Short-term Load Forecasting System for Power System Based on Big Data

Yang Donghui1,*

1Kunlun Tourism College Heilongjiang Institute of Technology, HeiLongJiang, 150040, Harbin, China

*Corresponding author e-mail: ydonghui0010@hljit.edu.cn

Abstract. Short-term load forecasting of power system is an important task of power distribution system. Accurate short-term load forecasting provides the best configuration for grid power generation and distribution, maximizing energy saving and ensuring stable operation. This paper aims to study the design of short-term load forecasting system of power system based on big data. On the basis of analyzing power system load forecasting algorithms, classification of load forecasting, characteristics of load forecasting and system design principles, each module of the system is designed in detail, and finally tested the performance of the system. The test results show that the system has no adverse reactions in the use of a large number of users and repeated operation for a long time. In the case of large throughput, the system has a satisfactory response time and relatively reliable system stability.

Key words: Big Data, Power System, Short-term Load Forecasting, System Design

1. Introduction

Smart grid is the direction of future grid development. In the smart grid environment, advanced information and management technology can ensure the safe, stable and efficient operation of the power system [1-2]. In order to improve the economic benefits of the power industry and increase the overall economic benefits of society, our country's energy industry is gradually turning to the process of improving and optimizing production plans, social resources and the optimal allocation of various resources, and strives to achieve the optimal allocation of various resources. The various entities, including power generation companies, energy planning and transmission departments, need to better understand changing legislation and increasing grid load trends [3-4]. Therefore, accurate prediction of grid load is an important condition for scientific production, transmission and distribution of energy, and also an important condition for ensuring the safe and stable operation of the grid.

Over the years, many scholars have conducted in-depth research on short-term load forecasting methods. Some scholars introduced a new linear regression-based model for calculating short-term system load forecasts. The model uses weighted least squares linear regression technology for robust parameter estimation and parameter estimation under heteroscedasticity, using reverse variables error technology to reduce the impact of potential errors in the explanatory variables on load forecasting, distinguish the maximum value of daily peak load forecasting and hourly load forecasting independent of time to prevent negative deviations in peak forecasting [5]. Some scholars said that due to the
general nonlinear mapping ability of prediction, artificial neural networks have played a vital role in predicting power load. Support vector machines have been successfully used to solve nonlinear regression and time series problems. However, support vector machines It is rarely used to predict electric load [6]. Some researchers have proposed a new method for power system load forecasting based on support vector machines. Compared with the artificial neural network that embodies the principle of risk minimization, the proposed algorithm embodies the principle of structural risk minimization and has more general performance and accuracy [7]. Other researchers introduced the latest neural network algorithm support vector machine into short-term power load forecasting, and compared its performance with the autoregressive model. The results show that the support vector machine is better than the autoregressive model that uses the same data to build and test the two models based on the root mean square error between the actual data and the predicted data [8]. At present, short-term power load forecasting research has become a key research content in the power field, and many results have been achieved in this regard, providing data support for the formulation of power grid planning schemes.

This paper has thoroughly reviewed the research background and related literature of power system load forecasting; established a load forecasting model by analyzing various forecasting algorithms, and finally designed and implemented a short-term load forecasting system for power system based on big data.

2. Design of Short-term Load Forecasting System for Power System Based on Big Data

2.1. Analysis of Power System Load Forecasting Algorithms

(1) Non-linear programming method
Power grid load forecasting can summarize many practical problems as optimization problems. In linear programming problems, all constraints related to the objective function are linear functions of variables. The most commonly used method for solving linear programming is the simplex method. However, there are other problems that make it difficult to express the objective function or constraint as a linear function. This type of programming problem, in which the objective function or constraint contains a nonlinear function. It is called the nonlinear programming problem [9-10].

(2) Regression analysis method
Regression analysis is a mathematical method (linear regression, nonlinear regression) for studying variables and their dependencies. Mainly analyze the relationship between weather changes (mainly temperature changes) and load. The regression equation is

$$y = a0 + a1x + a2x^2$$

(1)

Use the common least squares method to minimize the Q error between the regression curve and the actual data. Let \( y_i \) be the true value, \( a0 + a1x + a2x^2 \) is the adjustment (regression) value, and the sum of the squares of the difference of each adjustment value to avoid positive and negative errors due to other offsets.

$$Q = \sum_{i=1}^{n} \left[ y_i - (a_0 + a_1x_i + a_2x_i^2) \right]^2$$

(2)

Among them, \( y_i \) is the electricity load of the year.

(3) Time series method
Time series models include autoregressive (AR), moving average (MA), autoregressive-moving average (ARMA) and other models; determining the model type is the first thing to consider, you need
to use average and correlation functions to analyze the time series and analyze other parameters determine the model type. After determining the model, use the data from the initial sequence to estimate the parameters of the model [11-12].

(4) Characteristic curve similarity method
It searches historical data for daily load data with the same characteristics (weather, date, etc.), and predicts the trend of load changes due to power outage maintenance. This method is used to forecast holiday load. When the cycle characteristics of the load are particularly clear, the result is the best.

2.2. Classification of Load Forecasting
(1) Short-term load forecasting is a forecast of the electricity load of users in the next hour. This type of prediction is mainly used for emergency response, health control, quality control and safety monitoring;

(2) Short-term power load forecasting is usually used to predict the load within 1-7 days. It can coordinate the work of energy management departments, coordinate the coordination of energy production between hydroelectric power plants, and ensure the operating efficiency of the power sector and the accuracy of power transaction volume;

(3) Medium-term power load forecasting is usually one month to one year load forecasting, which is mainly used for fuel planning, unit maintenance, tank transportation, etc.;

(4) Long-term load forecasting is mainly used to formulate future development plans for large-scale power grids. Under normal circumstances, it is necessary to predict load prices over the years based on load price analysis and classification, grid capacity growth decision-making and grid monitoring, and rationally organize grid construction. Load forecasting provides macroeconomic decisions about power grid construction in the next few decades, enabling power supply to cope with rapid economic growth and rising living standards.

2.3. Characteristics of Load Forecasting

(1) Inaccuracy
The future growth of the load is clear or uncertain, it is easily affected by many factors, and various factors will change from time to time. Some changes can be estimated in advance, and some changes are difficult to estimate. When faced with the impact of sudden changes, it is the characteristic of inaccurate predictions.

(2) Multi-scheme
Due to inaccurate predictions and condition characteristics, all growth conditions that occur under load must be predicted, which will create a forecast plan under all possible conditions. The essence of load forecasting is to estimate or estimate the future growth trend based on internal laws. Therefore, it is necessary to systematically and scientifically analyze and summarize the many principles contained in load forecasting to provide scientific guidance for load forecasting.

(3) Conditional
Load forecasting is carried out under certain conditions. Judging from the situation, it can be divided into hypothesis and necessity. If the inner law of the load can be fully understood, the condition is inevitable. In this case, the prediction results obtained are usually more accurate or reliable. However, in some cases, due to the uncertainty of the future growth trend of the load, it is also necessary to analyze assumptions. In order to limit certain conditions in the forecasting process, relevant units can use the forecast results more conveniently.

2.4. System Design Principles
(1) Stability and safety
Ensuring system safety, reliability, and stable operation is an important principle for the design of power grid load forecasting systems. System stability is mainly manifested in extremely special circumstances. The load forecasting system can operate normally. When designing the system, carefully consider security factors such as system security, data security, network security, etc., create multiple IDs and network firewalls, and at the same time protect data, prevent data, and protect data privacy.

(2) Scalability and maintainability
Mature software must not only have many characteristics such as stability, compatibility, and readability, but more importantly, it must ensure the scalability and maintainability of the software. With the continuous advancement of technology, the requirements for the accuracy of prediction results are getting higher and higher, and the system needs to be manageable, maintainable, and upgradeable.

(3) Accuracy
The grid short-term forecasting system is an important part of the grid power management system. This is an important reference information for the transmission sector to formulate power generation and transmission plans. Therefore, the accuracy of algorithms and data is very important. Since the accuracy of the data directly affects the accuracy of the algorithm, missing data and abnormal situations need to be dealt with. The accuracy of the algorithm is reflected in the construction of the algorithm itself and the rationality of the algorithm's selection of data.

3. Experiment

3.1. System Overall Design
The short-term load forecasting system of power system based on big data mainly includes four modules. The detailed design is shown in Figure 1.

![Figure 1. System overall design](image)

3.2. Detailed design of System Function Modules

3.2.1 Data Management Module
(1) Information module
The so-called information part is also called the "environmental part". Its main function is to collect and classify factors such as weather, humidity, and temperature in the middle of the forecast process.
These factors affect load forecasting, and information collection and classification can play a leading role in decision-making, reducing the influence of external factors.

(2) Data interface module
This part is the connection point between the database and the system. As an important part of power load forecasting, this database contains all system data. In other words, it is the "brain" of the system. Because of the bitterness of the data, the management of the entire data is possible. For example, all the more standard functions, such as editing, deleting, and searching, require the help of the database. Based on this, all the characteristics of the users of the predictive interface system are sent through the data interface.

3.2.2 Data Processing Module
This module is mainly about the preprocessing of historical data. Since the load forecast is based on the original data, the accuracy of the forecast is also affected by the original data. The initial load data is collected in real time from the terminal. This improves the quality of some initial load data. This function preprocesses historical data to ensure the accuracy of prediction results. The operation of this module handles two main types of data anomalies (data skipping and data missing). When inputting historical data, the system will preprocess the data.

3.2.3 Load Forecasting Module
(1) Load correction
Some sudden factors and special events have a greater impact on historical load data, and the deviation of collected data is greater than that of normal historical data. Therefore, such data needs to be specially modified. This data is also called "abnormal data". The existence of abnormal data will affect the normal load forecast, and affect the results and the accuracy of the forecast. Therefore, such data must be specially processed to avoid such adverse effects. Therefore, a module is needed to process these data to install the system. The system provides manual and automatic batch correction of load data. In the manual load repair interface, after the user requests to specify the time interval, unit, and repair error limit, if the loading data error of the selected unit in the selected time period is greater than the repair error limit, the repair interface will be entered, and the data will be loaded manually by the user.

(2) Load forecast
There are two types of load forecasting: conventional daily forecasting and holiday forecasting. Before making a prediction, the user must first define the prediction format and related parameters. If the forecast results meet the requirements, you can modify the results to report or save the current forecast plan.

(3) Modification and Reporting of forecast results
If the forecast result meets the requirements, the user can choose to report it. If it does not meet the requirements, you can fix it and report it. At the same time, users can also request and change the forecast results for selected units and time periods.

3.2.4 System Management Module
The main function of system management is to efficiently manage logs, permissions, users, passwords, etc., ensure system security and standardization. Among them, user management refers to the addition, deletion, modification, and checking of users. Rights management is the allocation and cancellation of user rights. The password management manages all user login passwords. Log management is the management of system history.
4. Discussion
Simulate whether it can run normally under the load of the highest number of concurrent users, and how the resources, network and system response time, performance and traffic of each server will change after the load. If it does not meet the requirements, you need to find out the performance barriers of the test platform and configure it to solve platform performance problems such as memory overflow and short response time.

**Table 1.** Connection concurrency statistics

| Request transaction success rate | Number of transactions processed per second | Average response time/s | Concurrency duration/s |
|---------------------------------|---------------------------------------------|-------------------------|------------------------|
| 500 100%                        | 169.63                                      | 0.005856                | 291.84                 |
| 1000 100%                       | 222.4                                       | 0.00477                 | 446.825                |
| 2000 100%                       | 371.03                                      | 0.002687                | 267.781                |

![Figure 2. Average system response time](image)

As can be seen from Table 1 and Figure 2, the number of concurrent threads for all interfaces this time is 500, 1000, and 2000. In the test, the system was used by a large number of users and repeatedly operated for a long time, and the system did not show any adverse reactions. In the case of large throughput, the system response time was satisfactory and the system stability was relatively reliable.

5. Conclusions
Short-term load forecasting of the power system plays an important role in the safe operation of the power grid and the reliability of power supply. Therefore, it is the social responsibility of the power company to accurately predict the short-term load of the energy system. With the development of science and technology and the improvement of people's living standards, the way of electricity consumption has gradually changed, which also makes the forecast of power system load more difficult. In this case, if you want to accurately predict the short-term load of the power system, you need to carefully consider the characteristics of the load, carefully consider the factors that affect the power load, establish a suitable prediction model, and continuously modify and improve the modeling
based on the prediction results.

References
[1] Merkel G D , Richard I D , Brown R H . Short-Term Load Forecasting of Natural Gas with Deep Neural Network Regression[J]. Energies, 2018, 11(8):2008-.
[2] A J L , B T H A , C S C F . Benchmarking robustness of load forecasting models under data integrity attacks[J]. International Journal of Forecasting, 2018, 34(1):89-104.
[3] ZIEL, Florian. Modeling public holidays in load forecasting: a German case study[J]. Journal of Modern Power Systems & Clean Energy, 2018, 6(2):191-207.
[4] Li L, Wang H, D Yong. A VVWBO-BVO-based GM (1,1) and its parameter optimization by GRA-IGSA integration algorithm for annual power load forecasting[J]. Plos One, 2018, 13(5):e0196816.
[5] Carvallo J P , Larsen P H , Sanstad A H , et al. Long term load forecasting accuracy in electric utility integrated resource planning[J]. Energy Policy, 2018, 119(AUG.):410-422.
[6] Fu G . Deep belief network based ensemble approach for cooling load forecasting of air-conditioning system[J]. Energy, 2018, 148(APR.1):269-282.
[7] Shepero M , Dennis V , Munkhammar J , et al. Residential probabilistic load forecasting: A method using Gaussian process designed for electric load data[J]. Applied Energy, 2018, 218(MAY15):159-172.
[8] Lin L , Xue L , Hu Z , et al. Modular Predictor for Day-Ahead Load Forecasting and Feature Selection for Different Hours[J]. Energies, 2018, 11(7):1899-.
[9] Qiu X , Suganthan P N , Amaratunga G A J . Ensemble incremental learning Random Vector Functional Link network for short-term electric load forecasting[J]. Knowledge-Based Systems, 2018, 145(APR.1):182-196.
[10] Khomutov S O , Serebryakov N A . Creation of a short-term load forecasting neural network model of electrical engineering complex of section regional electric grid 6-35 kV[J]. Transportation Systems and Technology, 2020, 6(1):80-91.
[11] Ananthapadmanabha T , Veeresha A G , Kumar M , et al. An adaptive power system management with DG placement and cluster-based load forecasting by CS, K-means and ANN algorithms[J]. International Journal of Power Electronics, 2021, 13(3):380.
[12] Yang X , Yu X , H Lu. Dual possibilistic regression models of support vector machines and application in power load forecasting[J]. International Journal of Distributed Sensor Networks, 2020, 16(5):155014772092163.