total number of samples could rapidly drop the grade of prediction.

The kriging model is also commonly used in compressor performance prediction. It provides the best linear unbiased prediction of the intermediate values and reduces the computation complexity significantly [15]. It can also combine with different kinds of polynomial models to construct a more accurate predicting model for compressor characteristic [16].

Although the above-mentioned models are capable of performance prediction of compressors, they are mostly based on the mathematical or statistical analysis. Therefore, the discussion of the internal mechanism is missing, resulting in a limited ability on extrapolation. In order to improve the prediction precision of the centrifugal compressor performance using a limited number of data while maintaining extrapolation ability, this paper provides a loss-analysis-based regression model, which combines the classical curve fit method with compressor loss analysis.

In this paper, firstly, sample data are generated from simulation using Vista CCD. Next, a general investigation of centrifugal compressor thermodynamics is performed, followed by the study of the loss analysis. Then, the structures of loss-analysis-based regression models are determined by approximation and sensitivity analysis. In the meantime, kriging models based on the second-order polynomial and neural network models are constructed. Finally, the loss-analysis-based model will be applied to sample data, and compared with the kriging model and the neural network model, to examine the accuracy and calculation performance under different conditions.

2. Simulation Data Sets

Simulation is an efficient way to generate a bunch of data in a reasonable accuracy instead of time-consuming and high-cost experiments. In this paper, the sample data are produced by Vista CCD, which is a simulation program for the preliminary design of a centrifugal compressor with the capability of predicting performance maps. Required input settings for computing are listed in Table 1.

The calculation result of the Vista CCD is saved as a data table of efficiency, pressure ratio with mass flow rate, and rotational speed. Each table is composed of 17 rotational speed lines, and the speed is distributed evenly between 60,000 rpm and 140,000 rpm, numbered from 1 to 17. Each rotational speed line contains 25 points at equal intervals of flow rate, resulting in a total number of 425 data points. The characteristic map is shown in Figure 1. The simulation data are split into central set and extrapolation set, which will be used for further testing in Section 4.

3. Modelling Method

In this section, a regression model based on loss analysis is derived by examining the energy losses and simplification of the expressions. Then, the kriging model and the BP neural network are introduced for comparison.

3.1. Loss-Analysis-Based Regression Model. Loss analysis is the theoretical analysis of energy loss on each part of the compressor by investigating the thermodynamic process of the gas inside the compressor, resulting in the approximation of the actual work of the compressor. In the following discussion, one more step, that is, simplification of expressions is taken to address the relationship between pressure ratio, efficiency, and compressor parameters in a certain working status.

3.1.1. Representation of Characteristic Maps. Characteristic curves are drawn to express the interrelation of the main parameters of the compressor. Normally, it comprises two sets of curves, which are pressure ratio and isentropic efficiency versus corrected mass flow rate and corrected rotational speed, deriving from dimensional analysis and similarity principle. The curves contain most of the information we need to predict the performance of a certain centrifugal compressor; therefore, we focus on reconstructing the curves efficiently using a limited number of sample data. The curves of pressure ratio (π_c) and isentropic efficiency (η_{cs}) can be described as follows:

$$
\begin{aligned}
\pi_c &= F_\pi \left(G_c, n_c \right), \\
\eta_{cs} &= F_\eta \left(G_c, n_c \right),
\end{aligned}
\tag{1}
$$

TABLE 1: Required input settings for centrifugal compressor.

| Parameter names	Values	Unit
Design pressure ratio	1.86	
Design mass flow rate	0.5	kg/s
Design rotational speed	10,000	rpm
Inlet stagnation temperature	428	K
Inlet stagnation pressure	600,000	Pa
Inlet angle	0	degree
Radial distribution	Constant angle	
Incidence at shroud	1.5	degree
Relative velocity ratio	0.52	
Hub inlet diameter	30	mm
Stacking	Radial	
Main vanes	9	
Intervanes	9	
Back sweep angle	45	degree
Rake angle	30	degree
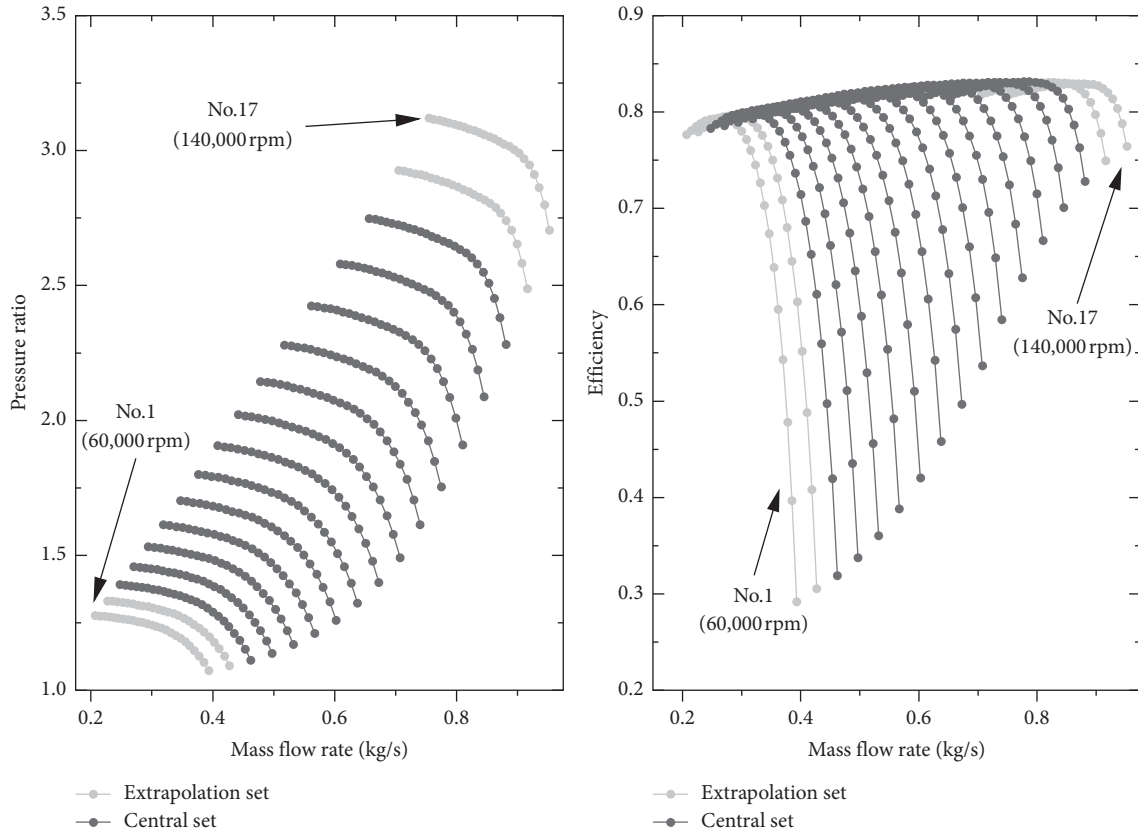
FIGURE 1: Characteristic map of centrifugal compressor.

where G_c (corrected mass flow rate) and n_c (corrected rotational speed) are used to eliminate the influence of inlet status, which are defined as follows [17]:

$$G_c = G \cdot \frac{101.325\,[\text{kPa}]}{p_{in}} \sqrt{\frac{T_{in}}{293\,[\text{K}]}},$$

$$n_c = n \cdot \sqrt{\frac{293\,[\text{K}]}{T_{in}}}, \tag{2}$$

where G denotes mass flow rate, n denotes rotational speed, and p_{in} and T_{in} are the pressure and temperature at the inlet of compressor.

Generally, the characteristic curves are represented into several constant speed lines, which starts from surge boundary and end up with choke limit.

3.1.2. Structures of Loss-Analysis-Based Models. The thermodynamic analysis of centrifugal compressor is performed

as follows, in order to construct the fundamental structure of the regression model.

The isentropic efficiency of compressor η_{cs} can be written as follows:

$$\eta_{cs} = \frac{h_{cs}}{h_c}. \tag{3}$$

The isentropic work, h_{cs}, and the actual compress work, h_c, can be written as follows [18]:

$$h_{cs} = c_p T_1 \left[\pi_c^{(\gamma-1)/\gamma} - 1 \right], \tag{4}$$

$$h_c = \frac{c_p T_1 \left(\pi_c^{(\gamma-1)/\gamma} - 1 \right)}{\eta_{cs}}. \tag{5}$$

From the aspect of energy loss, the isentropic work can be expressed as the difference value of impeller work Δh_{imp} and internal loss Δh_{int}, while the actual compress work can be expressed as the sum of impeller work Δh_{imp} and parasitic loss Δh_{par} [19]:

$$h_{cs} = \Delta h_{\text{imp}} - \Delta h_{\text{int}}, \tag{6}$$

$$h_c = \Delta h_{\text{imp}} + \Delta h_{\text{par}}. \tag{7}$$

Instead of directly using pressure ratio and efficiency as targets of regression, we can transform these variables into more suitable forms by substituting equations (6) and (7) into (4) and (5), respectively. The transformed variables π^* and η^* are defined as follows:

$$\pi^* = \pi_c^{(\gamma-1)/\gamma} = 1 + \frac{\Delta h_{\text{imp}} - \Delta h_{\text{int}}}{c_p T_1}, \tag{8}$$

$$\eta^* = \frac{\left(\pi_c^{(\gamma-1)/\gamma} - 1 \right)}{\eta_{cs}} = \frac{\left(\Delta h_{\text{imp}} + \Delta h_{\text{par}} \right)}{c_p T_1}. \tag{9}$$

When implementing regression, pressure ratio π_c and isentropic efficiency η_{cs} first transfer into π^* and η^* and then transfer back into π_c and η_{cs} after calculation.

3.1.3. Analysis of Loss Models.
In order to obtain the detailed expression of equations (8) and (9), the impeller work, internal loss, and parasitic loss are discussed below.

(1) Impeller Work. Assume that the compressor contains a radially bladed impeller and gas approaching the impeller from axial direction; the specific enthalpy can be written as follows:

$$\Delta h_{\text{imp}} = \sigma u_2^2. \tag{10}$$

For radial blades, the slip factor σ can be approximately calculated by Stanitz's correlation [19]:

$$\sigma = 1 - \frac{0.63\pi}{Z}. \tag{11}$$

Therefore, the slip factor is mainly determined by physical structure of the compressor. From the discussion above, it can be inferred that, the specific enthalpy of the impeller is proportional to the square of rotational speed.

(2) Internal Loss and Parasitic Loss. The loss in centrifugal compressor contains internal losses and parasitic losses. The disk friction loss, recirculation loss, and leakage loss together constitute the parasitic losses [20], as listed in the first three rows in Table 2. The internal loss consist of incidence loss, blade loading loss, skin friction loss, clearance loss, mixing loss, vaneless diffuser (VLD) loss, and vaned diffuser (VD) loss [21], which are listed in the other rows in Table 2.

3.1.4. Simplification of Loss-Analysis-Based Models.
To obtain the final structure of the models, we have to replace Δh_{imp}, Δh_{imp}, and Δh_{par} by the impeller work, internal loss, and parasitic loss models. However, directly substituting the comprehensive models into equations (8) and (9) will lead to a complicated expression with too many parameters and intermediate variables, which is hard to perform regression. Hence, it is necessary to simplify the model.

(1) Simplification of the Isentropic Efficiency Model. One way to implement the simplification is to separate the parameters and variables that form the expression of loss models. The parameters are dominated by the physical structure of compressors, which means they will not change once the compressor is set. Therefore, parameters are considered as constants during calculation. On the other hand, the variables are the status of gas in different stages of the compressor, including velocity, temperature, and pressure. By transforming some of these variables into forms regarding corrected flow rate G_c and corrected rotational speed n_c, the appropriate expression should be acquired for regression.

One of the most essential terms is the tangential tip velocity u_2, which can be calculated as follows [30]:

$$u_2 = \pi D_2 n, \tag{12}$$

where D_2 is the diameter at the outlet of the impeller. That means the tip velocity is proportional to the rotational speed since π and D_2 are constants.

Therefore, the expression of impeller work can be written as follows:

$$\Delta h_{\text{imp}} = \sigma u_2^2 = \sigma \left(\pi D_2 \right)^2 \cdot n^2. \tag{13}$$

By substituting u_2 and G into n_c and G_c, the parasitic losses change into

$$\Delta h_{\text{par}} = \Delta h_{\text{df}} + \Delta h_{\text{rc}} + \Delta h_{\text{lk}}$$
$$= f'_{\text{df}} \cdot \left(\frac{n_c^3}{G_c} \right) + f'_{\text{rc}} \cdot n_c^2 + f'_{\text{lk}} \cdot \Delta h_{\text{imp}}. \tag{14}$$

The coefficient f_x, in equations above, is hard to determine since it contains complex calculation of parameters and variables of the compressor, as shown in Table 2. However, they vary in relatively small ranges, therefore can be linearized into $f_x = f_1 + f_2 G_c + f_3 n_c$. Accordingly, the loss-analysis-based regression model of isentropic efficiency, equation (9) can be replaced as follows:

TABLE 2: A set of loss models of centrifugal compressors.

Loss mechanism	Loss model
Disk friction loss [22]	$\Delta h_{df} = f_{df}(\rho_1 + \rho_2)r_2^2 u_2^3/(8G)$
Recirculation loss [23]	$\Delta h_{rc} = 0.02\sqrt{\tan\alpha_2}D_f^2 u_2^2$
Leakage loss [24]	$\Delta h_{lk} = (G_{cl}/G)\Delta h_{imp}$
Incidence loss [25]	$\Delta h_{inc} = f_{inc}w_{1u}^2/2$
Blade loading loss [26]	$\Delta h_{bl} = 0.05 D_f^2 u_2^2$
Skin friction loss [27]	$\Delta h_{sf} = 5.6 f_{sf} L \overline{w}_2^2/D_{hyd}$
Clearance loss [28]	$\Delta h_{cl} = 2(s/b_2)[(r_h + r_s)/2r_2 - 0.275]u_2^2$
Mixing loss [29]	$\Delta h_{mix} = (1 + \tan^2\alpha_2)^{-1}[(1 - \varepsilon_w - b^*)/(1 - \varepsilon_w)]^2 v_2^2/2$
Vaneless diffuser loss [21]	$\Delta h_{vld} = c_p T_{02}[(p_3/p_{03})^{(\gamma-1)/\gamma} - (p_3/p_{02})^{(\gamma-1)/\gamma}]$
Vaned diffuser loss [21]	$\Delta h_{vd} = c_p T_{03}[(p_4/p_{04})^{(\gamma-1)/\gamma} - (p_4/p_{03})^{(\gamma-1)/\gamma}]$

$$\eta^* = \frac{\left(\Delta h_{imp} + \Delta h_{par}\right)}{c_p T_1}$$

$$= f_{\eta_0} + f_{\eta_1}n_c^2 + f_{\eta_2}n_c^3 + f_{\eta_3}\left(\frac{n_c^3}{G_c}\right) + f_{\eta_4}\left(\frac{n_c^4}{G_c}\right) + f_{\eta_5}n_c^2 G_c, \tag{15}$$

where f_{η_i} is the coefficient.

(2) Simplification of the Pressure Ratio Model. The loss-analysis-based regression model of pressure ratio cannot easily transform using the approach described in the previous section since the components of internal losses are too complicated to reduce to simple forms. Nevertheless, some approximation of these elements can be inferred. Here, some terms derived from corrected rotational speed and corrected flow rate are listed below:

$$G_c, G_c n_c, G_c n_c^2, G_c^2, G_c^2 n_c, G_c^2 n_C^2, n_c, n_c^2, n_c^3, n_c^4, \frac{n_c^3}{G_c}, \frac{n_c^4}{G_c}. \tag{16}$$

Combining these terms will result in a very long expression. Therefore, analysis of variance (ANOVA), a procedure for determining whether variation in the response variable rose within or among different population groups, is performed to seek the most significant terms. The result is shown in Figure 2.

Higher F value indicates higher significance. Consequently, the reduced expression for pressure ratio regression can be derived by eliminating the low-value terms:

$$\pi^* = f_{\pi_0} + f_{\pi_1}G_c n_c + f_{\pi_2}G_c^2 + f_{\pi_3}G_c^2 n_c + f_{\pi_4}n_c^2 + f_{\pi_5}\left(\frac{n_c^3}{G_c}\right), \tag{17}$$

where f_{π_i} is the coefficient.

Equations (15) and (17) are the final forms for regression. They will be applied to the sample data for further validation.

3.2. Kriging Model Based on Second-Order Polynomial. The kriging model estimates the unknown information of a point by weighing a linear combination of neighbour information within a certain range. The essence is to determine the vector of weight coefficient by minimizing the error variance of the estimated value. Therefore, the kriging model

is considered to be a linear unbiased estimation with the smallest variance. The kriging model assumes that the relationship between the system response and the input variable consists of a basic regression model and a nonparametric part, that is,

$$y(x) = F(\beta, x) + Z(x), \tag{18}$$

where $F(\beta, x)$ is the basic regression model:

$$F(\beta, x) = \sum_{j=1}^{k} \beta_j f_j(x). \tag{19}$$

Usually, the order and coefficients of regression model are constant.

$Z(x)$ in equation (18) is a zero-mean stochastic process with the following statistical properties:

$$\text{Cov}\left[Z(x_i), Z(x_j)\right] = \sigma^2 R(x_i, x_j), \tag{20}$$

where σ^2 is the process variance and R is the correlation coefficient matrix.

The elements in the matrix are functions of the distance between the two sample points, which characterize the spatial correlation between the input variables. Commonly used correlation function models include linear models, exponential models, Gaussian models, and spherical models. For modelling the characteristics of the compressor, the spherical model can obtain a smaller error [31], that is,

$$R(\theta, d) = 1 - 1.5 + 0.5\xi_k^3, \quad \xi_k = \min\{1, \theta d\}. \tag{21}$$

By observing the characteristic map of the compressor, we can find that the pressure ratio and efficiency curves are approximately parabolic on every single rotational speed line. Moreover, when the rotational speed is lower, the interval of rotational speed lines is smaller, while the interval of rotational speed lines is larger when the rotational speed is higher. Therefore, a second-order polynomial can be taken as the basic regression model of the kriging model to express the characteristic curves of the compressor better.

3.3. BP Neural Network Model. The neural network model is a computational model that mimics the structure and function of biological nerves and can be used to estimate and approximate functions. The neural network is calculated by a
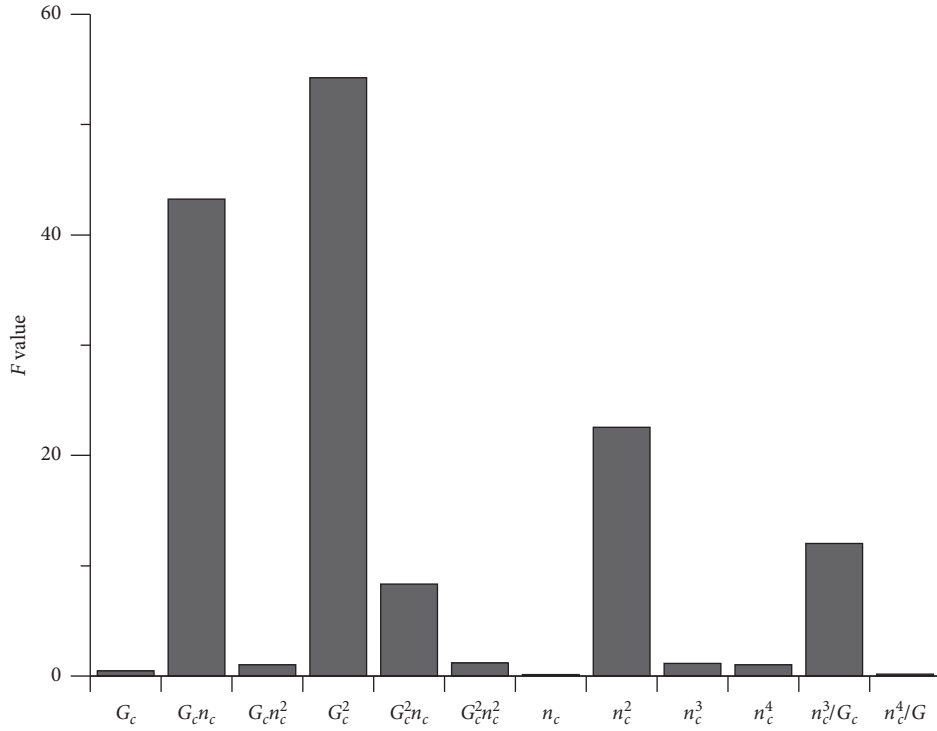
Figure 2: Result of F values.

large number of neuron nodes, usually including the input layer, the hidden layer, and the output layer. In this paper, a three-layer neural network is selected as the compressor performance prediction model, which is an input layer, a hidden layer, and an output layer. The number of hidden layer nodes has an important impact on the prediction results. Very few numbers will result in inaccurate prediction results. Too many samples will require more samples for training and may lead to overfitting. In this paper, a hidden neural network of 10 nodes is constructed to obtain good prediction results with fewer training samples. The activation function of the hidden layer is a sigmoid function, and its form is as follows:

$$S(t) = \frac{1}{1 - e^{-t}}. \tag{22}$$

The structure of the neural network is shown in Figure 3.

In this paper, two neural networks are constructed for the compressor pressure ratio and efficiency, respectively. The neural networks are trained by the backpropagation algorithm, which calculates the gradient of the loss function for all weights in the network and feeds back to the optimization algorithm, in order to update weights to minimize the loss function.

4. Results and Discussion

In this section, some typical schemes of training and test sets are used to test the proposed loss-analysis-based model and compare it with the kriging model and the neural network model. The training and test data are generated using Vista CCD, as described in Section 2. First, the central set is used
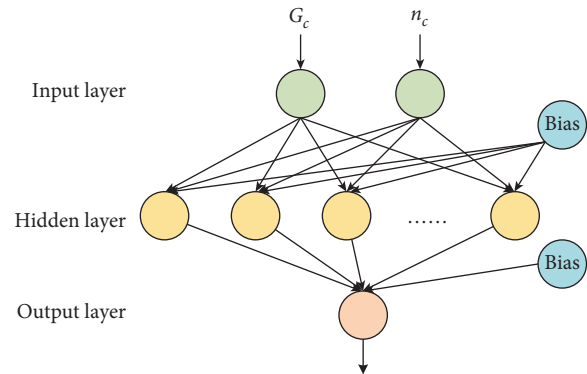

Figure 3: Structure of a neural network.

to examine the loss-analysis-based model. Then, we expand the data set to validate the model and compare with the kriging model and the neural network model.

4.1. Result of Loss-Analysis-Based Model on Central Set. In the following calculation, the division of training set and test set is listed in Table 3. Five speed lines are chosen, that is, No. 3, 6, 9, 12, and 15 (referring to Figure 1), and 10 points on each speed line are evenly selected to form the whole training set with 50 points in total. Then the speed lines between these training lines, that is, No. 4, 5, 7, 8, 10, 11, 13, and 14 are chosen to form the test set. Differently, all sample points on each test lines are selected.

For each training result, the data is applied to validate by root mean square error (RMSE) and coefficient of determination (R^2). The RMSE is the standard deviation of the

TABLE 3: Scheme of training set and test set on central set.

No.	rpm	Training set	Test set
3	70,000	•	
4	75,000		•
5	80,000		•
6	85,000	•	
7	90,000		•
8	95,000		•
9	100,000	•	
10	105,000		•
11	110,000		•
12	115,000	•	
13	120,000		•
14	125,000		•
15	130,000	•	
No. of points on each line		10	25
Total no. of points		50	200

residuals, and it can measure the average distance between the regression result and the original data. The R^2 is the proportion of explained variation to total variation, and it denotes how close the data are to the fitted regression line. These two statistical measures can be defined as follows:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \widehat{y}_i)^2},$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \widehat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \overline{y}_i)^2}. \tag{23}$$

The results are plotted in Figure 4. For lower rotational speed lines, the estimating results are almost coincided with the simulation data points, meaning that the model can provide very high accuracy predictions in relatively low rotational speed. For higher rotational speed lines, the estimating results slightly deviate from the origin points, but still remain very close estimations.

The results of RMSE and R^2 are listed in Table 4. The low value of RMSE indicates that the average difference of the regression result and the original data are very small, and the R^2 implies that the regression model is well fitted. The prediction on pressure ratio is slightly better than the prediction on efficiency. Overall, the results are good enough for performance prediction.

In summary, both efficiency and pressure ratio predictions are close to the data generated by the simulation program, and the trends of the characteristic maps are well preserved. Therefore, the loss-analysis-based model can be used to provide performance prediction for engineering applications.

4.2. Comparison of Models

4.2.1. Generating Training Set and Test Set.
In order to examine the influence of the number of data points and the number of rotational speed lines on the prediction results, the data set is divided into several groups and tested separately, instead of using all sample points in every test.

From the characteristic map in Figure 1, we can find that, in the same rotational speed line, that is, when the rotational speed is constant, the pressure ratio and efficiency change with the flow rate is approximated as a smooth second-order curve. Also, each rotational speed line starts from a minimum flow rate point (surge point) and end up with a maximum flow rate point (choke point). Therefore, at least three points need to be selected on each rotational speed line. If more than three sample data point should be chosen from a rotational speed line, it can be determined using uniform distribution between the maximum flow rate and the minimum flow rate.

Similarly, we can see from Figure 1 that the rotational speed lines are denser in lower rotational speed, and sparser in higher rotational speed; therefore, the pressure ratio and the efficiency can also be seen in a second-order relationship with the rotational speed, approximately. Hence, at least three rotational speed lines are required for the sample data, and it is preferable to select the minimum rotational speed, the maximum rotational speed, and the designed rotational speed. When more speed lines should be added, they can be selected evenly between the minimum and maximum rotational speed lines.

However, if three rotational speed lines are chosen and three or four points on each line are selected, the total number of sample points is too low to perform an effective regression. Hence, we set the minimum number of total sample points to 15, in order to avoid this situation.

The detailed division criteria are shown in Table 5.

In Table 5, T3 to T7 represent the training data set, and the digit after "T" denotes the number of rotational speed lines in it. The specific rotational speed lines selected are shown in the table above.

IP indicates interpolation test set, which contains even number of rotational speed lines from 4th to 14th, while EP indicates extrapolation test set, which contains the 1st, 2nd, 16th, and 17th rotational speed lines.

In addition, for every training set, a certain number of points are taken evenly from every selected rotational speed line to construct a complete training set so that the total number of sample points can roughly cover the interval from 15 to 70, which is also listed in the table above. Therefore, we can study the influence of the number of rotational speed line and the total number of sample points on the training effect.

Before each training session, 5% Gaussian noise is added into the current training set. To compare among the three models, RMSE is used as the index. The training process for every set is repeated 50 times to obtain the average result for the models.

4.2.2. Validation on Interpolation.
In order to illustrate the validity of the loss-analysis-based model, the training set is used to build the models and the results on the test set are calculated and compared with the results of the kriging model and the neural network model.

In this section, T3 to T7 are used to train the loss-analysis-based model, the kriging model, and the neural
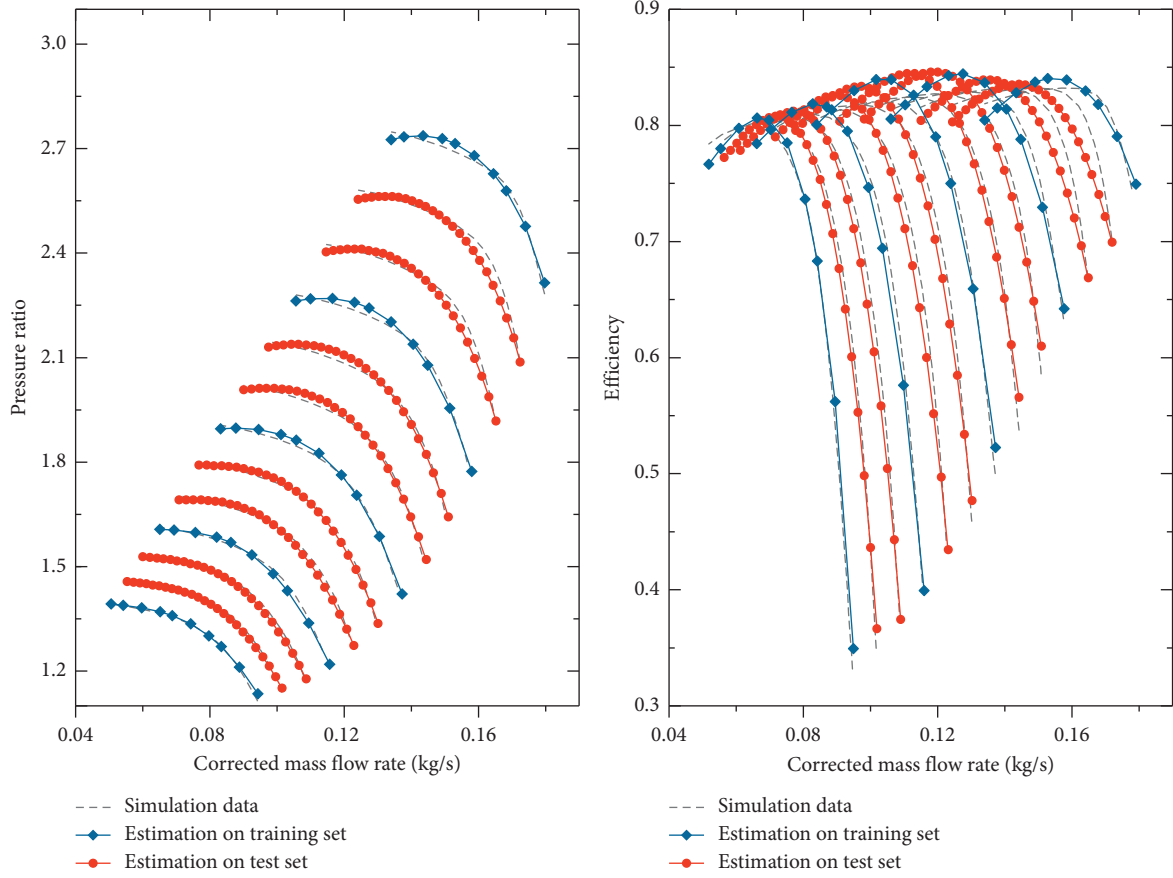
Figure 4: Results of loss-analysis-based model on central set.

Table 4: RMSE and R^2 results on central set.

Parameter	Training set	Test set
Pressure ratio		
RMSE	0.0208	0.0228
R^2	0.9982	0.9966
Efficiency		
RMSE	0.0153	0.0157
R^2	0.9806	0.9751

network, and the interpolation test results of the pressure ratio and efficiency characteristics of the centrifugal compressor are obtained using interpolation test set IP. The images are shown in Figures 5 and 6.

It can be seen that the loss-analysis-based model (a) and the kriging model (b) are very stable and the error is small. When predicting the pressure ratio of compressors, the results of the two models are close, and the loss-analysis-based model produces better results in predicting efficiency of compressors.

For the loss-analysis-based model, the increase in the number of rotational speed lines does not help to improve the calculation results, while the increase in the total number of sample points slightly reduces the prediction error when predicting the efficiency of the compressor. It is not necessary to use more than 30 points since the RMSE would not decrease significantly.

For the kriging model, the result of predicting pressure ratio is stable. For predicting the efficiency, increasing the number of speed lines and the total number of sample points can slightly reduce the error, but it is worse than the result of the loss-analysis-based model.

For the neural network model (c), in the case of a small number of speed lines, the model has an overfitting phenomenon, as shown by the lines T3 and T4. When the number of speed lines and the total number of sample points are larger, the prediction error is significantly reduced. It shows that the neural network model has better result when the sample points are denser and the number of speed lines is larger, and the prediction error is lower than that of the loss-analysis-based model and the kriging model, as shown in line T7. Therefore, the neural network model is only suitable when the training data set is large enough.

4.2.3. Validation on Extrapolation. Similarly, T3 to T7 are used to train the loss-analysis-based model, the kriging model, and the neural network model respectively, and the calculation results are examined by the extrapolation test set EP, as shown in Figures 7 and 8.

In general, the loss-analysis-based model performs the best in the extrapolation prediction, and the error is relatively small. It has advantages compared with the kriging model and the neural network model. This is because the

TABLE 5: Scheme of training set and test set on all data.

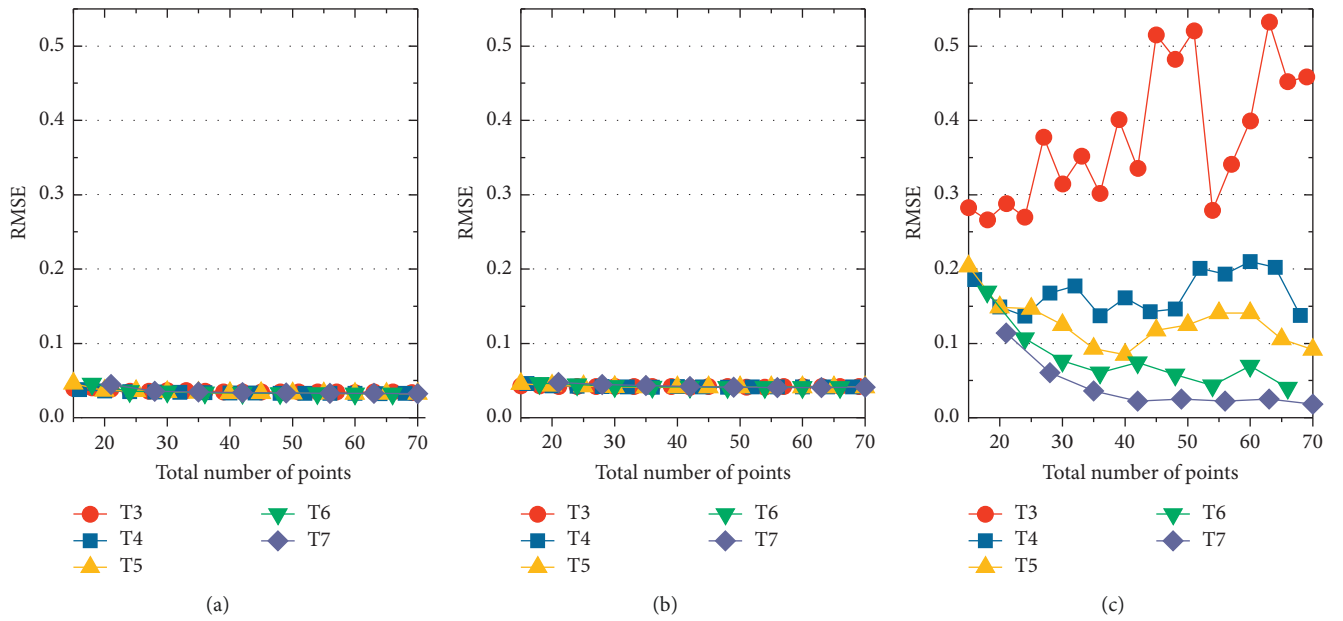| No.	rpm	T3	T4	T5	T6	T7	IP	EP
1	60,000							•
2	65,000							•
3	70,000	•	•	•	•	•		
4	75,000						•	
5	80,000			•	•	•		
6	85,000						•	
7	90,000		•		•	•		
8	95,000						•	
9	100,000	•		•		•		
10	105,000						•	
11	110,000		•		•	•		
12	115,000						•	
13	120,000			•	•	•		
14	125,000						•	
15	130,000	•	•	•	•	•		
16	135,000							•
17	140,000							•
No. of points on each line		5~23	4~17	3~14	3~11	3~10	25	25
Total no. of points		15~69	16~68	15~70	18~66	21~70	150	100

FIGURE 5: Results of predicting pressure ratio on the interpolation test set. (a) The loss-analysis-based regression model; (b) the kriging model; (c) the BP neural network model.

model is based on the loss analysis with simplification, which includes physical meaning, and can reflect the general trend of compressor characteristic. Therefore, it can rely on fewer data to reconstruct the performance curves in extrapolation calculation. When the total number of sample points or the speed lines increases, the extrapolation error decreases slightly. It means that to predict the performance outside the training data boundary, it still needs more data to generate the trend of characteristic maps.

The kriging model fluctuates when predicting efficiency, while the error is smaller in the pressure ratio prediction.

The prediction of pressure ratio is close to that of the loss-analysis-based model.

For the neural network model, when the number of speed lines is large, the error decreases with the increase of the number of sample points, but the predicting effect is still not as good as the loss-analysis-based model or the kriging model. This result may be explained by the fact that the neural network model contains little physical information, thus it cannot provide a proper prediction for the points outside the training data boundary. Generally, the neural network model is not suitable for extrapolation prediction.
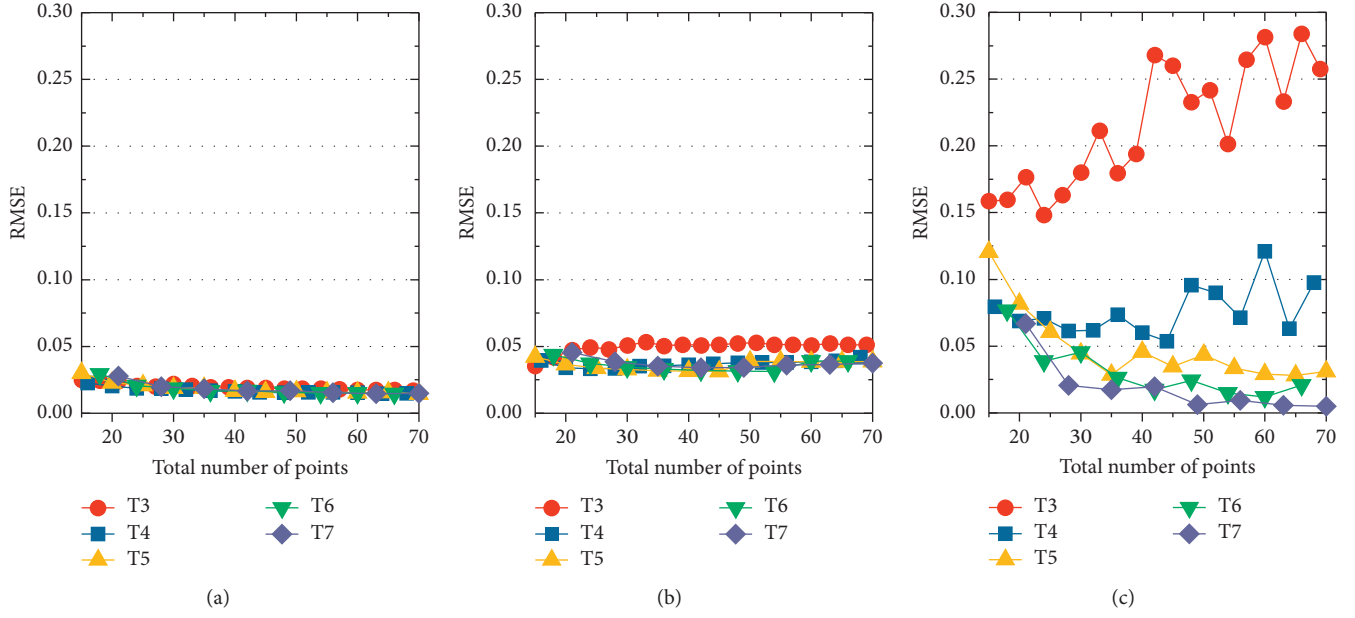
FIGURE 6: Results of predicting efficiency on the interpolation test set. (a) The loss-analysis-based regression model; (b) The kriging model; (c) The BP neural network model.
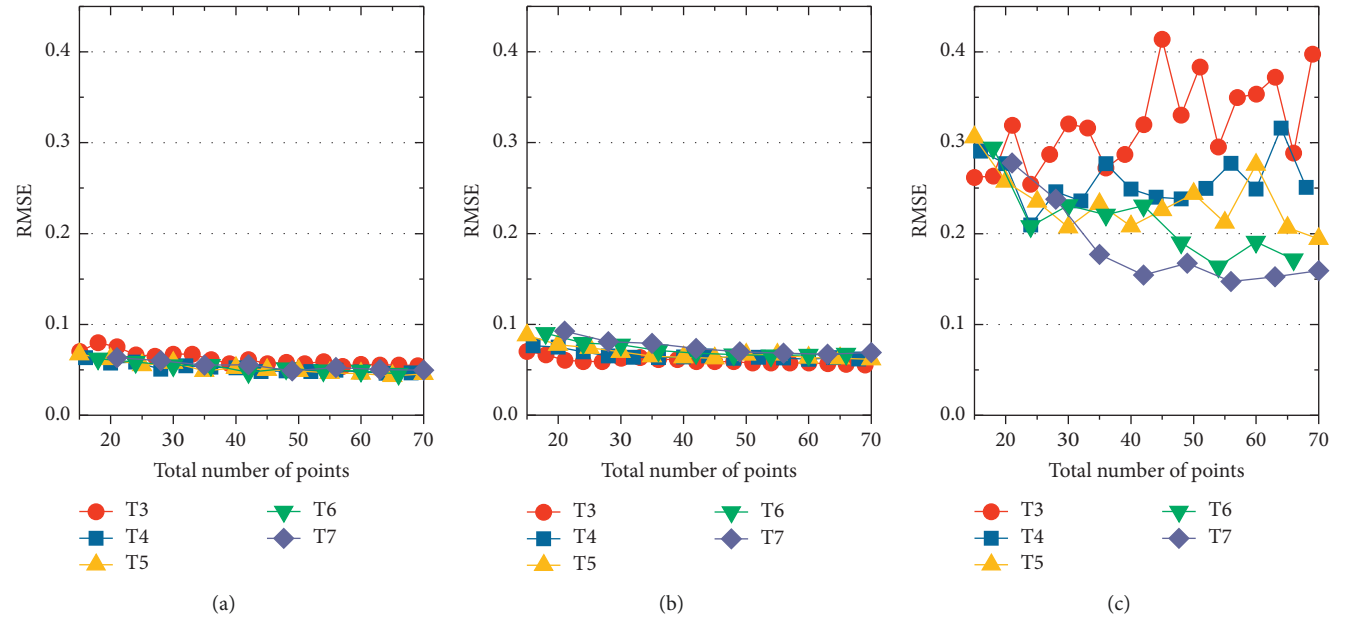

FIGURE 7: Results of predicting pressure ratio on the extrapolation test set. (a) The loss-analysis-based regression model; (b) the kriging model; (c) the BP neural network model.

4.3. Running Time of Models.

In the meantime, the whole training and test process is also clocked, and the total time is listed in Table 6. Each of the models is coded in MATLAB and run on Windows 7 operation system and on a workstation with Intel Xeon CPU E3-1230v5 3.4 GHz and 8 GB RAM.

It can be seen from the table that the neural network takes the longest time, whether it is model training or model prediction; the time of the loss-analysis-based model and the kriging model is similar, and the time required to train the loss-analysis-based model is shorter.

In summary, the loss-analysis-based model consumes far less computational resources than other models, and it is the appropriate model to provide performance predicting for both interpolation and extrapolation cases.
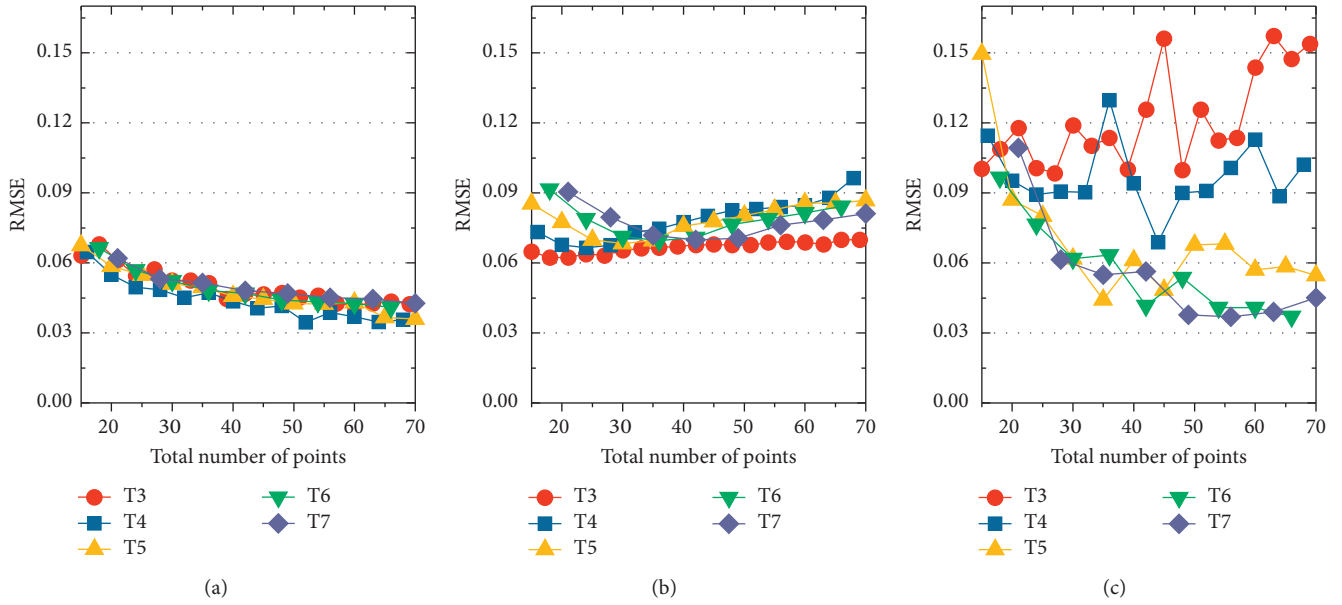
FIGURE 8: Results of predicting efficiency on the extrapolation test set. (a) The loss-analysis-based regression model; (b) the kriging model; (c) the BP neural network model.

TABLE 6: Running time of the models.

	Loss-analysis-based model (s)	Kriging model (s)	Neural network model (s)
Training time	50.48	122.53	711.78
Predicting time	14.73	14.11	58.58
Total	65.21	136.64	770.36

5. Conclusions

In this paper, analysis of energy loss of the centrifugal compressor was performed. Then, simplification and sensitivity analysis were carried out, and a loss-analysis-based regression model was obtained combining with the loss model of the compressor. The model contains certain physical meaning and can be easily applied to the fitting calculation. This model was fed by sample data generated by Vista CCD for testing. Results show that the loss-analysis-based model preserves the trends of both efficiency and pressure ratio, and the predicting result is very close to the original data.

On the other hand, the kriging model was constructed based on the second-order polynomial and the spherical correlation model, and the BP neural network model was also built for comparison. Then, the models were validated using simulation data with added noise. The calculation results show that the predicting results of the loss-analysis-based model and the kriging model are relatively stable, and the error is very small when the interpolation prediction is performed within a given data range. The predicting result of the neural network model improves as the number of points increases.

In addition, the loss-analysis-based model maintains a strong extrapolation predicting ability, which can be attributed to the inherent physical meaning of the model. The basic part of the kriging model uses a second-order

polynomial, which is close to the trend of the compressor characteristic curve; therefore, it also has extrapolation ability. The extrapolation ability of the neural network model is poor compared to these two models.

In terms of running time, since the loss-analysis-based model structure is not complicated and the calculation process is simple, it consumes minimal time. The required time for prediction using the kriging model is similar, but the training time is longer. The neural network model consumes significantly more time than other models, and there is no advantage under this condition. In general, the loss-analysis-based model is suitable for engineering applications.

Although this paper mainly discussed the model for centrifugal compressor, the method of constructing loss-analysis-based regression model, containing loss analysis, simplification, and sensitivity analysis, can be applied to similar mechanical components, such as turbine, fan, and pump. The models can also be included in a system-level simulation and optimization, providing accurate result while saving computational resource and time.

Nomenclature

Variable

b: Impeller width
b^*: Ratio of vaneless diffuser inlet width to impeller exit width
c: Absolute velocity
c_p: Specific heat at constant pressure
D: Diameter
D_f: Diffusion factor
f: Coefficient in correlations
G: Mass flow rate
G_c: Corrected mass flow rate
h: Specific enthalpy

L: Length of impeller flow
n: Rotational speed
n_c: Corrected rotational speed
p: Pressure
r: Radius
s: Clearance width
T: Temperature
u: Tangential impeller speed
w: Relative velocity
Z: Number of blades
α: Absolute flow angle
ε_w: Wake fraction of blade-to-blade space
η_c: Efficiency
γ: Specific heat ratio
π_c: Pressure ratio
ρ: Density
σ: Slip ratio.

Subscript
0: Total parameter
1: Inlet of impeller
2: Outlet of impeller/inlet of vaneless diffuser
3: Outlet of vaneless diffuser/inlet of vaned diffuser
4: Outlet of compressor
bl: Blade load
c: Compressor
cl: Clearance
cs: Isentropic process of compressor
df: Disk friction
h: Hub
hyd: Hydraulic diameter
in: Inlet of pressure
inc: Incidence
imp: Impeller
int: Internal
lk: Leakage
mix: Mixing
par: Parasitic
rc: Recirculation
s: Shroud
sf: Surface friction
u: Tangential component
vd: Vaned diffuser
vld: Vaneless diffuser.

Data Availability

The data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.
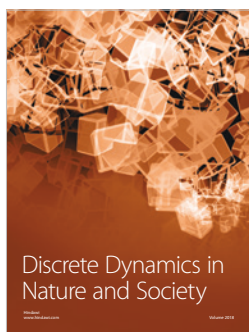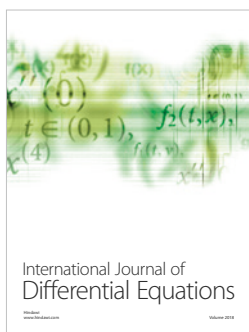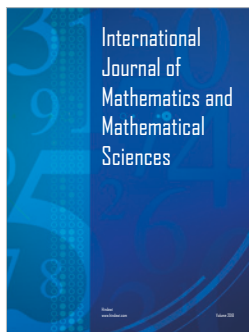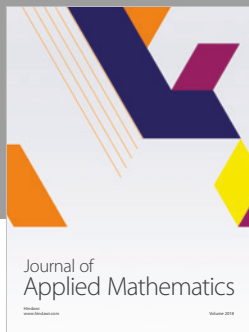
Acknowledgments

References

[1] S. A. Korpela, *Principles of Turbomachinery*, John Wiley & Sons, Hoboken, NJ, USA, 2011.

[2] C. Ji, C. Li, J. Fang, and Q. Sun, "Loss mechanism of static interstage components of multistage centrifugal compressors for integrated blade design," *Mathematical Problems in Engineering*, vol. 2018, Article ID 9025650, 16 pages, 2018.

[3] J. Kurzke and C. Riegler, "A new compressor map scaling procedure for preliminary conceptional design of gas turbines," in *Proceedings of the Volume 1: Aircraft Engine; Marine; Turbomachinery; Microturbines and Small Turbomachinery*, May 2000.

[4] A. M. El-Gammal, "An algorithm and criteria for compressor characteristics real time modeling and approximation," in *Proceedings of the ASME 1990 International Gas Turbine and Aeroengine Congress and Exposition*, American Society of Mechanical Engineers (ASME), Brussels, Belgium, June 1990.

[5] J. T. Gravdahl, A. Egeland, and S. O. Vatland, "Active surge control of centrifugal compressors using drive torque," in *Proceedings of the 40th IEEE Conference on Decision and Control (Cat. No.01CH37228)*, pp. 1286–1291, IEEE, Orlando, FL, USA, December 2001.

[6] I. Yazar, H. S. Yavuz, and A. A. Yavuz, "Comparison of various regression models for predicting compressor and turbine performance parameters," *Energy*, vol. 140, pp. 1398–1406, 2017.

[7] E. Tsoutsanis, Y. G. Li, P. Pilidis, and M. Newby, "Part-Load performance of gas turbines: Part II—multi-point Adaptation with compressor map generation and GA optimization," in *Proceedings of the ASME 2012 Gas Turbine India Conference*, pp. 743–751, American Society of Mechanical Engineers (ASME), Mumbai, Maharashtra, India, December 2012.

[8] E. Tsoutsanis, N. Meskin, M. Benammar, and K. Khorasani, "A component map tuning method for performance prediction and diagnostics of gas turbine compressors," *Applied Energy*, vol. 135, pp. 572–585, 2014.

[9] M. Zagorowska and N. Thornhill, "Compressor map approximation using Chebyshev polynomials," in *Proceedings of the 2017 25th Mediterranean Conference on Control and Automation (MED)*, pp. 864–869, IEEE, Valletta, Malta, July 2017.

[10] A. Lazzaretto and A. Toffolo, "Analytical and neural network models for gas turbine design and off-design simulation," *International Journal of Thermodynamics*, vol. 4, no. 4, pp. 173–182, 2001.

[11] Y. Yu, L. Chen, F. Sun, and C. Wu, "Neural-network based analysis and prediction of a compressor's characteristic performance map," *Applied Energy*, vol. 84, no. 1, pp. 48–55, 2007.

[12] K. Ghorbanian and M. Gholamrezaei, "An artificial neural network approach to compressor performance prediction," *Applied Energy*, vol. 86, no. 7-8, pp. 1210–1221, 2009.

[13] Z. Tian, B. Gu, L. Yang, and Y. Lu, "Hybrid ANN-PLS approach to scroll compressor thermodynamic performance prediction," *Applied Thermal Engineering*, vol. 77, pp. 113–120, 2015.

[14] K. Ghorbanian and M. Gholamrezaei, "Axial compressor performance map prediction using artificial neural network," in *Proceedings of the ASME Turbo Expo 2007: Power for Land, Sea, and Air*, pp. 1199–1208, American Society of Mechanical Engineers (ASME), Montreal, Canada, May 2007.

[15] M. Olivero, D. Pasquale, A. Ghidoni, and S. Rebay, "Three-dimensional turbulent optimization of vaned diffusers for

centrifugal compressors based on metamodel-assisted genetic algorithms," *Optimization and Engineering*, vol. 15, no. 4, pp. 973–992, 2014.

[16] J. El Hadef, G. Colin, Y. Chamaillard, and V. Talon, "Physical-based algorithms for interpolation and extrapolation of turbocharger data maps," *SAE International Journal of Engines*, vol. 5, no. 2, pp. 363–378, 2012.

[17] E. Logan Jr. and R. Roy, *Handbook of Turbomachinery*, CRC Press, Boca Raton, FL, USA, 2003.

[18] M. R. Galvas, *Analytical Correlation of Centrifugal Compressor Design Geometry for Maximum Efficiency with Specific Speed*, National Aeronautics and Space Administration, Washington, DC, USA, 1972.

[19] S. Lieblein, *Analysis of Experimental Low-Speed Loss and Stall Characteristics of Two-Dimensional Compressor Blade Cascades*, National Advisory Committee for Aeronautics, Washington, DC, USA, 1957.

[20] E. I. Gutiérrez Velásquez, "Determination of a suitable set of loss models for centrifugal compressor performance prediction," *Chinese Journal of Aeronautics*, vol. 30, no. 5, pp. 1644–1650, 2017.

[21] H. W. Oh, E. S. Yoon, and M. K. Chung, "An optimum set of loss models for performance prediction of centrifugal compressors," *Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy*, vol. 211, no. 4, pp. 331–338, 1997.

[22] J. W. Daily and R. E. Nece, "Chamber dimension effects on induced flow and frictional resistance of enclosed rotating disks," *Journal of Basic Engineering*, vol. 82, no. 1, p. 217, 1960.

[23] J. E. Coppage and F. Dallenbach, *Study of Supersonic Radial Compressors for Refrigeration and Pressurization Systems*, Garrett Corporation, AiResearch Manufacturing Division, Los Angeles, CA, USA, 1956.

[24] R. H. Aungier, "Mean streamline aerodynamic performance analysis of centrifugal compressors," *Journal of Turbomachinery*, vol. 117, no. 3, pp. 360–366, 1995.

[25] O. Conrad, K. Raif, and M. Wessels, "The calculation of performance maps for centrifugal compressors with vane-island diffusers," in *Proceedings of the Twenty-Fifth Annual International Gas Turbine Conference and Exhibit and Twenty-Second Annual Fluids Engineering Conference*, pp. 135–147, American Society of Mechanical Engineers (ASME), New Orleans, LA, USA, 1979.

[26] R. Galvas, *Fortran Program for Predicting Off-Design Performance of Centrifugal Compressors*, National Aeronautics and Space Administration, Washington, DC, USA, 1973.

[27] W. Jansen, "A method for calculating the flow in a centrifugal impeller when entropy gradient are present," in *Proceedings of the Royal Society Conference on Internal Aerodynamics (Turbomachinery)*, pp. 133–146, Institution of Mechanical Engineers, London, UK, 1970.

[28] E. P. Krylov and Y. A. Spunde, *About the Influence of the Clearance between the Working Blades and Housing of a Radial Turbine on its Exponent*, Foreign Technology Division, Wright-Patterson AFB, Dayton, OH, USA, 1967.

[29] J. P. Johnston and R. C. Dean, "Losses in vaneless diffusers of centrifugal compressors and pumps: analysis, experiment, and design," *Journal of Engineering for Power*, vol. 88, no. 1, pp. 49–60, 1966.

[30] S. L. Dixon and C. A. Hall, *Fluid Mechanics and Thermodynamics of Turbomachinery*, Butterworth-Heinemann, Oxford, UK, 2014.

[31] H. Tu and H. Chen, "Modelling of compressor characteristics using Kriging method," *Transactions of CSICE*, vol. 32, no. 4, pp. 377–383, 2014.

Advances in
Operations Research

Advances in
Decision Sciences

Journal of
Applied Mathematics

The Scientific
World Journal

Journal of
Probability and Statistics

International
Journal of
Mathematics and
Mathematical
Sciences

Journal of
Optimization

International Journal of
Engineering
Mathematics

International Journal of
Analysis

Hindawi

Submit your manuscripts at
www.hindawi.com

Journal of
Complex Analysis

Advances in
Numerical Analysis

Mathematical Problems
in Engineering

International Journal of
Differential Equations

Discrete Dynamics in
Nature and Society

International Journal of
Stochastic Analysis

Journal of
Mathematics

Journal of
Function Spaces

Abstract and
Applied Analysis

Advances in
Mathematical Physics