

Research Article

Method for Predicting Failure Rate of Airborne Equipment Based on Optimal Combination Model

WenQiang Li , Ning Hou, and XiangKun Sun

School of Mechatronics Engineering, Shenyang Aerospace University, Shenyang 110136, China

Correspondence should be addressed to WenQiang Li; lwqsss@163.com

Received 6 September 2021; Revised 8 November 2021; Accepted 13 November 2021; Published 20 December 2021

Academic Editor: Samuel N. Jator

Copyright © 2021 WenQiang Li 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.

Accurate prediction of airborne equipment failure rate can provide correct repair and maintenance decisions and effectively establish a health management mechanism. This plays an important role in ensuring the safe use of the aircraft and flight safety. This paper proposes an optimal combination forecasting model, which mixes five single models (Multiple Linear Regression model (MLR), Gray model GM (1, N), Partial Least Squares model (PLS), Artificial Neural Network model (BP), and Support Vector Machine model (SVM)). The combined model and its single model are compared with the other three algorithms. Seven classic comparison functions are used for predictive performance evaluation indicators. The research results show that the combined model is superior to other models in terms of prediction accuracy. This paper provides a practical and effective method for predicting the airborne equipment failure rate.

1. Introduction

Complex equipment plays an important role in military industry and civilian production. Once equipment fails or its performance deteriorates, it will not only seriously affect task execution and production efficiency but also lead to vicious incidents and cause immeasurable losses. Therefore, abnormal detection and fault diagnosis technology of this kind of equipment has been paid much attention by scholars [1]. As a typical complex equipment, airplanes play an important role in the development of military and civil fields, and their airborne equipment plays a key role as an important part of the aircraft system. Due to the influence of various factors such as external factors and the equipment itself, it will cause the failure of airborne equipment. The failure of airborne equipment directly affects the progress of the aircraft's development, testing, and delivery, and affects the normal use of the aircraft. Failure prediction is the pre-event of failure. Its essence is to use the acquired monitoring information to predict the degradation trend and predict the future failures under the premise that the system or components can still work normally [2]. It can be engaged in the transformation from post-maintenance to condition-based

maintenance and for health management. Fault prediction has more research value than post-fault diagnosis. It not only provides good repair and maintenance decisions but also improves the health management level of airborne equipment. The failure prediction of airborne equipment is generally divided into three aspects: failure time prediction, remaining life prediction, and failure rate prediction. The failure rate is the main basis for the allocation of failure samples in airborne equipment maintenance and testing. The prediction results directly affect maintenance and the credibility of the test results [3]. The failure rate prediction of airborne equipment has become the prerequisite and foundation for the integrity of the aircraft. The research and application of the failure rate prediction of airborne equipment is of great significance for ensuring the stability, reliability, and safety of aircraft systems. The main influencing factors of the failure rate of airborne equipment include its own factors, working hours, ambient temperature, ambient humidity, and maintenance quality.

In the past few decades, many researchers have used various models to predict equipment failure rates. The current failure rate prediction methods include traditional reliability predictions, statistical predictions, and data-

driven predictions. Traditional reliability prediction mainly includes the prediction method based on fault tree analysis [4] and the prediction method based on life distribution [5]. Al Badawi et al. designed a fault tree model to identify the failure mechanism in the EMD system. They collect and investigate the fault data of more than 100 electromagnetic drive system faults from the field, establish mathematical models and risk failure models, and the predicted results are accurate. Verma et al. established a prediction model for the failure rate of blood analyzer medical equipment based on Weibull life distribution. They collect failure data of five automatic blood analyzers and draw a Weibull probability map. The results show that the Weibull distribution can be well applied to the modeling of the blood analyzer failure data, and the obtained information is helpful for maintenance cost analysis and facilitates the decision-making for the maintenance mode of the analyzer. However, the characteristics of traditional reliability prediction methods are simple to predict and easy to operate. It is widely used, but the prediction accuracy is relatively low.

In addition to traditional reliability predictions, statistical predictions are also widely used in equipment failure rate prediction. Commonly used statistical prediction methods include Bayesian models and regression analysis models (mainly univariate linear regression analysis prediction method, multiple regression analysis prediction method, and nonlinear regression analysis prediction method), HMM, hidden semi-Markov model, etc. Bayesian-based failure rate prediction method (1996: Louis Hart et al.) [6] carried out related research; based on regression analysis model method, (Ma et al.) [7], first carried out regression analysis based on failure rate. In the research, a spare part failure rate prediction model based on regression analysis was established based on the failure rate data of a certain spare part, and the failure rate of a certain spare part was predicted using this model, and the failure rate was modeled and predicted by the unary linear regression method. The results show that based on the regression method, the correlation coefficient r is 0.9409, and the value is close to 1, and the number of use years has a very close linear relationship with the failure rate of spare parts. The forecast result is basically correct. However, the influence of environment, climate, storage method, and other conditions on the failure rate is ignored. Many scholars subsequently improved the method and gradually improved the accuracy of prediction. E.g. (Motiee et al. [8], Altinisik et al. [9]) have carried out regression analysis model research. At the same time, some scholars also conducted other statistical methods (including HMM and Hidden Semi-Markov Model [10]), and their effects are significant. The characteristic of the statistical-based fault prediction method is that it does not need to establish a physical model of the equipment. As long as the mapping relationship between the current state and the future fault is found, the prediction can be made, and the operability and achievability are strong. However, when the length of the prediction is greater than the length of the original data, the accuracy of the prediction result cannot be guaranteed when this method is used for prediction.

With the development of sensor detection and data processing technology, the operation status and faults of equipment can be effectively monitored by collecting and analyzing data. The data-driven prediction method is also widely used in the modeling and prediction of equipment failure rate. The use of this method improves the accuracy of predicting equipment failure rate. The data-driven fault prediction method is based on a large amount of monitoring historical data to analyze the mapping relationship between system input, output, and system state to establish a prediction model, which is used for fault prediction. In recent years, it has received extensive attention and research. Commonly used data-driven forecasting methods include: time series-based forecasting methods (including linear stationary models such as AR models, MA models, and ARMA models) and ANN-based time series forecasting (including BP neural networks, RBF neural networks, and wavelet neural networks). Based on the time series forecasting method (Ruiying et al. [11]), the data of the Boeing 757–700 aircraft in the past two years is analyzed and processed and the failure rate is predicted. The error between the predicted value and the true value is within 40%. The more traditional reliability prediction and reliability test evaluation results have been greatly improved, and the predicted results provide references for reliability maintenance. But, the error is larger. Subsequently, other researchers used the ARMA model to predict the equipment failure rate and improved the accuracy of the prediction. In 2015, Shengbiao [12] also carried out related research. At the same time (Yang et al. in 2017) [13] carried out research on the prediction of aircraft fault rate. They proposed a seasonal ARIMA model and introduced the mathematical model and modeling process of the system in detail. The application of the SARIMA model in aircraft failure rate prediction is analyzed through examples. The application examples show that the SARIMA model can make full use of the historical data and accurately predict aircraft failure rates by characterizing periodic fluctuations. Time series-based forecasting methods are convenient and quick to calculate, and can provide the identification of linear and stationary forecasting models. Many parameters need to be determined. However, it is difficult to fully reflect the complex nonlinear relationships in the series data. With the rapid development and application of artificial neural networks (ANN), (Kutyłowska et al.) [14] proposed an artificial neural network based on multi-layer perception to evaluate the failure rate of 143 centrifugal pumps in a refinery. In order to optimize and fully support the effect of the final preventive maintenance plan, scholars have also carried out ANN-based failure rate prediction research, such as in (Al-Garni et al.) [15], (Ruiying et al.) [16], (Al-Garni) [17], (Diryag et al.) [18], and (Chen) [19], etc. Although the ANN method has the characteristics of fast processing multivariate analysis, providing nonlinear prediction, and not requiring prior knowledge, it generally requires a large amount of data, which is not suitable for the case of a small amount of data.

In addition to the above prediction methods, many scholars have carried out research on other methods for failure rate prediction, such as the gray prediction failure

rate method, such as (Xiaodong et al.) [20], (Ming et al.) [21], (YanJun et al.) [22], and some other prediction algorithm research, such as (Wang et al.) [23], (Kutyłowska et al.) [24]. However, the prediction method does not consider the randomness of the system, and the medium and long-term prediction accuracy is poor. At the same time, some scholars use other methods to study the failure rate prediction, such as (Yang et al. [25]) which proposes a cloud-model-based uncertain reasoning method to predict the failure of power transmission and transformation equipment. The results of the case analysis show that when the equipment health index is greater than 60, the prediction result of this method is more in line with the actual equipment situation than the traditional inversion method.

However, for some equipment, due to the complexity of its system composition and failure mechanism, there is a strong coupling between different components, and single-item modeling is used for failure prediction. The existing prediction methods are difficult to achieve the desired prediction effect [26]. In addition, single-item failure prediction methods have their own limitations and applicable conditions. Therefore, if multiple prediction methods can be combined to perform fault prediction together, it can overcome the shortcomings of traditional single-item prediction methods and improve the accuracy of prediction. The combination prediction method is the main development direction of current research. For this reason, some scholars propose combined failure rate prediction methods, such as genetic neural network [27], Hidden Semi-Markov Model (HSMM) [10], life and data-driven models [28], Holt-Winters seasonal model [29], EMD and RVM-GM model [30], EMD-GMDH model [31], Weighted Hierarchical Bayes (WHB) method [32], Weibull distribution and time series [33], BP and dual-parameter Weibull distribution [34], cloud model and closely sub-model [35], radar chart and SVM model [36], and neural network and fuzzy recognition [37]. Researchers have carried out research on the use of combined models to predict equipment failure rates. The review of combined models in various scientific applications shows that these models have very high accurate prediction capabilities and can predict equipment failure rates well. However, these studies still have limitations. First, the number of models used in the mixed model is limited to two or three. Second, the prediction accuracy needs to be further improved. Third, there are certain deficiencies in the analysis and research of multi-factor influence. Compared with the single model, the prediction results of the combined model are improved to different degrees, but its prediction accuracy, reliability, and application still cannot meet the requirements of some predictions. At the same time, due to the complex fault cross-linking of airborne equipment, the high uncertainty of the cause of the fault, and the small number of samples of the fault, the above prediction methods have certain deficiencies.

In view of this, the article will consider the strong coupling of the airborne equipment operating state, the complexity of the failure mechanism, and the diversity of influencing factors; will study the failure rate prediction method of the airborne equipment; and propose a combined

model to improve the performance of the airborne equipment. In the combined model developed in this research, first consider the parameters that affect the failure of the airborne equipment. Since it is impossible to give all the parameters that affect the failure rate of the airborne equipment, the focus is on the flight time, the number of take-offs and landings, the maneuvering proficiency, and the abnormal ambient temperature. Abnormal environmental humidity, maintenance quality, and historical failure times information are effective variables to carry out research. Secondly, in order to overcome the limitations of the models in previous studies, the combined model uses five types of statistical and data-driven models, including multiple linear regression (MLR) model, grayGM (1,N) model, partial least squares (PLS) model, artificial neural network (BP) model, and support vector machine (SVM) model to predict the failure rate of the airborne equipment of a certain UAV flight control system. Finally, without increasing the complexity, based on the optimal weighted combination modeling theory, the minimum error square sum of the combination model is used as the objective function to solve the optimal weighting coefficients, and the optimal weighted geometric average combination prediction model is derived to combine the five models. A single predicted failure rate value of airborne equipment is regarded as an independent input variable, and the predicted failure rate value is used as a dependent output variable of the combined model for research. And, it verifies the validity and applicability of the proposed method in the failure prediction of airborne equipment, and provides an effective basis for its fault diagnosis and health maintenance. The rest of this paper is organized as follows: Section 2 introduces in detail the selection of airborne equipment failure rate samples and detailed analysis of the corresponding influencing factors. Section 3 analyzes the modeling process of the combined model, introduces the individual single models involved in the combined model in detail, proposes the optimal combination model, and analyzes the model evaluation indicators. Section 4 conducts detailed research and discussion on single models and combined models. Section 5 compares the prediction results and accuracy of each model in detail. Section 6 analyzes the applicability system of different models. Finally, Section 7 gives conclusions and corresponding suggestions.

2. Sample Selection and Influencing Factors Analysis of Failure Rate of Airborne Equipment

At present, UAVs have been widely used in the fields of military, police, agriculture, geology, meteorology, and urban management, and they play an important role in all walks of life. As a key airborne equipment, its flight control system is the core system for the drone to complete the entire flight process such as take-off, air flight, mission execution, and return and recovery. Once the airborne equipment fails, it will affect the normal operation of the drone. Operation and use may even endanger the flight safety of drones. The

airborne equipment of the flight control system can be divided into hardware parts and software components. The hardware part mainly includes gyroscopes, accelerometers, air pressure sensors, ultrasonic sensors, optical flow sensors, GPS modules, and related control circuits to complete the normal flight and attitude adjustment of the man-machine. The software part includes database management software, application software, and security software, which are used to complete the control and monitoring of UAVs and other related functions. In this paper, the airborne equipment of the UAV flight control system is taken as the research object. The airborne equipment mentioned below refers to the UAV flight control system.

According to statistical analysis, there are many factors that cause the failure of airborne equipment, and they cover the entire life cycle process of design, manufacturing, use, and maintenance. The failure rate of airborne equipment is based on the characteristics of the use of different airborne equipment, combined with many influencing conditions and factors, to make a reasonable and scientific prediction of the failure rate of airborne equipment. Factors affecting the failure rate of airborne equipment include the following aspects, as shown in Figure 1.

2.1. Flight Mission Status. Flying task includes the length of flight time, the number of take-offs and landings, and the number of missions. If the flight mission is heavy, the number of drones used will be more, which will lead to longer flight time and increase the number of take-offs and landings. The probability of equipment failure will be greater. Since the natural loss of airborne equipment is generally proportional to the flight mission, the length of flight time, and the number of take-offs and landings can be used as the main factors affecting its failure rate.

2.2. External Environmental Conditions. External environmental conditions mainly include factors such as weather conditions, ambient temperature, and humidity. Abnormal weather conditions and temperature and humidity will have a certain impact on the probability of failure of airborne equipment. Severe weather conditions and environments, those that exceed the reasonable temperature and humidity range, will lead to greater probability of failure. If extreme weather occurs, it will damage the onboard equipment. Excessive ambient temperature will cause the system hardware to overheat, which will cause the bonding of electronic parts and the deterioration of circuit stability. Too low ambient temperature will cause material shrinkage and reduced fluidity of the system hardware, causing hard and brittle materials, interlacing parts, and poor performance of electronic components. Humid environment will cause deterioration of the appearance, and physical, chemical, and electrical properties of the hardware of the onboard equipment, leading to failures such as surface condensation and loose material diffusion. Due to abnormal ambient temperature and humidity conditions, both will have a greater impact on the failure rate of airborne equipment.

2.3. Product Quality. Airborne equipment with good product quality, high system integration, and high reliability will reduce the failure rate of airborne equipment to a certain extent. At the same time, products with poor quality and low stability will also lead to increased failures of airborne equipment. During use, better-quality airborne equipment should be selected to ensure its own performance and quality. Airborne equipment with better product quality can reduce the overall failure rate of UAVs and play a key role in the normal use and integrity of UAVs.

2.4. Technical Quality of Control. UAV, as an unmanned aircraft, puts forward higher requirements on the proficiency of operators and their control capabilities. After systematic learning and training, and obtaining relevant level qualification certificates, technical personnel with certain driving experience and skilled manipulation play a good role in the overall success of airborne equipment and unmanned equipment, and will reduce the failure rate of airborne equipment. The technical quality factors of drone control will affect the direct or indirect failure rate of airborne equipment.

2.5. Comprehensive Airport Security Conditions. Airport security conditions include objective conditions such as the airport's runway quality, communications, navigation, power, and related facilities, all of which will have a certain impact on the failure rate of airborne equipment. If the airport's comprehensive support capability is stronger, the failure rate of airborne equipment will be correspondingly lower, while the airport's comprehensive support capability is weaker and the opposite is true.

2.6. Maintenance Quality. During the operation of airborne equipment, corresponding maintenance (external cleaning and internal dust removal, etc.), regular maintenance (weekly, monthly, and annual maintenance, etc.), and fault repairs are required. The maintenance quality determined by the abovementioned maintenance operation content, time interval, and the technical level of the maintenance personnel directly affects the success rate of the airborne equipment. Therefore, the maintenance quality can also be regarded as the main factor affecting the failure rate.

Since the failure rate of airborne equipment has many influencing factors and has a wide coverage, involving multiple dimensions and aspects, it is generally impossible to give all the parameters and related indicators that affect the failure rate of airborne equipment during the research process. Most of the key influences are selected. Factors are studied. The main influencing factors are flight time, number of take-offs and landings, control proficiency, abnormal environmental temperature, abnormal environmental humidity, and maintenance quality. These characteristic data are used as input vectors. At the same time, the failure rate of airborne equipment is expressed in thousands time failure rate calculation, that is, the number of failures occurring every 1000 h is used as the output vector of the failure rate

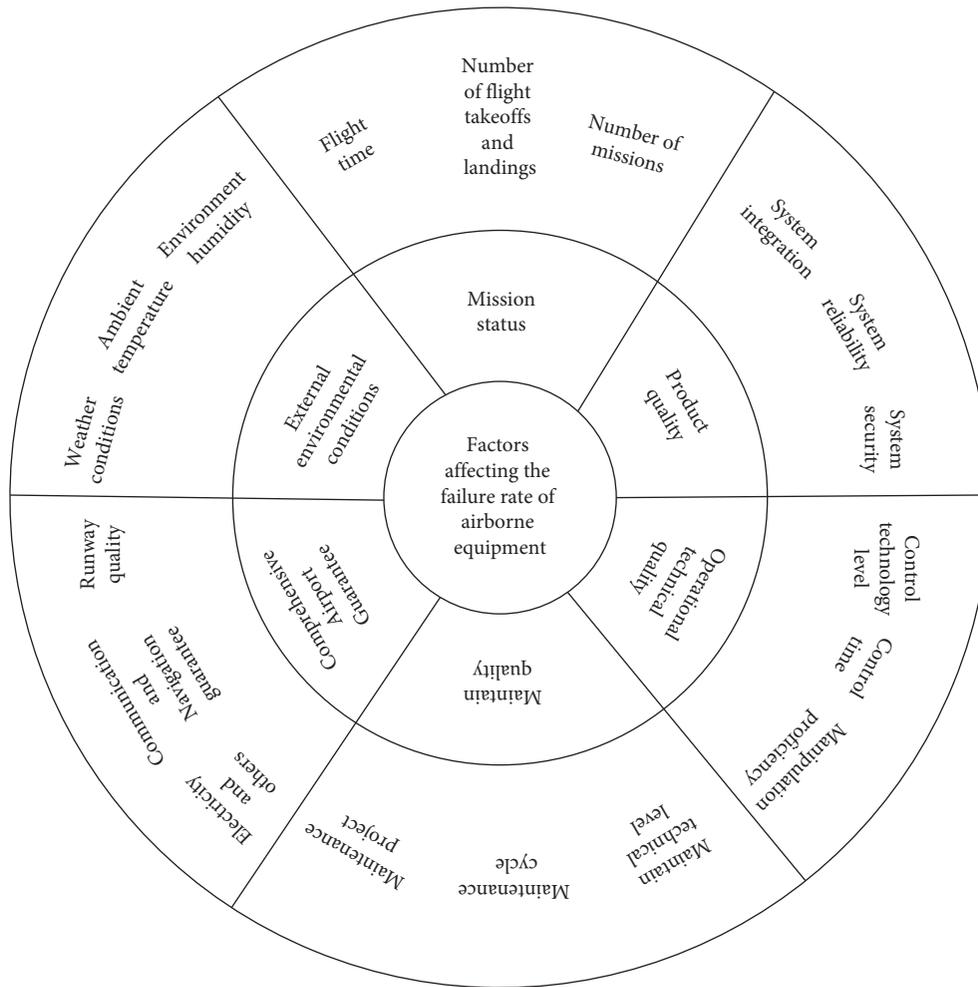


FIGURE 1: Factors affecting the failure rate of airborne equipment.

prediction model to carry out related research. Collecting information about the failure rate of airborne equipment with 25 failures from 2015 to 2018, the characteristic parameters, and data of the airborne equipment failures studied are shown in Table 1.

3. Combination Prediction Method for Failure Rate of Airborne Equipment

3.1. *Combined Model Modeling Process.* In order to study the failure rate prediction model of airborne equipment, an integrated hybrid combination model is used. The combined models proposed in this study include MLR model, GM(1, N) model, PLS model, BP model, SVM model, and make the optimal combination to form a combined forecasting model. The structure of the combined forecasting method is shown in Figure 2, and specifically includes the following four parts:

- (1) Information collection of failure rate and influencing factors: This stage analyzes and studies the failure of a drone's airborne equipment from 2015 to 2018, collects relevant data from the airborne equipment BIT, and also collects information about the failure rate from historical failure data. Information on

factors, including the number of failures, flight time, number of take-offs and landings, manipulation proficiency, abnormal temperature, abnormal humidity, maintenance quality, according to the degree of influence of each feature in the influencing factors on the failure rate, determine different influencing factors as characteristic data, and then the failure rate of the computer-borne equipment, the failure rate per thousand hours is regarded as the overall failure rate research.

- (2) The value of the influencing factor is regarded as an independent input variable, and the predicted failure rate value is used as the dependent output variable of the single model for research. Carry out modeling analysis and research one by one, and predict the failure rate of airborne equipment based on the relevant data of the input data model.
- (3) The realization of the combined model: At this stage, an innovative combined model is realized based on the five single models to predict the failure rate value (as the test data parameter) and the observed failure rate value (as the output parameter). Based on the optimal weighted combination modeling theory, the minimum

TABLE 1: The characteristics of the airborne equipment failures studied.

Parameter	Value
Types of airborne equipment (N)	1
Time period under study (year)	2015–2018
Flight time (h)	1000–2000
Number of take-offs and landings (N)	200–420
Operator proficiency (%)	60–95
Abnormal environmental humidity (%)	0.1–0.9
Abnormal environmental humidity (%)	0.1–0.9
Maintain quality (%)	70–95
Number of failures (N)	25
Failure rate per thousand hours N/1000 h	0.3–4.7

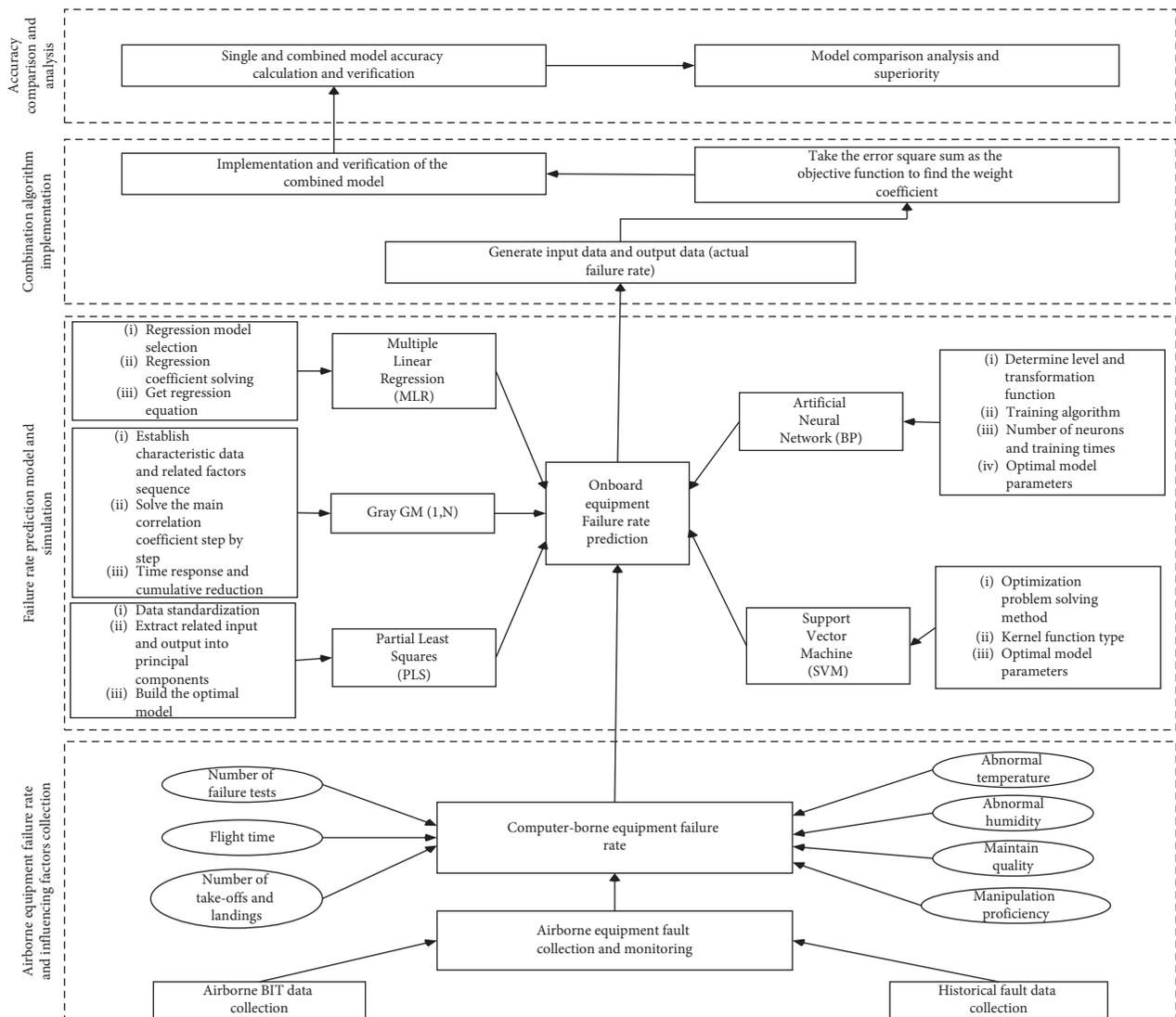


FIGURE 2: Failure rate prediction structure diagram.

error square sum of the combination model is used as the objective function to solve the optimal weighting coefficients, and the optimal weighted geometric average combination prediction model is derived to calculate the failure rate values of the airborne

equipment for the five individual predictions. The predicted failure rate is studied as the dependent output variable of the combined model, which provides a more effective model for accurately predicting the failure rate of airborne equipment.

- (4) Accuracy analysis and comparison of different models: Finally, evaluate the performance of the combined model and compare it with other single models. Using multiple indicators as evaluation criteria, the accuracy of the single model verification data is compared with the results of the proposed combined forecasting model, systematic comparison and analysis, and specific applications and verifications are carried out.

3.2. Single Model Analysis

3.2.1. Multiple Linear Regression (MLR) Model. The MLR analysis method is a commonly used analysis method to explain the change law of a single dependent variable depending on multiple independent variables, which is used to explore the relationship between independent variables and dependent variables. Assuming that there are n independent variables and y is a dependent variable, then there is $y = b_0 + b_1x_1 + b_2x_2 + \dots + b_nx_n$, where b_i is the regression coefficient. Assuming that β_i is the fitted value of b_i , the regression equation is obtained as the following formula:

$$\hat{y} = \beta_0 + \beta_1x_1 + \beta_2x_2 + \dots + \beta_nx_n \quad (1)$$

In formula (1), β_i is the partial regression coefficient, which can be solved according to the least square method and other optimization methods. Generally, the least square method can be used to solve the related regression parameters, and the required regression parameters need to be brought in. The regression equation is tested for statistics; usually, F test, T test, and multicollinearity test are used to determine whether the requirements are met. If the test requirements are not met, the independent variables need to be screened for re-modeling, and the output can be directly output until the requirements are met. Model, and finally obtain the regression coefficients of the multiple linear regression equation, and get the corresponding regression equation as formula:

$$\hat{y} = X\beta = X(X^T X)^{-1} X^T y. \quad (2)$$

3.2.2. Gray GM(1,N) Model. The gray system theory was first proposed by Professor Deng Julong to deal with “small samples, poor information” system prediction, and it has been widely used in many fields. The prediction steps of the gray GM(1, N) model are as follows:

Step 1. Use data to accumulate and generate calculations, in formula (3): $x_i^{(1)}(k) = \sum_{j=1}^k x_i^{(0)}(j)$, $(k = 1, 2, \dots, m)$.

$$X_i^{(1)} = \{x_i^{(1)}(1), x_i^{(1)}(2), \dots, x_i^{(1)}(m)\}. \quad (3)$$

Step 2. Generate a sequence of values immediately adjacent to the mean

Opposite the $X_i^{(0)}$ next to mean generation. $Z_1^{(1)} = (z_1^{(1)}(2), z_1^{(1)}(3), \dots, z_1^{(1)}(m))$

Among these: $z_1^{(1)}(k) = 1/2[x_1^{(1)}(k) + x_1^{(1)}(k - 1)]$, then the GM(1, N) model is: $\hat{x}_1^{(0)}(k + 1) + a\hat{z}_1^{(1)}(k + 1) = \sum_{i=2}^n b_i x_i(k + 1)$, the corresponding whitening equation is: $dx_1^{(1)}(k)/dt + ax_1^{(1)}(k) = \sum_{i=2}^n b_i x_i(k)$, among these: $z_1^{(1)}$ is the background value, a is the system development coefficient, and b_i is the system drive coefficient.

Step 3. Establish approximate time response

$$\text{Order } B = \begin{bmatrix} -z_1^{(1)}(2) & x_2^{(1)}(2) & \dots & x_n^{(1)}(2) \\ -z_1^{(1)}(3) & x_2^{(1)}(3) & \dots & x_n^{(1)}(3) \\ \vdots & \vdots & \dots & \vdots \\ -z_1^{(1)}(m) & x_2^{(1)}(m) & \dots & x_n^{(1)}(m) \end{bmatrix}, \quad (4)$$

$$Y = \begin{bmatrix} x_1^{(0)}(2) \\ x_1^{(0)}(3) \\ \vdots \\ x_1^{(0)}(m) \end{bmatrix}.$$

Then, the least square estimation parameter is listed as $\hat{a} = (B^T B)^{-1} B^T Y$, order the parameter list $\hat{a} = [a \ b_2 \ \dots \ b_n]^T$, the approximate time response is:

$$\hat{x}_1^{(1)}(k + 1) = \left[x_1^{(0)}(1) - \frac{1}{a} \sum_{i=2}^n b_i x_i^{(1)}(k + 1) \right] e^{-ak} + \frac{1}{a} \sum_{i=2}^n b_i x_i^{(1)}(k + 1). \quad (5)$$

In formula (5): $k = 1, 2, 3 \dots n - 1$.

Step 4. Accumulate and reduce the prediction model

$$\hat{x}^{(0)}(k + 1) = \hat{x}^{(1)}(k + 1) - \hat{x}^{(1)}(k). \quad (6)$$

3.2.3. Partial Least Squares (PLS) Model. Partial least squares model was first proposed by Wood (S. Wold) and Abano (C. Albano), etc., and it has been developed rapidly in theory, method, and application. The PLS model prediction steps are as follows:

Step 5. Data standardization processing

Standardize independent variables $X = (x_1, x_2, \dots, x_m)$ (m is the number of independent variables) and dependent variables Y . The standardization process is:

$$E_0 = x^* = \frac{(x_i - \bar{x}_i)}{S_{x_i}}, \quad (7)$$

$$F_0 = y^* = \frac{(y - \bar{y}_i)}{S_y}.$$

In formula (7): E_0, F_0 are the standardized vectors of x_i, y , respectively; \bar{x}, \bar{y} are the mean values of X, y , respectively; S_x, S_y are the mean square deviations of X, y respectively.

Step 6. Extract the input principal components t_h and output principal components u_h .

If there are m variables in the original data, use principal component analysis to recombine the information in the group of data, and extract h new comprehensive variables F_1, F_2, \dots, F_h ($h < m$) from them, so that the h comprehensive variables can summarize the information in the original data group at most, and take the new comprehensive variable as the principal component. The method of continuously extracting principal components t_h and u_h from the input variable E_h and output variable F_h is:

$$\begin{cases} t_h = E_{h-1} w_h & w_h = E_{h-1}^T F_{h-1} / \|E_{h-1}^T F_{h-1}\| \\ E_h = t_h p_h^T + E_{h-1} & F_h = t_h r_h^T + F_{h-1} \\ p_h = E_h^T t_h / \|t_h\|^2 & r_h = F_h^T t_h / \|t_h\|^2. \end{cases} \quad (8)$$

Step 7. Find the partial least squares model

Calculate in turn, if the rank of X is A , then we get:

$$\begin{aligned} E_0 &= t_1 p_1^T + \dots + t_h p_h^T, \\ F_0 &= t_1 r_1^T + \dots + t_h r_h^T + F_h. \end{aligned} \quad (9)$$

Equation (9) is reduced to the form of the regression equation of $y_k^* = F_{0k}$ on $x_j^* = E_{0j}$, and the regression equation is:

$$y_k^* = \alpha_{k1} x_1^* + \dots + \alpha_{kp} x_p^*. \quad (10)$$

3.2.4. Artificial Neural Network (BP) Model. In 1985, the BP algorithm proposed by Rumelhart systematically solved the problem of learning the connection weights of hidden units in multiple networks. At present, the BP model has become one of the important models of artificial neural networks and has been widely used.

The artificial neural network imitates the learning state of the human brain to judge and analyze things in the process of understanding things, and finally solves the problems in different fields. The model has good fault tolerance, and can describe the nonlinear relationship between variables and targets through its good nonlinear approximation ability, usually using BP neural network. The structure of the BP neural network is shown in Figure 3. The BP neural network is composed of an input layer, a hidden layer, and an output layer. The working process of BP neural network is divided into forward propagation of information flow and back propagation of error. During forward propagation, the data transfer direction is input layer \rightarrow hidden layer \rightarrow output layer, and each layer of neurons is only affected by the previous layer. If the error between the output data and the expected data is greater than the expected error, the back propagation of the error is entered. These two processes are performed alternately. The neural network performs gradient descent of the error function in the weight vector space, and dynamically iterates the weight vector values of the

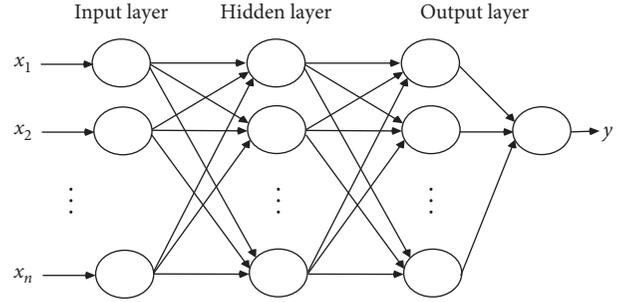


FIGURE 3: BP neural network structure.

neurons in each layer until the output data are within the error range and the calculation is stopped.

3.2.5. Support Vector Machine (SVM) Model. The SVM algorithm is a machine statistical learning theory based on minimizing structural risks. Through the learning strategy of maximizing the interval, it is finally transformed into the solution of a convex quadratic programming problem. SVM has the advantages of adapting to small sample learning and strong robustness.

The estimation function of the support vector machine is as follows:

$$f(x) = w^T \varphi(x) + b. \quad (11)$$

In formula (11), w is the normal vector and b is a constant. The value of w and b are obtained by minimizing and by the objective function. The objective function is as follows:

$$\begin{cases} R(w) = \frac{1}{2} w^T w + C \sum_{i=1}^n |y_i - f(x_i)|_\varepsilon \\ |w^T \varphi(x_i) + b - y_i| \leq \varepsilon. \end{cases} \quad (12)$$

In formula (12), the penalty parameter C , the insensitive loss function ε , ξ_i and ξ_i^* are the introduction of slack variables and the above formula are transformed into the following:

$$\begin{cases} \frac{1}{2} w^T w + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ (w^T \varphi(x_i) + b) - y_i \leq \varepsilon + \xi_i \\ y_i - (w^T \varphi(x_i) + b) \leq \varepsilon + \xi_i^* \\ \xi_i \geq 0, \xi_i^* \geq 0 \end{cases}. \quad (13)$$

Lagrangian function is introduced to equation (13), $\partial L / \partial w = 0$; $\partial L / \partial b = 0$; $\partial L / \partial \xi_i = 0$; $\partial L / \partial \xi_i^* = 0$. After solving, the following formula is obtained:

$$\left\{ \begin{array}{l} \max -\frac{1}{2} \sum_{i,j=1}^n (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*)K(x_i, x_j) + \\ \sum_{i=1}^n (\alpha_i - \alpha_i^*)y_i - \sum_{i=1}^n (\alpha_i - \alpha_i^*)\varepsilon \\ \text{s.t. } \sum_{i=1}^n (\alpha_i - \alpha_i^*) = 0; 0 \leq \alpha_i; \alpha_i^* \leq C \end{array} \right. \quad (14)$$

After solving the α value, the regression function is obtained:

$$f(x) = \sum_{i=1}^n (\alpha_i - \alpha_i^*)K(x_i, x) + b. \quad (15)$$

In formula (15), α_i and α_i^* are Lagrangian coefficients, and $K(x_i, x)$ is the kernel function.

3.3. Propose an Optimal Combination Model with the Error Sum of Squares as the Goal. Due to the complexity of certain phenomena and the diversity of various factors, the performance of single-term models in forecasting is always limited, so a combined forecasting model is established. Combination forecasting model is to use appropriate methods to combine multiple single forecasting models, comprehensively process the forecast results of various single forecasting models, and generate a total forecasting model containing the forecasting information of multiple forecasting models. Due to the complexity of the failure rate of airborne equipment and the influence of various factors on this phenomenon, the use of a single model has certain limitations in predicting the failure rate, which reduces the accuracy of the single prediction model. For this reason, five different single models are used in this study, including MLR model, GM (1, N) model, PLS model, BP model, and SVM model. On the basis of single prediction, a combined prediction method with the sum of squared errors as the objective function is proposed for combined prediction. At the same time, the following methods are used to solve the weights using the optimal combination model prediction method to select an objective function that can describe the prediction error, and then minimize the objective function to determine the optimal weight. Combining the failure rate prediction error of the five algorithms of MLR model, GM (1, N) model, PLS model, BP model, and SVM model, the error function can be determined. The sum of squares of the prediction errors is the objective function to construct the optimal combination model. Calculate the corresponding optimal weight. $x_i(t)$ are the prediction values of MLR, GM(1,N), PLS, BP, and SVM algorithms at time t , respectively; the prediction error at time t can be obtained by $e_{it}(t) = x(t) - x_i(t)$; forecast error information matrix is $E = [(e_i)_{n \times m}] [(e_i t)_{n \times m}]^T$; five kinds of prediction results are analyzed and calculated with unequal weight combination; $W = [w_1, w_2, \dots, w_n]^T$ are the corresponding output weight values of the two prediction methods, and the combined

prediction result is $\hat{X}(t) = \sum_{i=1}^n (w_i x_i(t))$, and $\sum_{i=1}^n w_i = 1$. Then, the sum of squared errors of the prediction model is: $S = \sum_{t=1}^m e_t^2 = \sum_{t=1}^m (\sum_{i=1}^n w_i e_{it})^2 = W^T E W$, the quadratic programming model will be solved with the minimum sum of square error S to determine the optimal weight as: $\min S = W^T E W$, s.t. $R^T W = 1$, and $R = [1, 1, \dots, 1]^T$, and by introducing the Lagrange multiplier λ , we get formula:

$$W^T E W - 2\lambda(R^T W - 1) = 0. \quad (16)$$

Calculate the derivation of W and λ , respectively, to obtain the $W = \lambda E^{-1} R$, $\lambda = 1/R^T E^{-1} R$. Finally, the corresponding optimal weight can be obtained as: $W = E^{-1} R / R^T E^{-1} R$.

The execution steps of the combined model are shown in Figure 4.

- (1) Collect detailed data on key factors affecting the failure rate of airborne systems, including flight time, number of take-offs and landings, manipulation proficiency, abnormal environmental temperature, abnormal environmental humidity, maintenance quality and failure rate data per thousand hours, and divide the collected data set into two parts; one part is used as test input data, and the other part is used as inspection data.
- (2) Use a single model to predict the failure rate of airborne equipment. The models in MLR, GM (1,N), PLS, BP, and SVM 5 will be used to predict the failure rate to obtain the corresponding actual and predicted values of the failure.
- (3) Based on five single prediction models, an optimized combination model with the error sum of squares as the goal is established. The failure rate value predicted by the single model selected in the previous step is the input variable of the combined model, and the combined model is formed to generate input and sum. Check the data set.
- (4) Determine the combination function on the basis of the combination weight, make an overall prediction of the failure rate, and use the corresponding accuracy evaluation indicators for comparison and analysis.

3.4. Model Evaluation Index. Evaluating the predictive performance of a predictive model requires measurement and analysis of the predictive results. Different evaluation indicators can be used, and different indicators have different emphasis and scope of adaptation [38–40]. This study uses five evaluation indicators to evaluate the applicability and accuracy of the model in the failure rate of airborne equipment, and to measure the error between the predicted value and the true value, so as to test the effectiveness of single and combined forecasting methods. Assume that the predicted value is $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ and the true value is $y = \{y_1, y_2, \dots, y_n\}$.

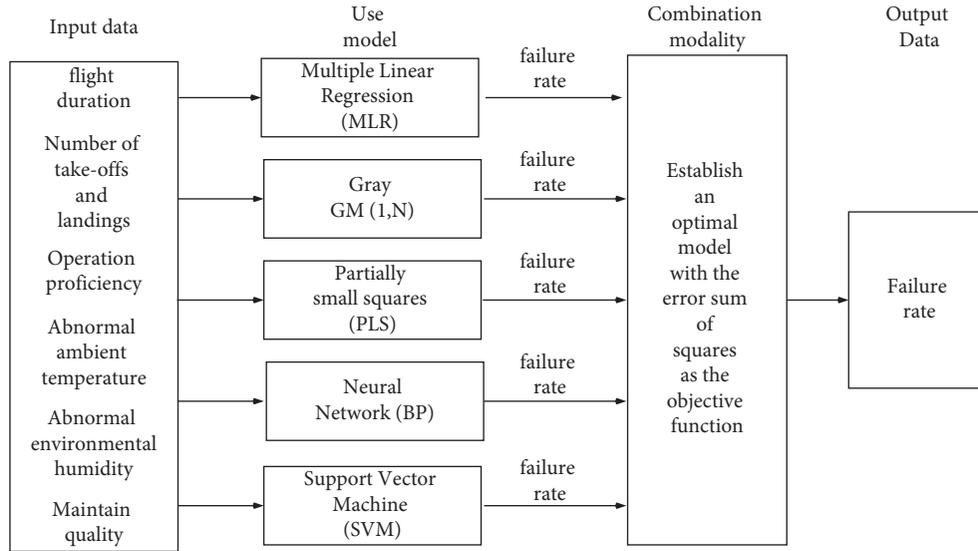


FIGURE 4: Execution steps of the combined model.

Mean absolute error: $MAE = 1/n \sum_{i=1}^n |\hat{y}_i - y_i|$

Root mean square error: $RMSE = \sqrt{1/n \sum_{i=1}^n (\hat{y}_i - y_i)^2}$

Mean absolute percentage error: $MAPE = 1/n \sum_{i=1}^n |\hat{y}_i - y_i|/y_i \times 100\%$

Normalized root mean square error: $MSPE = 1/n \sqrt{\sum_{i=1}^n (\hat{y}_i - y_i/y_i)^2}$

Decisive factor: $NRMSE = \sqrt{n \sum_{i=1}^n (\hat{y}_i - y_i)^2 / (\sum_{i=1}^n \hat{y}_i)^2}$

$R^2 = 1 - \sum_{i=1}^n (y_i - \hat{y}_i)^2 / \sum_{i=1}^n (y_i - \sum_{i=1}^n y_i/n)^2$

According to the definition of the above indicators, MAPE represents the average error of multiple prediction results and is used to evaluate the fitting performance of the model to existing data. The smaller the value, the higher the prediction accuracy. MSPE is also a method to reflect the error of multiple prediction values, but it can reflect the degree of dispersion of the error. The smaller the value, the smaller the fluctuation of the error to a certain extent. And, MAE, RMSE, and NRMSE are also used to test the deviation and volatility between the actual value and the predicted value. The smaller the value, the higher the accuracy of the prediction model. At the same time, the determination coefficient R^2 is the ratio of the change explained by the independent variable to the total change. The value of R^2 is [0,1]. The closer R^2 is to 1, the higher the prediction accuracy.

4. Model Research and Discussion

Take the statistical data of the UAV flight control system as an example to carry out the study of single and combined

forecasting models. In this study, in order to measure the accuracy of different models and prevent over fitting, the collected data are divided into two subsets: Test input data set (including 70% of data) and inspection data set (including 30% of data). Use the sample 1–25 group working time distribution data of 1000–2000 hours as an example for analysis. The first 17 sets of data of samples 1–17 are selected as modeling input, and the last 8 sets of data of samples 18–25 are data whose working hours are distributed between 1500 and 2000 hours for predictive model testing, and the MLR model and GM(1, N) are established, respectively. Model, PLS model, BP model, SVM model, taking flight time, number of take-offs and landings, manipulation proficiency, abnormal environmental temperature, abnormal environmental humidity, and maintenance quality as independent variables, and the thousand-hour failure rate as the dependent variable to obtain sample data is shown in the Figure 5. Related models are established and studied.

4.1. MLR Model Analysis. The MLR model process is shown in Figure 6.

Taking flight time (X1), number of take-offs and landings (X2), maneuvering proficiency (X3), abnormal environmental temperature (X4), abnormal environmental humidity (X5), and maintenance quality (X6) as the input independent variables of the multiple linear regression, it affects the machine. The thousand-hour failure rate (Y) of the load equipment failure rate is used as the dependent variable to establish a multiple linear regression equation. The MLR model obtained according to the steps described in section 3.2.1 is:

$$y = -2.2367 + 0.0004X1 + 0.0131X2 + 0.0044X3 - 1.4078X4 - 0.7413X5 + 0.0107X6. \quad (17)$$

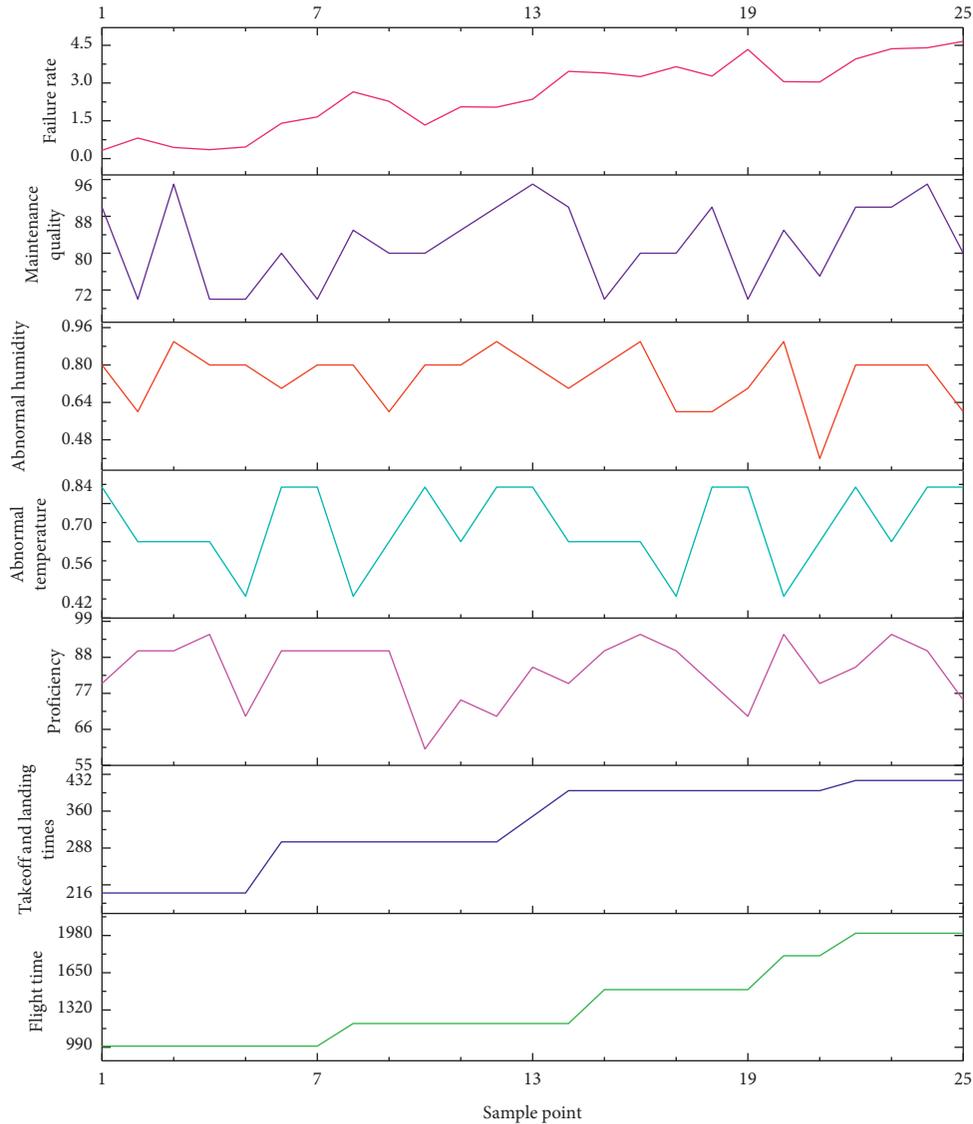


FIGURE 5: Collect 25 data sets.

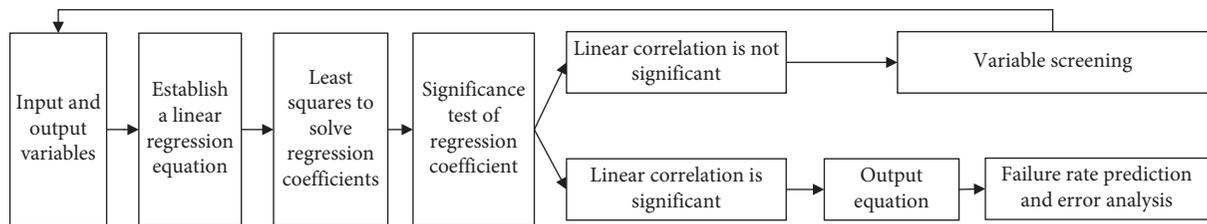


FIGURE 6: The MLR model process.

When performing statistical tests on formula (17), it is found that the variance inflation factor (VIF) corresponding to the failure rate (Y) is 13.94, and the VIF corresponding to the number of ups and downs (X2) is 13.26, indicating that the model is built There is serious multicollinearity (VIF >10), which may bring large errors to the regression results. Therefore, the input variables of the model need to be revised to eliminate multicollinearity.

The common method to eliminate multicollinearity is to eliminate the variable corresponding to the maximum value of VIF, namely, the number of take-offs and landings (X2). The independent variables of the MLR correction model after elimination are flight time (X1), handling proficiency (X3), abnormal ambient temperature (X4), abnormal ambient humidity (X5), and maintenance quality (X6). The MLR correction model obtained

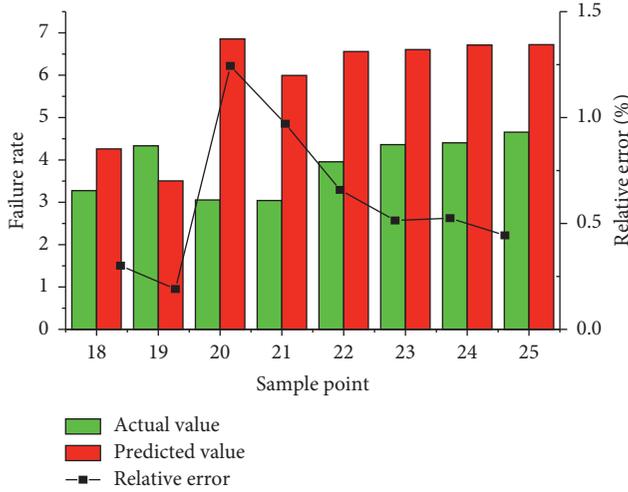


FIGURE 7: Partial failure rate prediction results of MLR model.

according to the steps described in Section 3.2.1 is model obtained according to the steps described in section 3.2.1 is:

$$y = -5.3841 + 0.0054X_1 + 0.0094X_3 + 0.2445X_4 - 2.3446X_5 + 0.0212X_6. \quad (18)$$

Through the test, formula (18) has eliminated the multicollinearity, and all the indexes of the F test and the T test meet the significance condition. The revised MLR model is used to predict the failure rate of 8 groups of airborne equipment in samples 18–25, and some of the predicted values and relative errors are shown in Figure 7.

4.2. GM(1,N) Model Analysis. The GM(1, N) model process is shown in Figure 8.

Select failure rate, flight time, number of take-offs and landings, manipulation proficiency, abnormal environmental temperature, abnormal environmental humidity, and maintenance quality as independent variables. Time series are the original data series:

$$X_j^{(0)} = \{X_1^{(0)}, X_2^{(0)}, X_3^{(0)}, X_4^{(0)}, X_5^{(0)}, X_6^{(0)}, X_7^{(0)}\}. \quad (19)$$

When the failure rate data sequence X1 is the system characteristic data sequence, flight time X2, takeoff and landing times X3, operator proficiency X4, abnormal ambient temperature X5, abnormal ambient humidity X6, and maintenance quality sequence X7 are all related factor sequences. Select samples 1–17 sets of data for modeling according to the steps described in 3.2.2, set the whitening equation of the GM(1,7) model as $dx_1^{(1)}/dt + ax_1^{(1)} = b_2x_2^{(1)} + b_3x_3^{(1)} + b_4x_4^{(1)} + b_5x_5^{(1)} + b_6x_6^{(1)} + b_7x_7^{(1)}$. Use China Southern Airlines Gray System Theory and

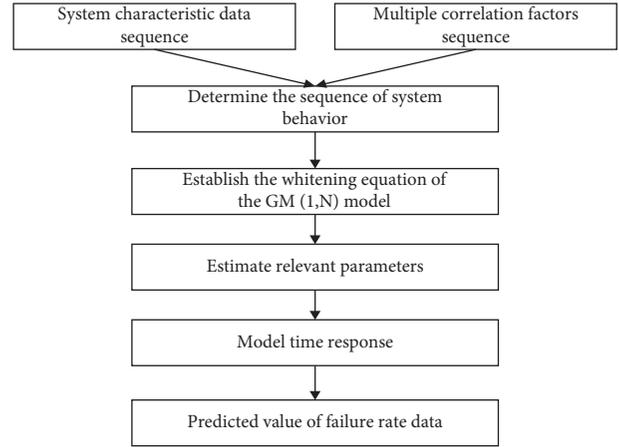


FIGURE 8: The GM(1,N) model process.

Application Software (GSTA V7.0) to analyze the data of samples 1–17 in Table 1, and obtain estimated parameters: $\hat{a} = [a \ b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ b_6 \ b_7]^T = [2.133 \ 0.545 \ 0.031 \ 0.056 \ 0.143 \ 0.532 \ 0.342 \ 0.251]$. The estimated model and the approximate time response are obtained as well as the GM(1,7) prediction model through cumulative reduction. On this basis, the prediction values of the last 8 groups of test data are obtained.

4.3. PLS Model Analysis. The PLS model process is shown in Figure 9.

Based on the calculation of the PLS modeling program, the cross-validity of each component is extracted, and the target of multivariate correction is directly positioned on the prediction. Therefore, the principle of determining the number of extracted principal components is to minimize the prediction error PRESS. Based on this, it is determined to extract 6 principal components. It can cover most of the information of the original data, and the model is more reasonable. According to the principle of minimum prediction error PRESS and the principle of cross-validity $Qh^2 \geq 0.0975$, the number of extracted principal components is the same in most cases. However, if the sum of squares of prediction errors (PRESS) is not much different between the extracted principal components, the principal components should be selected by referring to the size of the sum of squares of fitting errors (SS). The mathematical model of partial least squares standardized data obtained by calculation is:

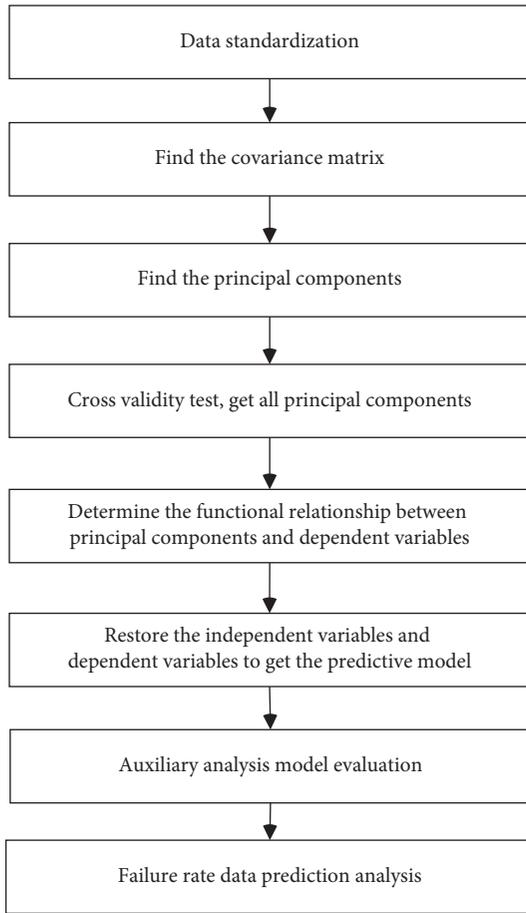


FIGURE 9: The PLS model process.

$$Y = 0.116X_1 + 0.096X_2 + 0.306X_3 + 0.108X_4 + 0.213X_5 + 0.221X_6 \tag{20}$$

And, we also get the original mathematical model as

$$Y = 8.281 + 0.675X_1 + 0.1091X_2 + 0.542X_3 + 0.115X_4 + 0.179X_5 + 0.635X_6 \tag{21}$$

And, the value of VIP_j , the projected importance index of the variable, is obtained: $X_1 = 1.559$, $X_2 = 1.718$, $X_3 = 0.366$, $X_4 = 0.477$, $X_5 = 0.421$, and $X_6 = 0.269$. The variable projection table is shown in Figure 10. The effect of each independent variable on the explanatory dependent variable is as follows: $X_2 > X_1 > X_4 > X_5 > X_3 > X_6$. According to the principle of $VIP_j > 1$ that X_j has an important role in explaining the dependent variable, flight take-offs and landings X_2 and flight time X_1 play an important role in interpreting the number of failure rates Y (number) of airborne equipment in the set of dependent variables, including the number of flight take-offs and landings. X_2 has the greatest effect on explaining the failure rate Y (number) of airborne equipment. Eight groups of independent variables of test data are brought into the equation, and the predicted values of eight groups of data are obtained on this basis.

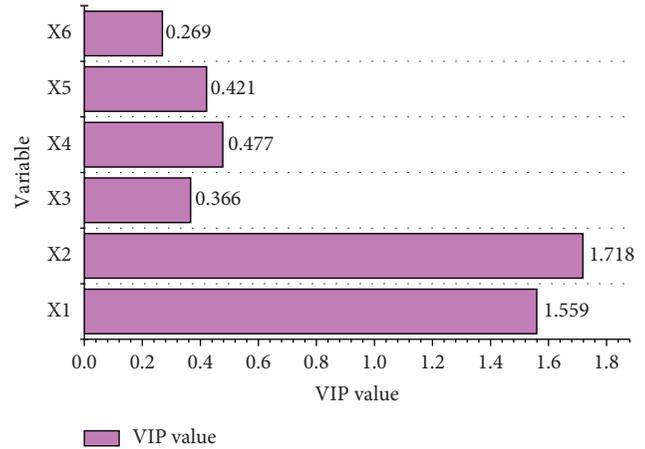


FIGURE 10: Variable projection table.

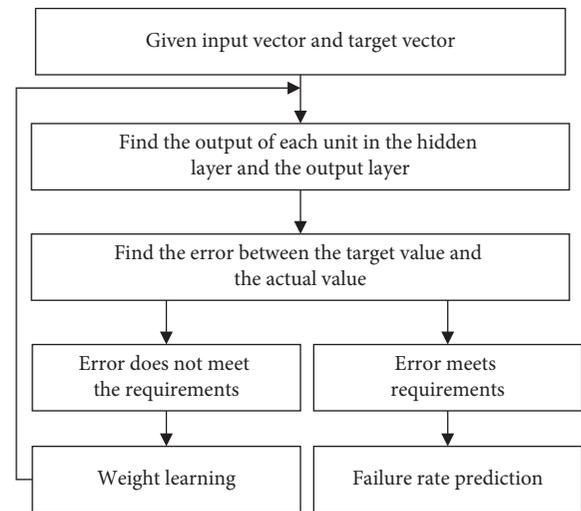


FIGURE 11: The BP model process.

4.4. BP Model Analysis. The BP model process is shown in Figure 11.

Using BP artificial neural network algorithm, created and implemented in MATLAB, first determine the number of input and output neurons. Since the failure rate of airborne equipment includes flight time, number of take-offs and landings, manipulation proficiency, abnormal environmental temperature, abnormal environmental humidity, and maintenance quality, it can be determined that there are 6 input layers in the neural network model. For each neuron, the output result only includes one output of failure rate, so the number of neurons in one output layer is determined. Secondly, determine the number of neurons in the hidden layer. How to determine the number of neurons in the hidden layer that enables the neural network to output the most accurate number of neurons in the hidden layer has not been given a definite rule in theory until now. In most cases, the trial and error method is used to determine the number of hidden layer nodes. According to preliminary experiments, when the hidden layer is 3 layers and the number of neurons is greater than or equal to 8, the training error has

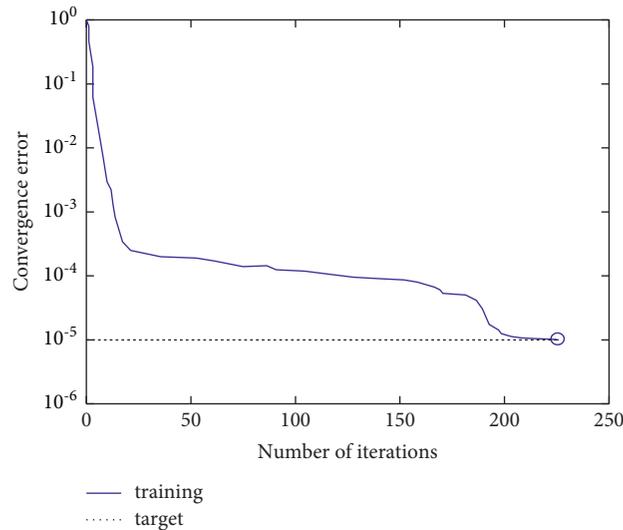


FIGURE 12: BP neural network output error change diagram.

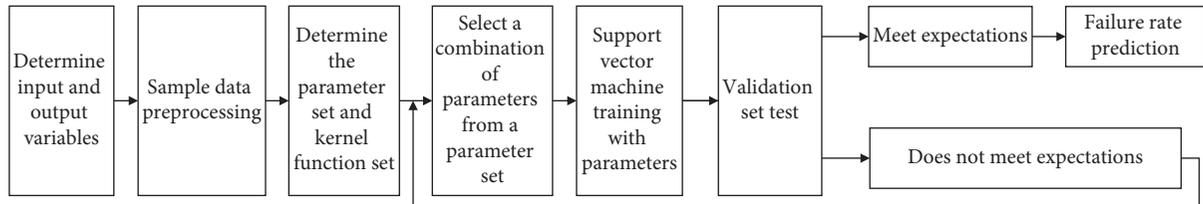


FIGURE 13: The SVM model process.

stabilized, so the number of hidden layer nodes in this model is 8. In the selection of activation function, the activation function of BP neural network includes logarithmic sigmoid (Log-sigmoid), hyperbolic tangent function (Tan-sigmoid), and linear function (Purelin) 3 types, this model is in MATLAB software. The logsig function is used as the activation function. The target convergence error is set to 0.00001, the maximum number of training times is set to 300, and the error function is selected as the variance performance analysis function (MSE). Therefore, the established neural network is a three-layer BP neural network with 6 neurons in the input layer, 8 neurons in the middle layer, and 1 neuron in the output layer. Input 1–17 sets of training data into the network for training, and the output error of the BP neural network is obtained as shown in Figure 12. After 226 iterations, the output error is less than the convergence error. Based on this algorithm, 8 sets of test data prediction values are obtained.

4.5. SVM Model Analysis. The SVM model process is shown in Figure 13.

Select the sequence data of sample 1–17 sets of data such as flight time, number of take-offs and landings, manipulation proficiency, abnormal environmental temperature, abnormal environmental humidity, maintenance quality, failure rate, and other sequence data to train the SVR model, and use the trained model to predict 18–8 groups of airborne equipment failure rate of 25. The gamma parameter g and

the penalty factor C of the support vector machine kernel function are the parameters that need to be determined in the model. The initial ranges of the given parameters C and g are both $[-10, 10]$. The optimal parameters C and g of the SVR model are $C = 1.325$, and $g = 0.256$ by using the interactive test method. According to the optimal parameters C and g , the SVR model was trained to predict the 8 sets of data and obtain 8 sets of test data prediction values.

4.6. Combination Algorithm Coefficient Solving and Construction. The combination algorithm coefficient solving and construction process is shown in Figure 14.

Based on the prediction of the abovementioned single model, the prediction values of the corresponding failure rates of 8 groups were obtained. Assuming that the predicted value of airborne equipment failure rate obtained by MLR model is \hat{y}_{1t} , the predicted value of airborne equipment failure rate obtained by GM(1, N) model is \hat{y}_{2t} , the predicted value of airborne equipment failure rate obtained by PLS model is \hat{d} , and the predicted value of airborne equipment failure rate obtained by BP model is \hat{y}_{3t} . The predicted value of equipment failure rate is \hat{y}_{4t} , and the predicted value of airborne equipment failure rate obtained by SVM model is \hat{y}_{5t} ; without increasing complexity, based on the optimal weighted combination modeling theory, the minimum error square sum of the combined model is used as the objective function to solve each optimal weighting coefficient, $w_1 = 0.093$, $w_2 = 0.212$, $w_3 = 0.321$, $w_4 = 0.209$, $w_5 = 0.165$;

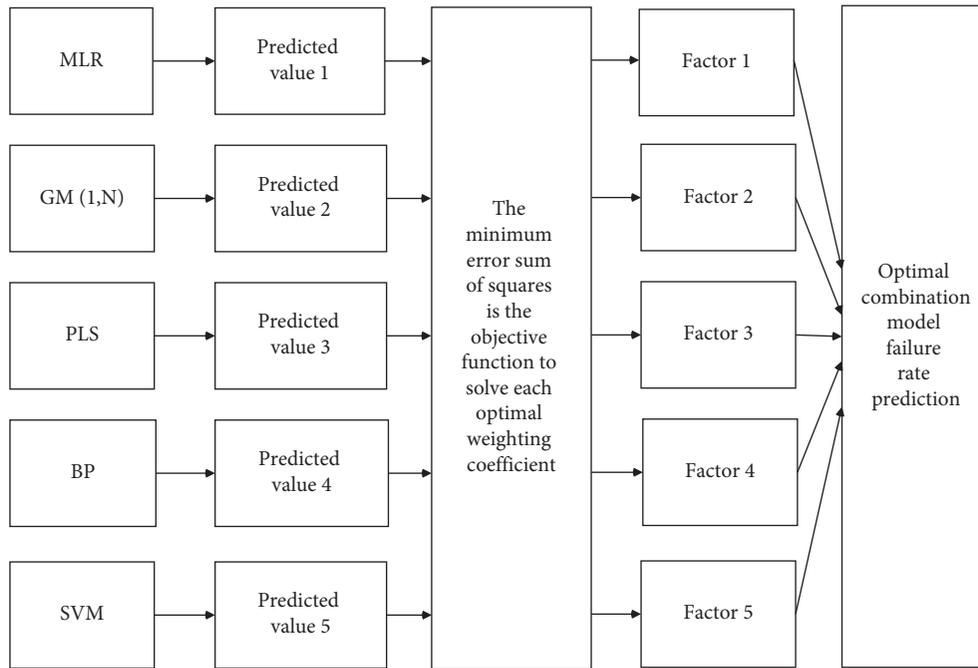


FIGURE 14: The combination algorithm coefficient solving and construction process.

therefore, the available combined model expression is: $\hat{y}_t = w_1\hat{y}_{1t} + w_2\hat{y}_{2t} + w_3\hat{y}_{3t} + w_4\hat{y}_{4t} + w_5\hat{y}_{5t}$. Bringing in the data obtained from the single model, the predicted value of 8 sets of test data is obtained through the combined model.

5. Comparative Analysis of Prediction Results of Various Models

The failure rate of the last 8 groups of data is predicted by combining the single item and combined prediction models, and the real-value failure rate and the corresponding failure rate prediction value change trend of each model are shown in Figure 15:

Figure 15 shows that the 20–25th sample point of the MLR model has a large deviation from the actual failure rate value. When establishing the MLR model equation, in order to meet the requirements of the statistical test, the number of ups and downs of one of the independent variables (X2) is caused by discarding. In MLR analysis, due to various reasons, such as incomplete consideration of the problem, meeting an indicator of the system, or when the explanatory variable cannot be measured or observed, some explanatory variables are often lost or omitted in the regression equation, and sometimes some explanatory variables that should not be included are included in the regression equation. Whether it is the loss or omission of explanatory variables, or the incorrect addition of explanatory variables, it will have an adverse effect on the authenticity of the regression equation and regression coefficients, and even obvious deficiencies and contradictions. This situation should be avoided in the engineering as much as possible. Otherwise, it will lead to poor forecasting effect, and the failure rate forecast needs to be further improved. Using other single prediction

models, it can be seen from Figure 15 that the predicted value of the airborne equipment failure rate required by different models is similar to the actual value. In the single prediction model, the deviation between the actual value and the predicted value of the BP model is small. The degree of compatibility is higher. However, the combined model has the highest degree of fit, and the deviation between the actual value and the predicted value is also the smallest.

Use the constructed prediction model to systematically predict the input of 1–17 groups of samples, and analyze and evaluate the accuracy and performance of various models for predicting failure rates through 6 evaluation indicators. Calculate the corresponding average absolute error (MAE), root mean square error (RMSE), average absolute percentage error (MAPE), mean square percentage error (MSPE), normalized root mean square error (NRMSE) coefficient of determination (R^2), and get the input data model and error as shown in Table 2.

From the comparison of the index data in Table 2, it can be seen that the MAE value of the BP model in the five single models is smaller, while the RMSE, MAPE, MSPE, and NRMSE values of the PLS model are the smallest, and the coefficient of determination R^2 of the BP model is the largest. The R^2 of the PLS model is second, indicating that the single model uses the BP model and the PLS model to predict deviation and volatility, and has the highest prediction accuracy among the individual models. In addition, it can be seen from Table 2 that in the single-phase prediction model, the GM(1, N) model also has good prediction accuracy, and all indicators are similar to the BP model and PLS model.

In order to choose the best prediction model, we compared the error indices of the verification data in

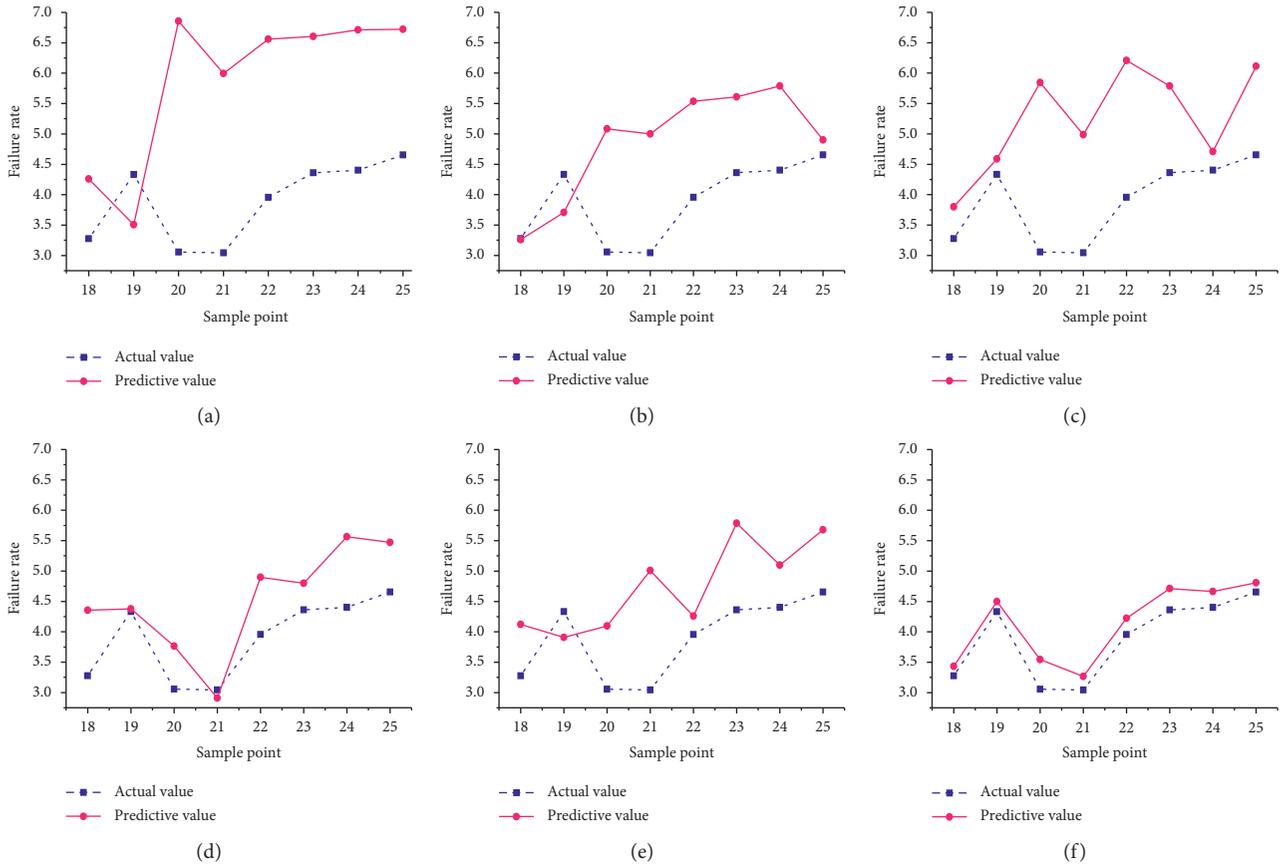


FIGURE 15: Comparison of predicted value and actual value of airborne equipment failure. (a) MLR Model, (b) GM (1, N) Model, (c) PLS Model, (d) BP Model, (e) SVM Model, and (f) Combined model.

TABLE 2: Input data model and error data table.

Model and deviation	MAE	RMSE	MAPE	MSPE	NRMSE	R^2
MLR	0.3782	0.4894	41.37	1.4749	0.5231	0.899
GM(1,N)	0.3363	0.3659	34.38	1.1408	0.3896	0.910
PLS	0.3359	0.2013	15.21	0.4199	0.2056	0.928
BP	0.1749	0.4184	28.31	1.0094	0.3092	0.964
SVM	0.4514	0.5248	48.21	1.5398	0.4271	0.909

different models, and selected 8–25 groups of the last 8 groups of test samples for single model and combined model for error calculation. The corresponding MAE, RMSE, MAPE, MSPE, NRSME, and R^2 using different models are obtained, respectively, as shown in Figures 16–21.

The comparison of the error index between the combined model and the five individual models in Figures 18 and 21 shows that the coefficient of determination R^2 index rate of the combined model has increased by 1.4% (compared with the BP model) to 10% (compared with the MLR model)). It also reduced the MAPE index rate by 59.7% (compared to the BP model) to 88.4% (compared to the MLR model). Moreover, RMSE is the square root of the ratio of the square of the deviation between the predicted value and the true value to the number of observations n . The smaller the RMSE index of the root mean square error, the stronger

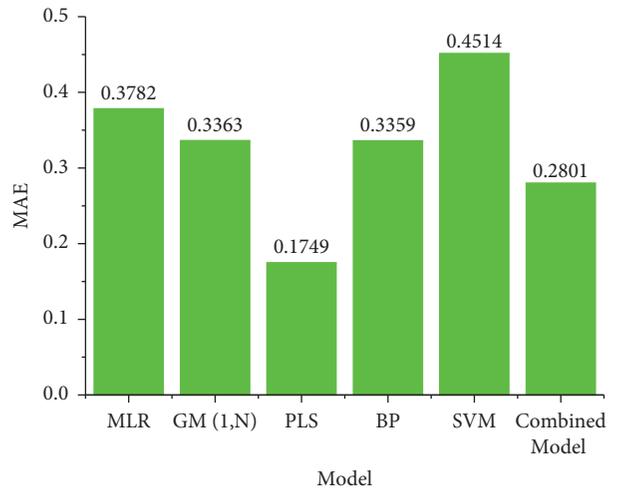


FIGURE 16: Comparison chart of single and combined model MAE.

the predictive ability of the model. As shown in Figure 17, the RMSE value is 0.2803, which is the smallest compared to other models. At the same time, the MAE error index is a measure of the average absolute deviation between the predicted value and the observed value. The lower the value of the index, the higher the accuracy and precision of the model. The MAE index value of the verification data

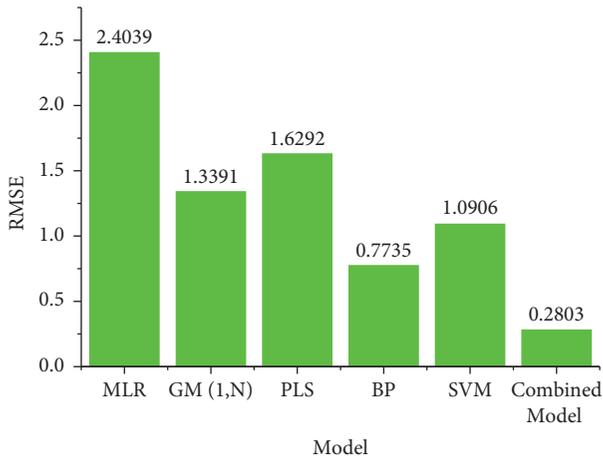


FIGURE 17: Comparison chart of single and combined model RMSE.

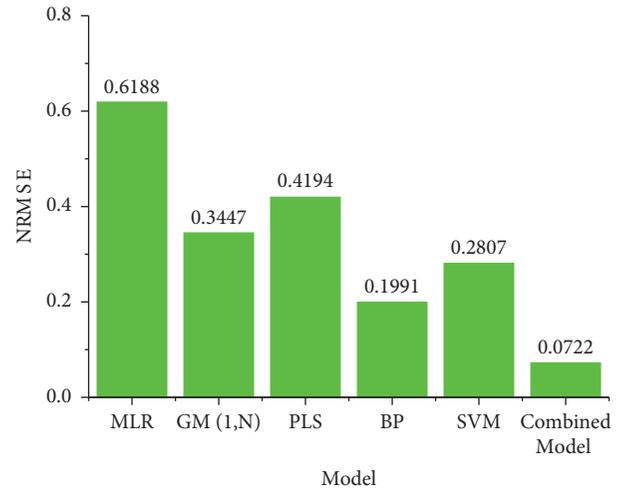


FIGURE 20: Comparison chart of single and combined model NRMSE.

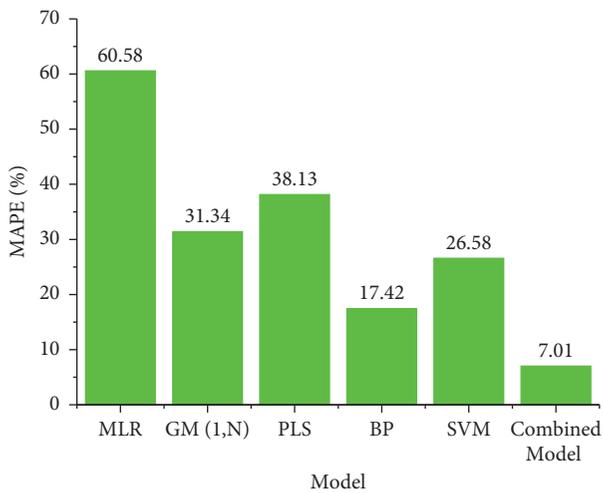


FIGURE 18: Comparison of single and combined model MAPE.

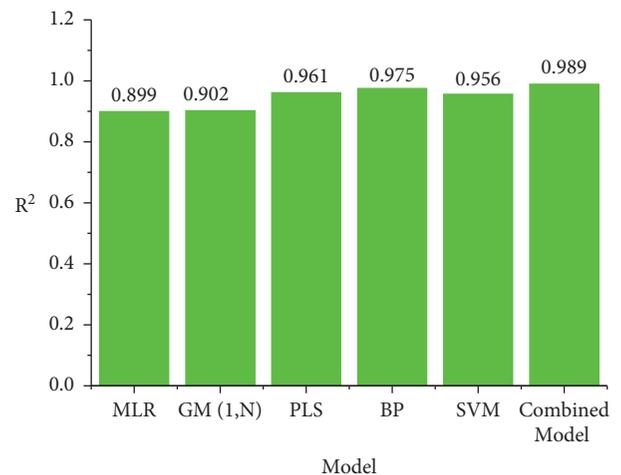


FIGURE 21: Comparison chart of single and combined model R².

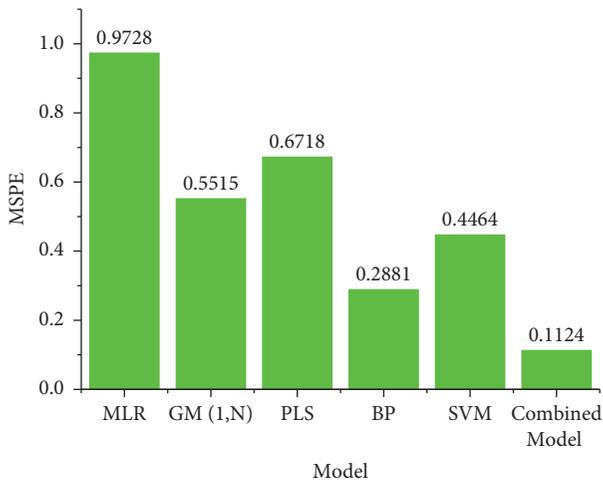


FIGURE 19: Comparison chart of single and combined model MSPE.

according to the combined model in Figure 16 is 0.2801. According to Figure 19, the MSPE index of the combined model is 0.1124, and the NRMSE index of the combined model of Figure 20 is 0.0722, which is the smallest compared with other models, which also shows that the combined model has better effects. Therefore, the combined model proposed in this paper improves the performance and accuracy of airborne equipment failure rate prediction.

The above error indicators only show the average value of the errors of the prediction model. In order to evaluate the model more effectively, we use the airborne equipment failure rate of the model in the inspection phase and the Theil inequality coefficient of the observations for analysis and research, which makes the prediction performance evaluation index reach seven classic functions. Theil inequality coefficient is used to calculate the consistency index of two

sequences. When the Theil inequality coefficient is closer to 0, the two sequences are more consistent, that is, the closer the predicted value is to the actual value, the better the model fitting effect. When it is closer to 1, the second sequence is more inconsistent (in 2021, Zhou Zhou, etc.) [41], which is:

$$\text{Theil} = \frac{[T^{-1} \sum_{i=1}^T (y - \hat{y})^2]^{0.5}}{[T^{-1} \sum_{i=1}^T (y)^2]^{0.5} + [T^{-1} \sum_{i=1}^T (\hat{y})^2]^{0.5}} \quad (22)$$

Among them, $y = \{y_1, y_2, \dots, y_n\}$ represents the actual value, $\hat{y} = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ represents the predicted value, and T represents the sample size. In order to calculate the Theil inequality coefficient of the model at the test stage, T is selected as 8, and the Theil inequality coefficient of each model is shown in Figure 22.

It can be seen from Figure 22 that the Theil inequality coefficient of the combined model is 0.034, which is the lowest compared with the single model, indicating that the prediction error of the combined model is smaller, and that it has accuracy, performance, and higher reliability.

In order to verify the accuracy of the combined model, the ARIMA model [42], the GRNN model [43], and the EMD-SVM model [44] methods are used for comparison and analysis. The accuracy indexes of different models are shown in Table 3.

Table 3 provides a comprehensive comparison of the prediction accuracy indicators between the proposed model and the ARIMA model, GRNN model, and EMD-SVM model. It can be seen from the table that compared with other models, the seven accuracy indicators of the combined model are all smaller. Therefore, it can be verified that the proposed model is better than other models. In other words, the proposed optimal combined model has better prediction performance.

6. Applicability Analysis of Different Models

Airborne equipment has the characteristics of small batches, multiple varieties, complicated system cross-links, and strong random failures, resulting in fewer data sources for failure rate information samples of airborne equipment, lack of effective fault characteristics, and diversified faults, making it difficult for the system to fail. Troubleshoot. It is difficult to accurately predict the failure rate. Due to the particularity of the prediction object and the complexity of the conditions, the used prediction model is not fixed. MLR model not only has simple structure and convenient calculation but also has the characteristics of wide application and strong applicability, which can cover the failure rate prediction applications of more airborne equipment. The GM (1, N) model can take advantage of the less historical sample data required for gray system prediction and high prediction accuracy. At the same time, it does not have to have typical distribution characteristics and regular requirements for the original data. It is suitable for airborne equipment with less original information and data. Lack of airborne equipment failure rate predictions. The PLS model can not only fully consider the multi-factor information that

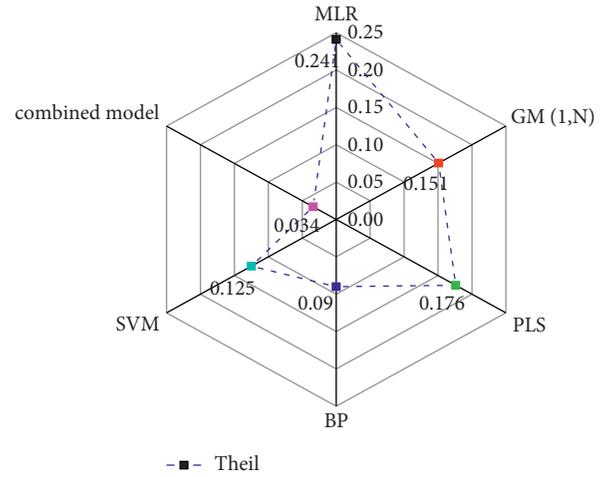


FIGURE 22: Comparison between single item and combined model Theil.

TABLE 3: Different model accuracy data.

Accuracy index	Model comparison			
	ARIMA	GRNN	EMD-SVM	Combined model
MAE	0.4628	0.7461	0.3167	0.2801
RMSE	1.0984	0.9871	2.0112	0.2803
MAPE	28.09	42.56	15.97	7.01
MSPE	0.3678	0.2453	0.7033	0.1124
NRMSE	0.7098	0.4328	0.5362	0.0722
R^2	0.901	0.899	0.897	0.989
Theil	0.363	0.229	0.105	0.034

affects the failure rate of airborne equipment, and effectively mine the potential variable relationships between the influencing factors and the failure rate data, but also can use a small sample size to obtain reliable predictive values. It reflects the dynamics, advancement, and effectiveness of using this method for small samples and multiple correlation data analysis, which is suitable for the prediction of the failure rate of airborne equipment with a small sample size and the correlation of various influencing factors. The BP model is more suitable for occasions where the fault samples of airborne equipment change randomly and nonlinearly, and has good generalization ability and fault tolerance. The SVM model also has certain advantages in the collection of small samples of failure influencing factors and failure rates, as well as the failure rate prediction for nonlinear and high-dimensional pattern recognition problems. Based on the combined model, all the information of multiple single-phase models can be comprehensively utilized, and the characteristics of the combined model can be used to make the prediction error smaller than the single model error, improve the prediction accuracy, and reduce the risk of greater accuracy. Suitable for airborne equipment failures rate prediction, the combination method has strong applicability, and provides a certain idea for the failure rate prediction of airborne equipment.

7. In Conclusion

In order to provide good maintenance and maintenance decision-making, and establish the health management mechanism of airborne equipment, and improve the prediction accuracy of the failure rate of airborne equipment, 5 single models (multiple line regression (MLR) model, gray GM (1, N) model, partial least squares (PLS) model, artificial neural network (BP) model, and support vector machine (SVM) model) based on the optimal combination prediction model were used.

The accuracy, performance, and predictive ability of the combined model are studied, and compared with different individual models. The use of the combined model increases the R² index rate from 1.4% (compared to the BP model) to 10% (compared to the MLR model). It also reduced the MAPE index rate from 59.7% (compared to the BP model) to 88.4% (compared to the MLR model). In addition, the RMSE and MAE MSPE indexes are small compared with other models, and the Theil inequality coefficient is also small. The accuracy of prediction has been further improved, indicating that the combined model has the accuracy and ability to predict the failure rate of airborne equipment.

It provides a new method of failure rate prediction for airborne equipment. This method can be used to predict the remaining life, failure time, and other aspects of airborne equipment. In addition, the prediction method has certain reference significance and engineering application value in the field of complex equipment fault diagnosis and prediction. In the future, new applications of combination models will be further explored, and complex multi-combination models will be developed to achieve more accurate prediction results.

Data Availability

The data can be obtained from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Doctoral Startup Foundation of Shenyang Aerospace University (Grant no. 19YB30) and Scientific Research Funding Project of the Education Department of Liaoning Province (General Program) (Grant no. LJKZ0202).

References

- [1] W. Li, B. X. Zhou, D. N. Jiang, and X. J. Sun, "Fault prediction of power supply vehicle based on multi-state time series prediction learning," *Journal of Jilin University (Engineering and Technology Edition)*, vol. 50, no. 4, pp. 1532–1544, 2020.
- [2] H. Ai and D. H. Zhou, "Fault prediction approach for dynamic system," *Journal of Huazhong University of Science and Technology (Natural Science Edition)*, vol. 37, no. Sup. 1, pp. 222–225, 2009.
- [3] X. D. Hou, W. YP, and J. P. Yang, "Research summary of weapon electronic equipment fault Prediction based on state," *Systems Engineering and Electronics*, vol. 40, no. 2, pp. 360–367, 2018.
- [4] F. S. Al Badawi and M. AlMuhaini, "Reliability modelling and assessment of electric motor driven systems in hydrocarbon industries," *IET Electric Power Applications*, vol. 9, no. 9, pp. 605–611, 2015.
- [5] A. Verma, A. Narula, A. Katyal et al., "Failure rate prediction of equipment: can Weibull distribution be applied to automated hematology analyzers?" *Clinical Chemistry and Laboratory Medicine (CCLM)*, vol. 56, no. 12, pp. 2067–2071, 2018.
- [6] H. Louis and H. Barry, "Power converter failure rate prediction BY bayesian analyses," *Quality and Reliability Engineering International*, vol. 12, no. 5, pp. 365–370, 1996.
- [7] H. Max, J. S. Song, and S. F. Dong, "Model of spare parts failure rate based on linear regression," *Computer Simulation*, vol. 18, no. 11, pp. 6–8, 2003.
- [8] H. Motiee and S. Ghasemnejad, "Prediction of pipe failure rate in Tehran water distribution networks by applying regression models," *Water Supply*, vol. 19, no. 3, pp. 695–702, 2019.
- [9] A. Altinisik and U. Yildirim, "Failure prediction in electrical connector assembly: a case in automotive assembly process," *Assembly Automation*, vol. 40, 2020.
- [10] Y. Peng and M. Dong, "A prognosis method using age-dependent hidden semi-Markov model for equipment health prediction," *Mechanical Systems and Signal Processing*, vol. 25, no. 1, pp. 237–252, 2011.
- [11] L. I. Rui-Ying and R. Kang, "Research on failure rate forecasting method based on ARMA model," *Systems Engineering and Electronics*, vol. 30, no. 8, pp. 1588–1591, 2008.
- [12] S. Li, "Time series prediction of failure rate based on ARMA model," *Automation & Instrumentation*, vol. 21, no. 12, pp. 218–219, 2015.
- [13] Y. Yang, H. Zheng, and R. Zhang, "Prediction and analysis of aircraft failure rate based on SARIMA model," in *Proceedings of the 2017 2nd IEEE International Conference on Computational Intelligence and Applications (ICCIA)*, pp. 567–571, Beijing, China, September 2017.
- [14] M. Bevilacqua, M. Braglia, M. Frosolini, and R. Montanari, "Failure rate prediction with artificial neural networks," *Journal of Quality in Maintenance Engineering*, vol. 11, no. 3, pp. 279–294, 2005.
- [15] A. Z. Al-Garni, A. Jamal, A. M. Ahmad, A. M. Al-Garni, and M. Tozan, "Neural network-based failure rate prediction for De Havilland Dash-8 tires," *Engineering Applications of Artificial Intelligence*, vol. 19, no. 6, pp. 681–691, 2006.
- [16] R. Li and R. Kang, "Failure rate forecasting method based on neural networks," *Acta Aeronautica et Astronautica Sinica*, vol. 29, no. 02, pp. 357–363, 2008.
- [17] A. Z. Al-Garni and A. Jamal, "Artificial neural network application of modeling failure rate for Boeing 737 tires," *Quality and Reliability Engineering International*, vol. 27, no. 2, pp. 209–219, 2011.
- [18] A. Diryag, M. Mitić, and Z. Miljković, "Neural networks for prediction of robot failures," *Proceedings of the Institution of Mechanical Engineers, Part C: Journal of Mechanical Engineering Science*, vol. 228, no. 8, pp. 1444–1458, 2014.

- [19] F. Chen, "Safety evaluation method of hoisting machinery based on neural network," *Neural Computing and Applications*, vol. 33, no. 2, pp. 565–576, 2021.
- [20] X. D. Hou, J. P. Yang, Y. Zhang, and J. J. Xu, "Failure rate prediction of T/R component based on improved GM(1,1) model," *Journal of Naval University of Engineering*, vol. 30, no. 3, pp. 71–75, 2018.
- [21] M. Zhu, Z. R. Wang, H. Liang, Y. Q. Zhu, and X. Tong, "Prediction of the elevator failure rate based on the residual error correcting model of GM(1, 1)," *Journal of Safety and Environment*, vol. 17, no. 5, pp. 1701–1704, 2017.
- [22] Y. J. Shao, C. M. Ma, H. X. Pan, and Y. J. Liu, "Failure rate prediction research of weapon spare parts based on discrete GM(1,1)," *Journal of Gun Launch & Control*, vol. 18, no. 2, pp. 96–98, 2013.
- [23] H. Wang, L. Zhao, J. S. Liu, and X. Ji, "Prediction of electrical equipment failure rate based on improved drosophila optimization algorithm," in *Proceedings of the 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 1915–1921, Guilin, China, July 2017.
- [24] M. Kutylowska, "Application of MARSplines method for failure rate prediction," *Periodica Polytechnica Civil Engineering*, vol. 63, no. 1, pp. 87–92, 2019.
- [25] J. Yang, P. Liu, W. Hu, and W. Wu, "Method for predicting failure rate of power transmission equipment based on cloud theory," *High Voltage Engineering*, vol. 40, no. 8, pp. 2321–2327, 2014.
- [26] X. Wang, J. Wu, C. Liu, S. Wang, and W. Niu, "A hybrid model based on singular spectrum analysis and support vector machines regression for failure time series prediction," *Quality and Reliability Engineering International*, vol. 32, no. 8, pp. 2717–2738, 2016.
- [27] J. Cheng, Z. X. Xia, and L. G. Hu, "Prognostics for aeronautic equipments based on genetic neural network," *Journal of Air Force Engineering University(Natural Science Edition)*, vol. 12, no. 1, pp. 15–19, 2011.
- [28] Y. Feng, X.-D. Huang, R.-J. Hong, and J. Chen, "Online residual useful life prediction of large-size slewing bearings-A data fusion method," *Journal of Central South University*, vol. 24, no. 1, pp. 114–126, 2017.
- [29] Y. Yang, H. Yu, and Z. Sun, "Aircraft failure rate forecasting method based on Holt-Winters seasonal model," in *Proceedings of the 2017 IEEE 2nd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA)*, pp. 520–524, Chengdu, China, April 2017.
- [30] R. Q. Wang and C. Lu, "Fault rate prediction based on EMD and grey model," *Journal of Ordnance Equipment Engineering*, vol. 39, no. 7, pp. 127–130, 2018.
- [31] T. X. Xu, C. Lu, and H. Wang, "Failure rate prediction method based on relevance vector EMD and GMDH reconstruction," *Journal of Vibration, Measurement & Diagnosis*, vol. 38, no. 6, pp. 1275–1285, 2018.
- [32] W. Qiu, Q. Tang, Z. Teng, W. Yao, and J. Qiu, "Failure rate prediction of electrical meters based on weighted hierarchical Bayesian," *Measurement*, vol. 142, pp. 21–29, 2019.
- [33] J. Wang and H. Yin, "Failure rate prediction model of substation equipment based on weibull distribution and time series analysis," *IEEE Access*, vol. 7, pp. 85298–85309, 2019.
- [34] X. Song, C. Li, J. Wang, Y. Zhi, and Z. Wang, "A dynamic failure rate prediction method for chemical process system under insufficient historical data," *International Journal of Industrial Engineering*, vol. 26, no. 2, 2019.
- [35] J. Lei and Q. Gong, "Comprehensive prediction method for failure rate of transmission line based on multi-dimensional cloud model," *IET Generation, Transmission and Distribution*, vol. 13, no. 9, pp. 1672–1678, 2019.
- [36] C. Xiao, Z. Liu, T. Zhang, and L. Zhang, "on fault prediction for wind turbine pitch system using radar chart and support vector machine approach," *Energies*, vol. 12, no. 14, Article ID 2693, 2019.
- [37] Y. Lyu and Y. Jiang, "Examination on avionics system fault prediction technology based on ashy neural network and fuzzy recognition," *Journal of Intelligent & Fuzzy Systems*, vol. 38, no. 4, pp. 3939–3947, 2020.
- [38] Z. Zhang and W.-C. Hong, "Application of variational mode decomposition and chaotic grey wolf optimizer with support vector regression for forecasting electric loads," *Knowledge-Based Systems*, vol. 228, Article ID 107297, 2021.
- [39] M.-W. Li, Y.-T. Wang, J. Geng, and W.-C. Hong, "Chaos cloud quantum bat hybrid optimization algorithm," *Non-linear Dynamics*, vol. 103, no. 1, pp. 1167–1193, 2021.
- [40] Y. Dong, Z. Zhang, and W.-C. Hong, "A hybrid seasonal mechanism with a chaotic cuckoo search algorithm with a support vector regression model for electric load forecasting," *Energies*, vol. 11, no. 4, Article ID 1009, 2018.
- [41] Z. Zhou, J. W. Ling, R. L. Mei, and X. H. Tian, "Combined forecasting model of gas daily load based on weight distribution of ant colony algorithm," *Journal of Harbin Institute of Technology*, vol. 53, no. 6, pp. 177–183, 2021.
- [42] H. Liu, H.-Q. Tian, and Y.-F. Li, "Comparison of two new ARIMA-ANN and ARIMA-Kalman hybrid methods for wind speed prediction," *Applied Energy*, vol. 98, pp. 415–424, 2012.
- [43] J. Qi, G. Jiang, G. Li, Y. Sun, and B. Tao, "Surface EMG hand gesture recognition system based on PCA and GRNN," *Neural Computing and Applications*, vol. 32, no. 10, pp. 6343–6351, 2020.
- [44] J. B. Chen, Y. Z. Liu, Y. Chen, and K. Nie, "Fault diagnosis of airborne fuel pump based on EMD and SVM," *Journal of Beijing University of Aeronautics and Astronautics*, vol. 47, no. 8, pp. 1687–1696, 2021.