Predicting Spare Parts Inventory of Hydropower Stations and Substations Based on Combined Model

Zhenguo Ma, Bing Tang, Keqi Zhang, Yuming Huang, Danyi Cao, Jiaohong Luo, and Jianyong Zhang

1Changzhou Power Supply Branch, State Grid Jiangsu Electric Power Co., Ltd., Nanjing, Jiangsu 213022, China
2College of Science, Hohai University, Changzhou, Jiangsu 213022, China

Correspondence should be addressed to Jianyong Zhang; hohaizhangyi@hhu.edu.cn

Received 23 December 2021; Accepted 4 March 2022; Published 7 April 2022

Academic Editor: Jinyan Song

Copyright © 2022 Zhenguo Ma 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.

In this paper, a combined model is proposed to predict spare parts inventory in accordance with equipment characteristics and defect elimination records. Fourier series is employed to process the periodicity of the data, autoregressive moving average (ARMA) is used to deal with the linear autocorrelation of the data, and backpropagation (BP) neural network is used to settle the nonlinearity of the data. The prediction results, comparisons, and error analyses show that the combined model is accurate and meets the practical requirements. The combined model not only fully utilizes the information contained in the data but also provides a reasonable decision basis for the procurement of spare parts, making the inventory in a safe state and saving holding costs.

1. Introduction

As an important part of the power grid, the stable operation of hydroelectric power stations and substations is of great significance to the safety of the power grid [1, 2]. To ensure the stable operation of the system, the inventory of equipment spare parts needs to be in a safety status. When the quantity of inventory is insufficient, it cannot guarantee the system’s demand for spare parts and the timely elimination of defects, and when the quantity of inventory is too much, it will occupy too much storage space and capital.

Effective inventory forecasting not only ensures stable system operation but also reduces inventory costs and improves capital utilization of enterprises. There are a large number of research results on inventory forecasting, the forecasting objects are distributed in different industries such as supply chain, supply side, and demand side, and the forecasting methods are often focused on single-quantity model or improved single-quantity model [3].

There have been many research results for forecasting the spare parts demand for hydropower stations and substations. In these studies, scholars have gradually shifted from research on the management and optimization of inventories to that on inventory forecasting. Inventory management is receiving more and more attention as a condition for sustainable production. From the point of view of analytical techniques, researches on inventory management are focused on descriptive analytics, as well as on predictive and prescriptive analytics [4]. These applications and studies could be developed as an optimal resource allocation method that helps the inventory managers and engineers to optimize their inventory policy. For some power equipment, the long service life often leads to an excessive intermittent demand for spare parts, which is a challenge for inventory control, and to solve this problem, a regular review inventory control system is theoretically considered to be an optimal approach [5]. In [4, 5], the authors did not make full use of the information contained in the data. With the improvement of data mining techniques, many studies started to use machine learning [6], deep learning [7], and classification methods [4, 8] to make simple predictions about the inventory, and these predictions bring benefits to the actual inventory management such as improving the reliability of the system, decreasing
The failure rate of the equipment, and reducing the cost of operation and maintenance.

The application of artificial intelligence technology in spare parts inventory forecasting technology has significantly improved the accuracy of forecasting results. Ding [9] proposed an algorithmic model based on the modified BP neural network for grid material demand forecasting. The author simply improved the BP model to improve convergence. Similar studies based on artificial intelligence methods are machine learning [6, 10], deep learning [7], and so forth.

Qualitative forecasting models such as the expert meeting method [11] and the Delphi method [12] rely too much on experience and expert scoring, for short-term predictions of large data volumes are not valid, so their applicability and reliability are relatively poor. Quantitative models focus on mathematical models that use historical data or factor variables to forecast demand and apply certain mathematical methods to reveal the regular links between relevant variables, which have a high degree of objectivity. The conventional quantitative forecasting models are the autoregressive moving average (ARMA) model [13, 14], regression analysis [15, 16], and gray forecasting model [17].

Although fruitful results have been achieved based on these quantitative models, we have to consider the following three aspects. The first one is prediction accuracy. Any forecasting models have their limitations, which originate from the model itself or the data. For example, the ARMA model is suitable for time series with linear relationships and does not predict well for data with nonlinear relationships. While data often have multiple characteristics such as linear relationships, nonlinear relationships, and periodic characteristics at the same time, and if a single forecasting model is adopted, it will inevitably reduce the forecasting accuracy. An effective way to solve this kind of problem is to combine multiple models. Study [18] showed that when combining two single forecasting methods, its error was reduced by 7.2%, and when the combination of methods was increased to five, its error was reduced by 16.3%. Using an appropriate combination of forecasting models can overcome the limitations of a single model and improve the accuracy and diversify the forecasting risk as much as possible [19–21]. Researchers [19] used the recursive least square method combined with the 2RC model to obtain the best prediction of the state of charge of lithium batteries. However, in literature [20, 21], researchers considered fewer influencing factors. When the influencing factors are added, it will increase the difficulty of the study. The second one is that any prediction should be considered from a practical point of view. For solving practical problems, it is also necessary to start from the actual problem of the enterprise; for example, in power companies, the prediction value of spare parts must be slightly higher than the real value to prevent the failure of protection devices caused by insufficient spare parts when an unexpected event occurs, resulting in a power failure. The third point is that many factors affect inventory, such as internal and external factors. In this paper, we consider the inventory problem under 10 factors in a comprehensive manner, which is shown in Section 4.1. The proposed combined prediction model based on the above three aspects is the main innovation of this paper.

At present, power enterprises like DH Hydropower Station and JC Power Supply Company stockpile spare parts using the agreed inventory procurement [22]. This form of procurement relies on experience to speculate on the approximate demand rather than some theoretical model, such as the minimum life-cycle cost (LCC) [23] for spare parts for the next year. It has the disadvantage of insufficient spare parts inventory caused by both internal and external factors, which often leads to insufficient spare parts inventory and the need for temporary procurement. There is also over-purchase of certain spare parts, which causes inventory backlog and increases the cost of holding materials, thus taking up a lot of capital and storage space.

Based on the above reasons, this paper proposes a combined model to predict the inventory for the DH Hydropower Station excitation unit and substation relay protection equipment in JC Power Supply Company. To the best of our knowledge, no literature combines Fourier series, ARMA, and BP neural network to study inventory prediction, and there is no literature that uses this kind of mixed model to design a prediction scheme for inventory materials in Hydropower Station and substations. Moreover, the mixed model developed in this paper has a high prediction accuracy compared with the models developed by using the ARMA method alone and combining ARMA and BP neural network, and the model validity is good because the RMSE and MAE of the mixed model given in this paper are smallest. The first part is about the current status of research on the issue. The second part is about the related works, including the principle of Fourier series form of sequence, ARMA model, and BP neural network. In the third section, we put forward a combined model based on the Fourier series, ARMA, and BP neural network for inventory. The fourth part is the simulation of the real case and comparisons, and the last section is about the conclusions and a brief perspective of the future work.

2. Related Works

2.1. Fourier Series of Time Series. When analyzing defect elimination records, we find that the spare parts have a certain periodicity in usage; for example, during the rainy season, the usage of certain spare parts is higher than usual. To make the prediction more accurate, the periodic characteristics of the data need to be extracted using the Fourier series [24].

Let the sequence \(x_1, x_2, \ldots, x_m\), which can be considered as a point in an \(n\)-dimensional space coordinate system, form a set of bases for any \(n\)-dimensional orthogonal vectors for a given \(n\)-dimensional space. Thus, the sequence \(\{x_t, \ t = 1, 2, \ldots, n\}\) can be represented by a linear combination of orthogonal trigonometric functions as

\[
y(t) = \sum_{k=0}^{[m/2]} \left( a_k \cos \frac{2k\pi t}{m} + a_k \sin \frac{2k\pi t}{m} \right) + \epsilon_t. \tag{1}
\]
2.2. ARMA Model. ARMA model is a relatively mature forecasting model for studying linear time series, which requires time series data to be stationary to build the model, and has certain requirements on the magnitude of the data. Its model expression is

\[
\begin{align*}
    y_t &= \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 u_{t-1} + \cdots + \theta_q u_{t-q}, \\
    \phi_p &\neq 0, \theta_q \neq 0, \\
    E(u_t) &= 0, \text{Var}(u_t) = \sigma^2_n, E(u_t u_s) = 0, s \neq t, \\
    E(y_t u_t) &= 0, \forall s < t,
\end{align*}
\]

where \( y_t \) is the time series, \( p \) is the autoregressive order, \( q \) is the moving average order, and \( \phi_i (i = 1, \ldots, p) \), \( \theta_j (j = 1, \ldots, q) \) is the coefficient to be determined. \( u_t \) is the error.

The stationarity of the data is an important prerequisite for ARMA, and it can be verified by the Augmented Dickey-Fuller (ADF) test [25]. If the data are not stationary, stationary processing can be performed by using the \( n \) th-order difference method with the following equation:

\[
\begin{align*}
    \Delta^n y_t &= \sum_{i=0}^{n} (-1)^{n-i} C^n_i y_{t+i} \\
    &= \sum_{i=0}^{n} (-1)^{n-i} C^n_i y_{t+n-i}.
\end{align*}
\]

To obtain the best prediction model, the Akaike information criterion (AIC) [26–28] can be used to determine the parameters \( p \) and \( q \) of the model. Different AIC values can be obtained when fitting the data by selecting different \( p \) and \( q \).

\[
AIC = -2 \ln \sigma^2_i (p, q) + \frac{2(p + q)}{N},
\]

where \( \sigma^2_i (p, q) \) is the variance of residuals and ARMA (\( p, q \)) is considered the best prediction model when AIC has the smallest value.

2.3. BP Neural Network Model. The learning process of the BP neural network consists of two processes: forward propagation of signal and backward propagation of error. By continuously adjusting the value of network weights, the final output of the network is made as close as possible to the desired output for training purposes. Therefore, it can capture the nonlinear features and trends in it well when dealing with time series and thus is widely used for time series prediction.

It is composed of an input layer, an output layer, and one or more hidden layers, each of which consists of several neurons, and the individual neurons of adjacent layers are fully connected. Figure 1 shows a typical multilayer neural network.

The input vector is \( \mathbf{y} = (y_1, y_2, \ldots, y_m) \), the output vector is \( \mathbf{Y} = (Y_1, Y_2, \ldots, Y_n) \), and the input of each neuron in the \( l \) th hidden layer is \( \mathbf{h}_l^{(i)} = (h_1^{(i)}, h_2^{(i)}, \ldots, h_t^{(i)}) \), where \( t \) is the number of neurons in the \( l \) th layer. Let \( w_{ij}^{(l)} \) be the connection weight between the neuron from the \( j \) th neuron in the \( (l - 1) \) th layer and the \( i \) th neuron in the \( l \) th layer and \( b_i^{(l)} \) be the bias of the \( i \) th neuron in the \( l \) th layer. Then,

\[
\text{net}_i^{(l)} = \sum_{j=1}^{t-1} w_{ij}^{(l)} h_j^{(l-1)} + b_i^{(l)},
\]

where \( h_i^{(l)} = f (\text{net}_i^{(l)}) \), where \( \text{net}_i^{(l)} \) is the output of the \( i \) th neuron in the \( l \) th layer and \( f (\cdot) \) is the neuronal activation function. Usually a nonlinear activation function is used in multilayer neural networks.

To ensure the prediction accuracy of the BP neural network, the loss function is defined as follows:

\[
I f (k) = \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left[ x_n (i) - x_{pk} (i) \right]^2},
\]

where \( x_n \) denotes the predicted value of the BP model after the \( k \) th round of training. The function \( If (k) \) means the root mean square error of the true value \( x_n \) and \( x_{pk} \). A smaller
If \( f(k) \) indicates adequate training and more accurate prediction.

### 3. Combined Model

For the time series \( \{x_t\} \), in this paper, it not only reflects periodicity but also reflects certain linear autocorrelation and nonlinear nature of the residuals. To make the prediction accurate, in this section, we construct a combined model to forecast the data. The idea of the model is shown in Figure 2.

#### 3.1. Data Processing

For a given time series data \( \{x_t\} \), firstly, it can be transformed into a time series \( \{w_t\} \) with periodic properties by using the Fourier series of the time series in equation (1) and then verify the smoothness of the data to predict the autocorrelation part of the series. The data processing in this study is shown in Algorithm 1.

#### 3.2. Prediction of Linear Autocorrelation

For the smooth time series \( \{w_t\} \) with periodic characteristics, the ARMA model is used to get \( \{\hat{w}_t\} \), and \( \{\hat{w}_t\} \) is the predicted result of the linear autocorrelation of the time series. The residual \( \{e_t\} \) is also calculated, which is the input of the prediction of the nonlinear residual. The linear autocorrelation prediction process is shown in Algorithm 2.

#### 3.3. Prediction of Nonlinear Autocorrelation

In this section, we are going to construct BP neural network model to forecast the residuals, and the main steps are shown in Algorithm 3.

#### 3.4. Combined Model

Taking together Algorithms 1–3, we can generate a combined model to computer the predicted results \( \{\hat{X}_t\} \) for the time series \( \{x_t\} \). The details are listed in Algorithm 4.

Note: \( f(\) the case of first-order difference.

### 4. Experiment, Results, and Comparisons

#### 4.1. Experimental Data

The spare parts in this paper are the liquid crystal display (LCD), the central processing unit (CPU), and DC220V Power Supply (DCPS), which are installed in the excitation system of hydroelectric generating units and substation relay protection equipment. They have the following features:

- A wide variety and distribution and high value and integration
- It is difficult to obtain monitoring data, and the operation and maintenance (O&M) process requires a lot of manpower. It has a long procurement cycle
- Service life is influenced by environmental factors and human factors

The data from the defect elimination records stored in DH Hydropower Station and JC Power Supply Company contain 10 field variables: Component Name, Installation Date, Manufacturer, Damage Date, Design Life, Temperature, Humidity, Installation Location, Interval, and O&M (Operation and Maintenance) Shift, where Design Life is a...
**Algorithm 1: Data processing.**

**Input:** \( \{x_t\} \)

Step 1. Get \( \{y_t\} \) from \( \{x_t\} \) by equation (1)

Step 2. ADF test
- If \( \{y_t\} \) is stationary, then
  - \( w_t = y_t \)
- Else
  - \( w_t = y_t - y_{t-1} \)

**Output:** \( \{w_t\} \)

**Algorithm 2: Prediction of a linear autocorrelation.**

**Input:** \( \{w_t\} \)

Step 1. Calculate \( p \) and \( q \) for ARMA \( (p, q) \)
  - For \( p = 0 \) to \( M \) \( M \) is a positive integer
  - For \( q = 0 \) to \( K \) \( K \) is a positive integer
  - \( \text{AIC}(p, q) = -2 \ln \sigma_t^2(p, q) + 2(p, q)/N \)

Step 2. Forecast the time series \( \{w_t\} \) using ARMA \( (p, q) \).

Step 3. Get \( \{\bar{w}_t\} \) \% prediction values

Step 4. \( e_t = w_t - \bar{w}_t \) \% residuals

**Output:** \( \{\bar{w}_t\} \) and \( \{e_t\} \)

**Algorithm 3: Prediction of the nonlinear residual.**

**Input:** \( \{e_t\} \)

Step 1. Configure BP neural network

Step 2. Construct the training set \( \{(i, e_i), \ i = 1, 2, \ldots, n\} \)

Step 3. Train BP model
  - While \( |f_k(i + 1) - f_k(i)| > \epsilon \) or \( f_k(\cdot) > \alpha \)
  - Train BP model
  - Determine the BP model

Step 4. Get predicted values of the residuals \( \{\hat{e}_t\} \)

**Output:** \( \{\hat{e}_t\} \)

**Algorithm 4: Combined model.**

**Input:** \( \{x_t\} \)

Step 1. Get \( \{w_t\} \) by Algorithm 1

Step 2. Obtain \( \{\bar{w}_t\} \) and \( \{e_t\} \) via Algorithm 2

Step 3. Calculate \( \{\hat{e}_t\} \) by Algorithm 3

Step 4. Compute \( \{\hat{X}_t\} \):
- If \( \{x_t\} \) is smoothing, then
  - \( \hat{X}_t = \hat{e}_t + \bar{w}_t \)
- Else
  - \( \hat{X}_t = \hat{e}_t + \bar{w}_t + \bar{X}_{t-1}^{-1} \)

**Output:** \( \{\hat{X}_t\} \)
numeric variable, Installation Date and Damage Date are temporal variables, and the others are textual variables.

The overall trend in the annual usage of a certain relay protection equipment is smooth, and we can consider the data as a time series. Each year, in the case of high temperature and high humidity or low temperature and high humidity, the use of spare parts will increase; that is, the data will reflect a certain periodicity. At the same time, the service life of the equipment, that is, Damage Date minus Installation Date, will be affected by the working conditions. Therefore, the data will reflect a certain nonlinear characteristic.

4.2. Inventory Prediction Based on ARMA Only. In this section, only the ARMA model is used to forecast three kinds of components. We can find that all data are stationary after
testing by the ADF method. The parameters $p$ and $q$ are determined according to AIC in (5), the details are listed in Table 1, and the forecast results are shown in Figure 3.

From Figure 3, we can conclude that the predicted values match too poorly with the true values. This indicates that the data have a weak linear autocorrelation, especially for the data of CPU and DCPS. But the ARMA model can only deal with data with linear relationships by nature; therefore, it can be concluded that the prediction based on this model only is invalid.

4.3. Inventory Prediction Based on ARMA and BP. To deal with the nonlinear relationships in the data, a BP neural network model is added to the ARMA model. The settings of the parameters in the ARMA model are shown in Table 1.

In the BP neural network, the input and output layers have one node each, and three hidden layers include a fully connected layer containing 10 nodes, a discarded layer with a discard rate of 20%, and a fully connected layer with 15 nodes, respectively. Training termination conditions $\varepsilon = 0.001, \alpha = 0.05$. To avoid the gradient explosion and gradient disappearance problems in BP neural networks, simplify the computation process and improve the computation speed, and the activation function that we adopt in this paper is the Rectified Linear Unit (ReLU) [29, 30]:

$$f(x) = \max(0, x).$$  \hspace{1cm} (8)

The predicted values for LCD, CPU, and DCPS based on the ARMA model and the BP neural network are shown in Figure 4.

![Graphs showing prediction values for LCD, CPU, and DCPS based on ARMA and BP models.](image-url)
Comparing Figures 3 and 4, we can see that the degree of agreement of the predicted values based on the ARMA model and BP neural network is higher than the ARMA model’s, which indicates that there is indeed a nonlinear relationship in the data.

At the same time, it can be seen from Figure 4 that the results are not bad for LCD, but for CPU, there are more cases where the predicted values are lower than the true values, and for DCPS, throughout the forecast curve, there are several cases where peaks and valleys occur at the same time point, resulting to a large difference in predictions, which is due to the fact that the ARMA and BP neural network model have not fully explored the information of the data.

### 4.4. Inventory Prediction Based on the Combined Model

According to the experience of operation and maintenance, in periods of high temperature and humidity, or low temperature and humidity, the demand for spare parts is higher, and the data reflect a certain periodicity. Figure 5 is the description of the loss function of the BP neural network versus the training time. The smaller the loss function values are, the more accurate the neural network model is and the better the prediction result is. In Figure 5, we conclude that the data after the Fourier series extraction of features will make the BP neural network have smaller loss function values.

In this section, we are going to implement a combined model to forecast the number of spare parts. We set $n = 48$ for the Fourier series in Algorithm 1. In Algorithm 2, we recalculate the AIC of ARMA, the values of $p$ and $q$ are obtained and shown in Table 2, and in Algorithm 3, we adopt the same BP neural network used in Section 4.3. The prediction results are shown in Figure 6.

As can be seen from Figure 6, the predicted and true values of the three devices are in good agreement with the trend of change. Moreover, the predicted values are all slightly greater than or equal to the true values. Such prediction results are in line with the safety requirements of hydropower stations and substations for spare parts, and this prediction model is more advantageous for the stability of the system.

Through the comparison of predicted results, the combined model we constructed in Section 3 is the best one for this type of predicting problem.

### 4.5. Error Analyses

To further illustrate the validity of the combined model, we conduct error analysis for the three prediction models in this paper. The corresponding root means square error (RMSE) and means absolute error (MAE) under the three methods are discussed, which are calculated as follows:

![Figure 5: The relationship between loss function and training times.](image-url)

**Table 2: Parameters of the ARMA model.**

<table>
<thead>
<tr>
<th>Equipment</th>
<th>AIC</th>
<th>$p$</th>
<th>$q$</th>
</tr>
</thead>
<tbody>
<tr>
<td>LCD</td>
<td>327.925</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>CPU</td>
<td>395.599</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>DCPS</td>
<td>400.587</td>
<td>2</td>
<td>1</td>
</tr>
</tbody>
</table>
RMSE and MAE of three prediction methods.

<table>
<thead>
<tr>
<th>Equipment</th>
<th>RMSE</th>
<th>MAE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.1365</td>
<td>0.8416</td>
</tr>
<tr>
<td>LCD</td>
<td>ARMA</td>
<td>ARMA + BP</td>
</tr>
<tr>
<td>CPU</td>
<td>1.3693</td>
<td>0.9128</td>
</tr>
<tr>
<td>DCPS</td>
<td>1.7440</td>
<td>1.5679</td>
</tr>
</tbody>
</table>

where $x_a$ and $x_p$ denote the predicted and true values, respectively. If the RMSE and MAE of a certain algorithm are smaller, it means that the algorithm is more effective. The RMSEs and MAEs of the three prediction models are shown in Table 3.

From Table 3, it is known that the RMSE and MAE based on the combined model are both minimal, which further illustrates the effectiveness of the combined model.
5. Conclusions and Prospect

According to the prediction results and their comparison, the data of defect elimination records have certain periodic characteristics, linear autocorrelation, and nonlinear relationship, so extraction of the periodic characteristics via the Fourier series is a key step. The ARMA model does its work in processing the linear autocorrelation, and this is the main part of the forecasting results. The BP neural network model fully exploits the nonlinear relationship of the data, and its purpose is to improve the prediction accuracy. Three different methods take full advantage of their respective strengths and make the combined model perform well.

The combined model was used in inventory forecasting for three different devices at the same time, and all of them demonstrated good forecasting results, which indicate that this model has some strong portability and stability. The analysis of the forecasting results also shows that the model has some advantages over the single model or two models in forecasting.

The combined model used in this paper embodies three practical applications: the first is the accurate calculation, high portability, and stability; the second is to improve inventory management capability and save inventory cost and storage space; the third is the ideas in this paper that can be generalized to another field, such as the primary equipment of power system.

For electric power enterprises, the problem of predicting inventory of equipment without monitoring devices has a high practical value, and subsequent research efforts can add more influencing factors, such as product batches, idle time, and extreme weather conditions, but data on these influencing factors are currently unavailable. Under the many factors, we can reconstruct the prediction model and choose the remaining useful life (RUL) [31, 32] of the equipment as the main influence variable while considering the life-cycle costing (LCC) [23, 33] of the equipment to achieve a more accurate prediction of the equipment inventory, which is more conducive to the operation and maintenance of relay protection equipment, as well as reducing inventory expenditures.

Data Availability

All data included in this study are available upon request with the corresponding author.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

The authors are very grateful to Professor Binwu Zhang from the Institute of Intelligent Optimization and Control of Hohai University for their enthusiastic guidance and constructive suggestions. This work was supported by the Science and Technology Project of State Grid Corporation Limited (SGJSCZ00KJJS2100747) and in part by the National Nature Science Foundation of China (NSFC) under Grant no. 52077058.

References


