Construction and Application of Data Mining Model of Unit Scale Cost of Electric Grid Project

Peng Rong1*, Guoming Liu1, Peng An2, Xiang Qi1, Qing Wang1, Caiyang Yue2, Wei Zhang1, Han Zhang1, Yong Dai1

1State Grid Shandong Electric Power Company Taian Power Supply Company, Shandong Province, Taian; 271000, China
2State Grid Shandong Electric Power Company, Shandong Province, Jinan; 250000, China
*Corresponding author’s e-mail: 14259069@qq.com

Abstract. With the further promotion of the power system reform, the transmission and distribution price reform is in-depth, from "building mechanism" to "strong supervision", and the effectiveness of investment has become the focus of supervision by all parties, so that grid companies must pay more attention to the efficiency and efficiency of investment. This paper mainly studies how to build a project cost analysis and data mining model based on the massive feasibility study estimate, preliminary design estimate and final account data of power grid infrastructure projects, mining the value laws and characteristics of power grid infrastructure projects from multiple scenes and perspectives, supporting the investment amount estimate in the planning stage, the investment estimate review in the feasibility study stage, and the investment control target prediction in the planning stage, so as to improve the investment efficiency of electric grid project.

1. Research background
As a typical capital intensive enterprise, the investment and construction of engineering project, as an important part of enterprise investment, has an important support for ensuring production safety and reliable operation, and also has an important impact on the company's investment planning, cost management and efficiency improvement. With the further promotion of the power system reform, the transmission and distribution price reform has been comprehensively deepened, from "building mechanism" to "strong supervision", and the effectiveness of investment has become the focus of supervision by all parties, so that grid companies must pay more attention to the efficiency and efficiency of investment. Therefore, in order to meet the regulatory requirements of transmission and distribution price, it is imperative to determine the project investment estimation scientifically and reasonably, plan the investment plan reasonably and arrange the project reasonably, apply the optimization cost prediction method in combination with big data mining, and improve the fine control level of investment, so as to provide important support for the stability of transmission and distribution price of the company.

2. Construction of data mining model for unit scale cost of Electric Grid Project
Based on the main technical conditions, final account data, general design scheme type, price information of main equipment and other historical data of project cost of electricity grid construction
projects of 35kV and above, the real-time cost information basic database of the whole process of power grid infrastructure construction is constructed. Based on the common design or common cost typical scheme database, the typical sample database items are selected from the basic cost information database, the unit scale cost level of various types of projects is calculated by classification, and the reference value standard database of unit cost investment is established. By mining the law and characteristics of the value of power grid infrastructure projects from multiple scenarios and perspectives, we can support the estimation of investment amount in the planning stage, the evaluation of investment estimation in the feasibility study stage, and the prediction of investment control objectives in the planning stage. The idea of model construction is as follows:

![Figure 1. Construction idea of data mining model for unit scale cost of electric grid project.](image)

3. Building data mining model of unit scale cost of electric grid project

3.1 Construction of real-time cost basic database

3.1.1 Design data template and integrate the whole chain cost information. Comprehensively sort out the cost related information generated in the whole process of electricity grid construction project, design the basic database template of electricity grid construction project cost, define the fields required in the basic database of electric grid project cost, and cover the cost data of completed final account electricity grid construction project and single project in the whole process of 35kV and above. It includes three categories: basic information, technical conditions and value information of the whole process of electric grid project. It constructs the basic database field collection of electric grid project cost and integrates the whole chain project cost information.

3.1.2 Sort out data sources and build a real-time cost database
First, sort out the cost data sources of the whole process of electricity grid construction projects. Investigate and sort out the data of main technical conditions of current electricity grid construction projects (main network, distribution network), such as substation capacity, outgoing circuit number, distribution device type, line length, as well as data acquisition sources, standardization degree and difficulties of current data acquisition, such as feasibility study estimate, preliminary design estimate, completion final account, etc. For structured data, sort out the rules of automatic analysis of cost
analysis related data such as final accounts of completion; for unstructured data, study the data batch acquisition and batch processing technology based on artificial intelligence.

The second is to establish a dynamic update mechanism for the cost data of the whole process of electricity grid construction projects. In order to strengthen the rolling real-time update of data, the development department cooperates with finance, infrastructure construction and other relevant business departments to establish a dynamic update mechanism of cost data in the whole process of electricity grid construction project. In addition to the design of the system's automatic integration access route, the responsibility and division of labor for cost related data maintenance and management, data maintenance specification requirements, etc. shall be specified for the fields that do not yet have the system's automatic access to data, so as to ensure that the cost related data of the completed final account project can be updated to the cost basic database in real time.

The third is to form the basic database of real-time cost of historical projects. Collect 298 electricity grid construction projects and 953 single projects of 35kV and above that have been put into operation in 2013-2018. Considering the impact of replacing business tax with value-added tax, 250 projects and 780 single projects in 2016-2018 are reserved after excluding the data from 2013-2015. In addition, 223 single items missing from final account data are further eliminated, and 557 single items of effective cost basic database are available, which are samples of real-time cost basic database of historical projects.

3.2 Building data mining model of unit scale cost

3.2.1 Study on the main factors affecting the project cost. By using the methods of multiple linear regression, principal component analysis and correlation analysis, this paper analyzes the relationship between the construction cost, installation cost, equipment purchase cost, other costs and their related influencing factors item by item, dissects the hierarchical relationship between the direct and indirect influencing factors, and excavates the key driving factors affecting the project cost as the basis for scheme classification based on key technical parameters in the next step. The main influencing factors of substation and line single project in planning and feasibility study stage are as follows:

| Stage                      | Single item type          | Main technical conditions for classification scheme                                      |
|----------------------------|---------------------------|------------------------------------------------------------------------------------------|
| Planning stage             | Substation engineering    | Voltage level, Type of Substation, Type of power distribution device                      |
|                            | Line engineering          | Voltage level, Loop number, Conductor type                                               |
| Feasibility study stage    | Substation engineering    | Voltage level, Type of Substation, Type of power distribution device                      |
|                            | Line engineering          | Voltage level, Loop number, Conductor type, Icing, Wind speed                           |

3.2.2 Define the main technical conditions and determine the classification scheme

Through the study of the main factors that affect the cost of electric grid project, sort out the available technical conditions and parameters in the planning stage and feasibility study stage, determine the classification method and main technical conditions with reference to the typical general cost scheme, and clarify the reference value classification scheme of unit investment.

First, analyze the technical condition data that the model can obtain in different application stages. Combining the data conditions of planning system planning, investment and other related modules, and combining with the accessibility of data in planning stage, feasibility study stage, substation project and line project, determining the main technical conditions parameters for the classification of unit scale cost mining model.
Secondly, based on the typical scheme type of common cost, simplify and merge it, and determine the classification scheme of unit investment reference value. The classification scheme applied in the planning stage is mainly determined by merging and replacing the typical scheme of common cost; For the classification scheme applied in the feasibility study stage, it is basically consistent with the classification granularity of the typical scheme of common cost.

After in-depth research and analysis, at present, only three technical conditions of voltage level, substation type and distribution device type are considered for the application of substation project in the planning stage to classify. Meanwhile, referring to the typical common cost scheme, the application classification scheme of 110-750 kV electricity grid construction project in the planning stage is divided into 16 categories, and the application classification scheme in the feasibility study stage is divided into 30 categories. In the planning stage, the voltage level, circuit number and conductor model are considered for the overhead line project. In the feasibility study stage, the voltage level, circuit number, conductor model, icing and wind speed are considered for classification. In the planning stage, the application classification of 35-500kV electric grid projects is divided into 50 categories, and in the feasibility study stage, the application classification is divided into 100 categories.

Taking 110 (66) and 220 kV substation projects as examples, the classification results in the planning stage are shown in the green shading part of the figure below. When applied in the feasibility study stage, the classification standard is consistent with the classification of typical common cost schemes, and the classification results are shown in the blue shading part of the figure below.

![Figure 2. Example of application classification of substation engineering planning and feasibility study stage.](image)

### 3.2.3 Classified calculation of reference value of unit investment cost

Based on the completed final account project in 2016-2018, the sample database of unit scale investment mining model is constructed, and the unit scale cost level of each classification scheme of substation project and overhead line in recent two years is calculated by using the final account data of completion. At the same time, referring to the cost control line of power transmission and transformation project of State Grid Corporation of China (2015 Edition), the cost control line of double main transformers and the unit cost conversion coefficient of single main transformer, after
calculating the unit capacity cost level of double main transformers by using the model, according to the conversion coefficient, the unit capacity cost level of single main transformer of each classification scheme is obtained.

For the overhead line project, since the line project in the sample library usually corresponds to a variety of common design schemes, in order to ensure the accuracy of the calculated unit investment reference value results of each classification scheme, the cost control line and double circuit coefficient of single circuit line project in the State Grid Corporation of China transmission and transformation project cost control line (2015 version) are taken as the unit cost discounts corresponding to different circuit numbers calculate the coefficient, and calculate the weight of each circuit number path length according to the corresponding unit cost conversion coefficient multiplied by the number path length of each circuit in the sample project of overhead line. According to the calculated weights of different circuit numbers, the static investment (final accounts) of the sample line project is allocated, and then the unit path length cost of the sample project is calculated.

At the same time, according to the sample calculation results of each classification scheme and the method of "mean value ± standard deviation", the reference range of initial unit investment level is calculated.

3.2.4 Form the standard system of unit investment reference value

One is to establish the feedback and correction mechanism of abnormal results. On the basis of the above initial forecast unit cost level, analyze and calculate the rationality of the results of each classification scheme one by one. When the results are abnormal, return to the sample database to eliminate the abnormal sample data, and gradually correct the calculation results until the calculation results are reasonable.

Second, considering the difference between the final accounts and the cost level in the early stage of the project, appropriate margin shall be reserved to further adjust the calculation results. On the basis of the initial unit scale cost level calculated based on the final accounting data, the historical project investment balance level is considered as the adjustment coefficient to form a unit investment reference system suitable for the planning and feasibility study stages, which is divided into project types and design schemes.

4. Application of data mining model for unit scale cost of electric grid project

4.1 Planning stage: estimation of investment scale of auxiliary projects

In the planning stage, the planning management personnel can import the main technical parameter information in the project planning database in the system, such as the input voltage level of the substation project, the type of the substation, the type of the distribution device, the power transformation capacity, etc. if the input technical parameters can be completely consistent with the technical conditions of the classification scheme in the model, the model can automatically match to get the general cost typical scheme; if not The model uses neural network, support vector machine and other classification algorithms to automatically predict the investment amount of the project.

Taking the new 110kV substation project of Rizhao Chuanzi (Qiaonan) as an example, in the planning stage, the basic parameters of the single project in the planning stage are: three 50MVA main transformers will be installed in the long term, and two 50MVA main transformers will be installed in this stage. The 110kV power distribution device adopts the indoor GIS equipment layout, the 10kV power distribution device adopts the user's internal opening and closing cabinet layout, and the main transformer outdoor layout. Through the input of planning stage parameters, the model output total investment reference value (static investment) is 36.237-41.26 million yuan, and the feasibility study (static investment) is 40.7 million yuan, within the predicted reference range. The specific input and output results are as follows:
4.2 Feasibility study stage: supporting investment economic review

In the stage of feasibility study investment decision-making review, the preliminary management personnel of the project can input relevant basic information of the project, the system can automatically output the investment reference value of the feasibility study stage of the project through the model, and can compare and analyze with the investment filling value and common cost reported by the designer, so as to support the development of the feasibility study stage of the project and the financial department to review the feasibility of the investment scale economy of the project help project investment make accurate decisions.

Taking the new construction of Zhangjiatun 110kV substation as an example, the feasibility study stage, the basic parameters of the single project in the feasibility study stage are: install 2 sets of 63MVA in the current period, install 3 sets of 63MVA in the long-term; 2 times of 110kV outlet and 28 times of 10kV outlet; Capacitor 2 × (4.8 + 4.8) Mvar, using the common design 110-A2-4, feasibility study static investment of 39.87 million yuan. By inputting the feasibility study stage parameters or the common design scheme, the model output total investment reference value (static investment) is 3908.06~5138.12 yuan, and the feasibility study static investment is 39.87 million yuan, within the predicted reference interval. Specific input and output results are as follows:
4.3 Plan preparation stage, assist in forecasting investment control objectives

By calculating the historical investment cost balance level between the feasibility study investment estimation, the approved budget estimate and the final accounts of the completed projects of different voltage levels and project types, the investment cost control coefficient of all kinds of electricity grid construction projects is set as the upper limit of the total project investment control when the annual investment of new projects is prepared or the plan of the continued projects is adjusted, so as to further reduce the investment of electricity grid construction projects capital scale, arrange more investment projects as much as possible, improve the ability of investment planning department to arrange investment projects reasonably, and assist in accurate preparation or plan adjustment of annual investment plan.

5. Conclusion

This paper mainly takes the historical cost data of electricity grid construction projects of 35kV and above as the object, identifies the main influencing factors of cost through qualitative and quantitative analysis methods, innovatively and flexibly uses the typical scheme classification method of common cost for reference, replaces the method of common cost value information with historical data value information, establishes the data mining model of unit cost, and conducts multi scene and multi angle mining the law and characteristics of the value of electricity grid construction projects assist in the accurate decision-making of power grid infrastructure investment and improve the investment efficiency and efficiency of power grid enterprises.

References

[1] Li, J.Q., Wang, D.J., Liu, L., Zhu, X. H., (2016) Research on cost management of electric grid project based on construction drawing budget mode. Engineering and Construction.
[2] Yu, C., (2016) Study on the influencing factors and evaluation model of power grid project cost. Beijing: North China Electric Power University.

[3] Xie, W.B., (2017) Research on the design of power grid engineering cost information system. Hubei University of Technology.

[4] He, G.P., (2017) Research on cost management of intelligent substation based on life cycle theory. Beijing: North China Electric Power University.

[5] Chang, X.G., (2018) Discussion on general cost management mode of power grid construction project. Project Management.

[6] Liu, Q., Wang, M.L., Liu, T., Yan, W., Ding, C., (2019) Research on standardized technology and economic management mode of power grid construction. Project Management Technology.

[7] Zhang, J.P., Wang, J.J., Ding., W.M., Hou, X.H., (2019) Study on the whole process lean management of power engineering cost. Equipment Management and Maintenance.

[8] Wang, L.Z., (2019) Analysis of problems and Countermeasures in the early management of distribution network construction project. China Management Informationization.

[9] Ye, Z.W., Tang, Y.M., Jiang, N., Zhang. J., Wang, J.Y., (2019) Study on cost characteristics and control measures of power transmission and transformation project.

[10] Wu, D., (2019) Research on the application of the theory and method model of transmission and transformation project cost analysis. Digitcw.