Check for updates

Standard Cost Forecasting Model for Power Grid Production and Operation Under Multidimensional Lean Management Model

Yanqin Wang(✉), Yong Wang, Ning Xu, Shunli Wang, and Xian Zhao

State Grid Hebei Electric Power Co., Ltd. Economic and Technological Research Institute, Shijiazhuang 050000, Hebei, China
tammychen2022@163.com

Abstract. The current conventional standard cost forecasting model for power grid production and operation is mainly based on the analysis of historical data of power grid operation by combining with big data technology, which leads to poor forecasting effect due to the lack of analysis of cost drivers. In this regard, we propose a standard cost forecasting model for power grid production and operation under the multidimensional lean management model. The interval number is used as the scale of the operating cost of the grid enterprise, and the interval number is processed to remove uncertainty, and the drivers of the grid operating cost are analyzed and merged, and the cost prediction model is constructed by combining with BP neural algorithm. In the experiments, the proposed model is verified for prediction accuracy. The analysis of the experimental results shows that when the proposed model is used to predict the grid operation cost, the root mean square error value of the model is low and has a high prediction accuracy.

Keywords: time series · kinetic merging · neural networks · multidimensional lean management · grid operations · cost forecasting

1 Introduction

Cost management is an important part of keeping the grid enterprise running normally and gaining more revenue. Cost management involves forecasting the operating cost of the grid enterprise, operating cost planning [1]. Cost management involves forecasting, planning, and analyzing, accounting, and checking the operating costs of a grid enterprise. The cost of the grid enterprise may be a stable data, more there is a certain degree of uncertainty, especially in the grid enterprise in the grid enterprise in the process of historical data collection, it is impossible to precise to a specific data, strictly speaking, the grid enterprise operating costs should be between a certain area is more reasonable, that is, the cost data should be an uncertainty interval data. Project operating cost forecast refers to the project operation process of all the activities generated by the cost of the forecast, the project in the pre-construction feasibility study process for the project operating period will occur during the cost of the forecast, including all costs associated

© The Author(s) 2023
N. Akhtar et al. (Eds.): PMIS 2023, AHIS 8, pp. 830–838, 2023.
https://doi.org/10.2991/978-94-6463-200-2_86

with operating revenue, can be based on previous years operating cost data and will occur in the operation of the data on operating cost forecast, including direct materials, direct labor This includes direct materials, direct labor, production costs, administrative costs, financial costs, etc. [2]. A common method of project operating cost forecasting is the elemental forecasting method, i.e., the operating expense method, which uses a simple linear relationship to forecast each cost and then sums to calculate the total operating cost. The advantage of this method is that it is simple and easy to operate. The disadvantage of this method is that using linear relationship to forecast operating costs is only a rough estimate of operating costs, ignoring the characteristics of non-linear costs, and there will be a large deviation from the actual operating costs. The job costing method is a costing method that takes jobs as the accounting object, uses job cost drivers to calculate the amount of jobs, and allocates indirect costs to multiple products through job cost driver allocation rates. The job costing method is generally applicable to the situation where the proportion of indirect costs to total costs is high, the activities of production and operation of grid enterprises are more operations, the production process is more complex, and the types of products produced are more. For grid enterprises with high degree of computerized accounting, products requiring different levels of technical services, and large scale of the company, the advantages of the job costing method are that it can make product costs more accurate, facilitate cost control, and provide support for strategic management of grid enterprises. The characteristics of the job costing method are that it can not only allocate and calculate costs, but also analyze the flow of resources according to the cause-effect relationship. In this paper, we propose a standard cost prediction model for power grid production and operation under the multidimensional lean management model, analyze the power supply cost of power grid from the perspective of big data, and dig deeper and analyze the production and operation cost data of power grid [3].

2 Selection of Operational Cost Scales and Information Processing in a Multidimensional Lean Management Model for Power Grids

Multidimensional lean management has a certain demand for the original value data system of grid power grid enterprises, and the management requirements are gradually increasing. However, the current business data and value data lack a clear logical relationship, the data can not be fully integrated, can not systematically carry out in-depth insight analysis of data to achieve the purpose of financial-driven business improvement [4]. Therefore, it is necessary to further break through the business and finance data barriers through digital construction and new technologies, and clarify the desk management system. In the multidimensional lean management model, the power grid enterprise operation process involves many influencing factors, in the process of cost management of power grid enterprises is crucial to the effective prediction of the cost of power grid enterprises, in the process of cost forecasting, after years of research overall cost forecasting is generally divided into three categories, one is the direct use of market research information and personal experience of decision makers to the future product operation The second category is through the impact of the grid enterprise operating costs of factors to analyze to predict the grid enterprise costs of correlation analysis method; the third category is the grid enterprise operating costs of the past data for statistical composition of the historical time series, and the use of the historical data to build a mathematical model with extrapolation to the grid enterprise operating costs of time series forecasting method. The time series forecasting method. The above three types of forecasting methods, overall time series forecasting methods are more scientific and reasonable, more data-based, more objective, and therefore time series models are more often used in cost forecasting [5].

The use of time series forecasting methods for power grid enterprise cost forecasting process, mainly to the power grid enterprise operating costs of historical data collection, these historical data to establish the corresponding mathematical model, used to forecast the future cost of the power grid enterprise [6]. In the process of data collection on the operating costs of power grid enterprises, the operating costs of power grid enterprises in the statistical process involves a variety of influencing factors, especially in the process of accounting for operating costs, there are certain external accounts and other aspects of costs, in the process of accounting for costs can not be a definite data, but should be a data between a certain interval, so in the process of data collection on the operating costs of power grid enterprises In the process of collecting historical data, it is more scientific and reasonable to choose the interval as the scale of the operating cost of the grid enterprise, which can better reflect the objective facts of the operation of the grid enterprise. In this paper, we consider the interval as the scale of the operating cost of grid enterprises to construct a cost prediction model based on the uncertain environment [7].

Since modeling and calculating interval historical data are complex and computationally intensive, it is necessary to de-intervalize interval data in the process of modeling interval historical data, i.e., to convert interval numbers into model-friendly real numbers. The uncertainty removal function can be used to convert interval numbers into model-friendly real numbers [8]. In the historical data statistics of the operating cost of the power grid enterprises, the historical cost data collected according to the actual situation are given in the form of interval numbers, assuming that the time series of interval-type historical cost of the power grid enterprises obtained by statistics are:

$$X = \{x(1), x(2), ..., x(n)\} \tag{1}$$

The collected interval-type historical cost values are processed using a de-uncertainty function to obtain the de-uncertainty time series expressions as shown below.

$$Y = \{y(1), y(2), ..., y(n)\} \tag{2}$$

Through the above steps, the interval number is selected as the scale of grid enterprise operation cost, which provides data support for the subsequent construction of the standard cost prediction model of grid operation.

3 Consolidation of Standard Cost Drivers for Grid Production and Operation Operations

In order to construct a standard cost prediction model for grid production and operation operations, the factors that affect costs need to be analyzed. Studies have shown that combining fully correlated cost drivers does not have a significant impact on the final operating cost prediction results and can also improve the prediction accuracy of the model. In order to reduce the complexity of the cost allocation process, highly correlated job cost drivers can be merged [9]. By comparing the advantages and disadvantages of job cost drivers selection, it is clear that the cost drivers determined by using cluster analysis and principal component analysis are more accurate for the calculation results of product cost, and according to the comparison of related studies, the separation sum of squares, the distance between samples must be Euclidean distance, and in practical application, the separation sum of squares classification is better and more widely used. Therefore, in this paper, we first use correlation coefficient clustering analysis to get the correlation coefficient matrix first, then get the clustering dendrogram according to the correlation coefficient matrix, and then conduct principal component analysis to get the number of cost drivers, and combine with the clustering dendrogram to determine the final cost drivers [10].

Cluster analysis is a statistical analysis technique that classifies samples into relatively homogeneous classes. In current practical applications, systematic clustering and K-means clustering are the two most widely used classes. It is to first treat each sample as a class, and then define the distance between classes systematic clustering, also known as hierarchical clustering, merges the smallest pair of distance between classes to form a new class, and then calculates the distance between the newly generated classes and other classes, merges the two closest classes, and repeats the merging, each time merging will Each time the merger is repeated, one class is reduced until finally all related classes are merged into one class [11]. K-mean clustering, on the other hand, views the data as points on a K-dimensional space, and uses distance as the criterion for cluster analysis, dividing the samples into specified K classes, i.e., the number of categories needs to be developed in advance. The number of categories cannot be determined in advance, so systematic clustering is chosen. There are various methods for calculating the distance between classes, and the common distances are absolute value distance, Marxian distance, Euclidean distance, Chebyshev distance, Lang's distance, and Gapkov distance. In practical application, the difference sum of squares classification is better and most widely used, all use the difference sum of squares to calculate the distance between classes, that is, the Euclidean distance, also called Ward's method [12].

Firstly, suppose there are n cost drivers in the process of grid production and operation $N_j(j = 1, 2, ..., n)$, and these drivers are divided into m classes $M_i(i = 1, 2, ..., m)$, where m is determined according to the accuracy of cost calculation, and m clustering centers are the final clustering results. The P-dimensional vector of the kth cost driver in the class is denoted by X_{kt}, the number of cost drivers is denoted by n_t, and the center of gravity of the class is denoted by X_t. Then the sample sum of squares of deviations in the class M_t is shown below.

$$S_t = \sum_{k=1}^{n_t} (X_{kt} - \overline{X_t})^T (X_{kt} - \overline{X_t}) \tag{3}$$

According to the above formula, the optimal clustering scheme can be selected as shown below.

$$S = \min S = \sum_{t=1}^{m} S_t \tag{4}$$

Principal component analysis is a further mining of information from the perspective of variance contribution. By combining matrix theory, it is concluded that the number of cost drivers should not be smaller than the rank in the coefficient matrix, so that there will be better accuracy, and it is experimentally proved that the rank in the system matrix is equal to the number of the cumulative variance contribution of 100% in the principal component analysis. Therefore, principal component analysis is feasible for determining the number of cost drivers, and the principal component contribution method can be used to evaluate the selection of cost drivers. In this paper, we use spss software to directly obtain the results of cost drivers calculation, and the specific steps are as follows.

Firstly, we construct n categories, each of which contains only one variable, then calculate the similarity coefficients between variables to form a correlation coefficient matrix, draw a clustering dendrogram based on the similarity coefficient matrix, and then use principal component analysis to determine the final number of cost drivers. Through the above steps, the analysis and consolidation of cost drivers for grid production and operation can be completed, which can help to improve the accuracy of the cost prediction model.

4 Power Grid Production and Operation Standard Cost Forecasting Model Construction

The relationship between the operating cost of power grid production and operation and its influencing factors is complex and difficult to represent by a linear function, so it needs to be expressed by a complex nonlinear function, and i.e., artificial neural network can approximate a nonlinear curve with arbitrary accuracy, so use i.e., artificial neural network to establish the nonlinear relationship between cost and influencing factors, and then make predictions of cost [13].

Neural network a large number of input samples usually differ in meaning, units, nature and other criteria, normalization can make the basic unit of measurement unified, and there are differences in the range of values of each field, when the difference between the range of values is large, it will have a different degree of impact on the network, larger values will have a greater impact on network learning, which will lead to a slower network learning speed or even can not converge, that is, into the s type function's saturation zone, so in order to reduce the difference between the range of values and speed up the network convergence, the input sample data is preprocessed, that is, the sample data is normalized so that the sample data is normalized between [0,1], so that the sample is convenient for the unified analysis of different categories of sample data, and the preprocessed data will not adversely affect the network learning, but when some sample data is normalized to 0 or 1, when 0 and 1 happen to be the extreme points of the S-type function in the BP artificial neural network, if you want to output suitable results, you must set larger weights, and the increase of the weights will increase the learning time and slow down the training speed, so the sample data will be normalized in the interval less than [0,1], for example, between [0.05–0.95], when the normalization formula is shown below.

$$Y_i = 0.9 \times \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} + 0.05 \tag{5}$$

where, Y_i represents the normalized result, x_i represents any value in the sample data, $x_{\max}$ represents the maximum value in the sample data, $x_{\min}$ represents the minimum value in the sample data. The weights and thresholds in the BP artificial neural network are random, and the selection of the initial weights will affect the convergence speed of the network and the convergence of the output results, but usually a value close to zero is selected first, usually between [0, 1], because The initial value if too large will make the network training prematurely into the saturation zone, the convergence speed becomes very slow, the neuron output results convergence is also relatively weak, specifically explained as follows: the weight and threshold will affect the input value of the s-type function, when the weight and threshold is small, the s-type function input value is closer to 0. This region function curve is steeper, the slope is larger, that is, the regulation performance is stronger, the convergence speed of the function is stronger, the output results convergence is also stronger. The convergence of the output results is also better, on the contrary, if the beginning of the weights and thresholds are larger, the s-type function input value in the direction of positive infinity and negative infinity, this region s curve is relatively flat, the slope is smaller, that is, the regulatory performance is poor, the convergence speed of the function is very slow, the output results tend to -1 and 1, the convergence is relatively poor. In addition to the weight algorithms are gradient descent method, gradient descent method with momentum, adaptive lr gradient descent method, adaptive lr momentum gradient descent method, L-M optimization algorithm, after the effect of the following program run, this paper selects adaptive lr momentum gradient descent method, i.e. traingdx optimization algorithm. Then, in the forward propagation process of the prediction model, the input layer input sample x_i is adjusted by the weighted average and threshold to obtain the following output value [14].

$$h_j(x_i) = \sum w_{ij}x_i + \vartheta_j \tag{6}$$

where, w_{ij} represents the weighted average of the samples and ϑ_j represents the threshold value. The above results are used as the input values of the implicit layer, and the following output values, i.e., the standard cost prediction values, are obtained by processing the neuron summary function of the implicit layer.

$$Z_j(x_i) = f[h_j(x_i)] \tag{7}$$

The above steps will lead to the construction of a standard cost forecasting model for grid operations. Combining the contents of this section with the aforementioned grid operations cost scale selection and information processing, and the consolidation of standard cost drivers for operations, the standard cost forecasting model for grid production and operations under the multidimensional lean management model has been designed [15].

5 Testing and Analysis

5.1 Test Preparation

In order to prove that the prediction effect of the standard cost forecasting model for grid production and operation under the multidimensional lean management model proposed in this paper is better than the conventional standard cost forecasting model for grid production and operation, an experimental session is constructed to verify the actual prediction effect of the model after the theoretical part is designed. In order to improve the reliability of the experimental results, two conventional standard cost forecasting models for grid production and operation are selected as the objects of comparison, and the experimental results of the three methods are compared to prove the effectiveness of the method in this paper. The conventional methods selected for this experiment are the standard cost prediction model of grid production and operation based on gray theory and the standard cost prediction model of grid production and operation based on influence factor analysis.

Since the method in this paper selects the BP neural network method as the model construction method, the parameters in the method need to be set. The model proposed in this paper uses the tangent function of the S-type function as the excitation function of the implicit layer, the linear function Purelin as the excitation function of the output layer, the momentum term coefficient is 0, the learning factor is 1, and the number of neurons in the input and output layers are set to 3. After the training of the models is completed, the root mean square error of the three prediction models is calculated to test the prediction accuracy of the models.

5.2 Analysis of Test Results

The evaluation index selected for this experiment is the prediction accuracy of the model, and the specific comparison index is the root mean square error of the model, the lower the value, the higher the prediction accuracy of the model is represented, and the specific calculation formula is shown below.

$$\sigma = \sqrt{1/n \sum_{i=1}^{n} (S_i - \overline{S_i})^2} \tag{8}$$

where, σ represents the root mean square error, S_i represents the actual predicted value, $\overline{S_i}$ represents the deviation value, and n represents the number of test samples, the specific experimental results are shown below.

Analysis of Fig. 1 shows that there are also differences in the prediction accuracy of different models when predicting the standard cost of grid production and operation operations. The numerical comparison shows that the root mean square error of the proposed multidimensional lean management model is lower than 4.0%, while the root mean square error of the two conventional cost forecasting models is above 7.0%. Therefore, it can be proved that the output value predicted by the cost prediction model in this paper is a good fit to the actual value curve and has a better prediction accuracy.

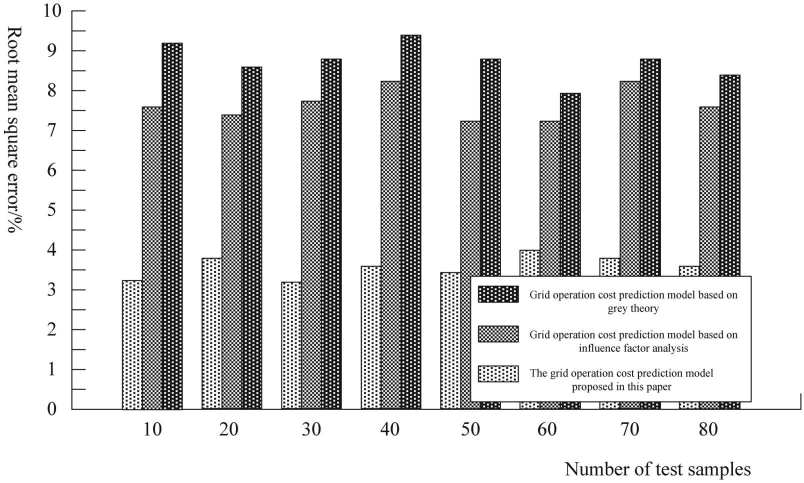

Fig. 1. Comparison results of root mean square error

6 Conclusion

This paper proposes a prediction model for the standard cost of grid production and operation in a multidimensional lean management model. By making a historical time series of the operational cost data of grid enterprises, and using the historical data to build a mathematical model with extrapolation, the training effect of the model can be effectively improved, based on which a cost prediction model is constructed by combining BP neural network. Compared with the traditional method, the method proposed in this paper can apportion the cost to each operation unit, which is more helpful to realize the fine control of operation cost.

References

1. Envelope F, Adm A. Multiple Linear Regression Model for Improved Project Cost Forecasting [J]. Procedia Computer Science, 2022, 196:808-815.
2. Paulescu M, Paulescu E, Badescu V. Nowcasting solar irradiance for effective solar power plants operation and smart grid management - ScienceDirect[J]. Predictive Modelling for Energy Management and Power Systems Engineering, 2021:249–270.
3. Dhanalakshmi J , Ayyanathan N . State Wise Modeling and Descriptive Analytics with Data Visualization of Indian Southern Power Grid Operation[J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(10):5931-5944.
4. Yu J , Ji H , Song Q , et al. Design and implementation of business access control in new generation power grid dispatching and control system [J]. Procedia Computer Science, 2021, 183(22):761-767.
5. Isakov I , Todorovic I . Power production strategies for two-stage PV systems during grid faults[J]. Solar Energy, 2021, 221:30-45.
6. Kumar A, Khan I . Harmonics Analysis and Enhancement of Power Quality in Hybrid Photovoltaic and Wind power System for Linear and Nonlinear Load using 3 Levels Inverter [J]. International Journal of Recent Technology and Engineering, 2021, 10(1):296-307.

7. Taheri S S , Seyedshenava S , Mohadesi V , et al. Improving Operation Indices of a Micro-grid by Battery Energy Storage Using Multi Objective Cuckoo Search Algorithm[J]. International Journal on Electrical Engineering and Informatics, 2021, 13(1):132-151.
8. Manzoor U , Baig S A , Sami A , et al. Evaluating the Impact of Lean Practices and Supply Chain Management Practices on Firm Performance-An Empirical Evidence from Manufacturing Industries[J]. Asia Proceedings of Social Sciences, 2021, 7(1):57-59.
9. Kennedy L , Francis F H , Alpini F G . Macrophage-Specific SCAP Promotes Liver and Adipose Tissue Damage in a Lean NAFLD Model: Lean, Mean, Proinflammatory Machine[J]. Cellular and Molecular Gastroenterology and Hepatology, 2022, 14(1):236-238.
10. Jiang Y , Yu H , Jiang J . Optimization of Multidimensional Clinical Information System for Schizophrenia[J]. Complexity, 2021, 2021(3): 1-10.
11. Zhang J , Zhu W , Wu X , et al. Traffic State Detection Based on Multidimensional Data Fusion System of Internet of Things [J]. Wireless Communications and Mobile Computing, 2021, 2021(1):1-12.
12. Ada M . Decoupling in lean management accounting practices [J]. Zeszyty Teoretyczne Rachunkowości, 2021, 45(1):75-92.
13. Fu H. Optimization Study of Multidimensional Big Data Matrix Model in Enterprise Performance Evaluation System[J]. Wireless Communications and Mobile Computing, 2021, 2021(2): 1-12.
14. Teixeira J , Bernardi F A , Rui P , et al. Proposal for a health information management model based on Lean thinking [J]. Procedia Computer Science, 2021, 181(5):1097-1104.
15. Zhou R , Wang W . Simulation of Network Traffic Prediction Model Based on Deep Neural Network [J]. Computer Simulation,2021(5) 475-479.

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.