Postprint of Short-Term Power Load Forecasting Algorithm Based on BP Neural Network and Maximum Deviation Similarity Criterion

Authors: Luo Yuhui, Cai Yanguang, Qi Yuanhang, Huang Helie

Date: 2018-08-13T00:00:00+00:00

Abstract

To address the issues of strong randomness, low stability, and unsatisfactory prediction accuracy in enterprise power load, a short-term power load forecasting algorithm based on BP neural network using the maximum deviation similarity criterion is proposed. First, the maximum deviation similarity criterion algorithm is modified, and the distance between the load feature vector of the forecast day and the class center load features after clustering is utilized to determine the similar day category of the forecast day. Then, the load data of the similar day category after clustering is used as the training data for the BP network, outputting the 96-point load values for three consecutive days starting from the forecast day. Experimental results show that the proposed short-term power load forecasting method achieves significant improvements in both accuracy and network training time, demonstrating high effectiveness and practicality.

Full Text

Short-Term Power Load Forecasting Algorithm Based on BP Neural Network with Maximum Deviation Similarity Criterion

Luo Yuhui[1], Cai Yanguang[1]†, Qi Yuanhang[2], Huang Helie[1]
(1. School of Automation, Guangdong University of Technology, Guangzhou 510006, China;
2. Zhongshan Institute, University of Electronic Science and Technology of China, Zhongshan Guangdong 528402, China)

Abstract: To address the challenges of strong randomness, low stability, and poor prediction accuracy in enterprise power load forecasting, this paper proposes a short-term power load forecasting algorithm based on BP neural network with maximum deviation similarity criterion (MDSC). The method first modifies the MDSC algorithm and uses the distance between the forecast day' s load feature vector and the cluster centers obtained through MDSC-based clustering to determine the similar day category for the forecast day. Subsequently, the clustered similar-day load data serves as training data for the BP network, which outputs 96-point load values for three consecutive days starting from the forecast day. Experimental results demonstrate that the proposed method significantly improves both prediction accuracy and network training time, exhibiting high effectiveness and practicality.

Keywords: demand response; power load forecasting; BP neural network; maximum deviation similarity criterion; clustering algorithm

0 Introduction

Accurate and fast forecasting methods represent a critical research focus in power load analysis. Currently, China is implementing comprehensive reforms in power demand-side management across numerous cities. Demand-side management involves enterprises and users establishing refined electricity management systems to monitor and optimize their power consumption. Through technical measures such as equipment retrofitting and system-wide optimization, enterprises can improve their electricity usage patterns, achieving energy conservation, emission reduction, and cost reduction. Demand response (DR) serves as a key mechanism for coordinating electricity consumption with other grid components, with DR-oriented power load forecasting forming the core of demand-side management and representing a crucial means to ensure grid stability and achieve energy savings.

Short-term power load forecasting predicts hourly power load values for the upcoming day or several days. Accurate load forecasting is economically and operationally vital for power grids. High-precision, rapid forecasting methods are currently a hot research topic. Existing short-term load forecasting methods fall into two categories: traditional approaches and artificial intelligence algorithms. Traditional methods include grey prediction and linear regression, while AI-based approaches primarily comprise neural networks, support vector machines, and fuzzy systems. Power load forecasting has long been a research focus among scholars worldwide, yielding numerous excellent methods such as fractal extrapolation and probability density-based support vector machine regression.

However, enterprise power load is influenced by numerous complex and highly uncertain factors, and single prediction models often exhibit limitations that prevent them from achieving required accuracy levels. Consequently, practical applications typically require consideration of multiple influencing factors to extract principal and effective features for dimensionality reduction, often employing combined models. For instance, one study proposed a cascaded BP-RBF neural network where non-load factors serve as inputs to the front-end BP network, which outputs peak and valley loads for the forecast day to roughly determine the load level. This load level then selects the similar day type, determining which class of load data should be fed into the RBF neural network for 24-point daily forecasting, achieving favorable results. Another approach used wavelet decomposition with fuzzy grey clustering to categorize weather factors, using the forecast day' s weather factors to determine similar day types for neural network input, showing improved performance. A third method performed systematic clustering on historical load data before forecasting, grouping data by inter-data distances into three categories and using same-class data as neural network inputs, thereby improving accuracy.

Nevertheless, many scholars rely solely on peak and valley loads to describe the forecast day' s approximate load level, overlooking enterprise-specific consumption characteristics and patterns. This approach struggles to accurately characterize the forecast day' s load level and results in large-scale input data for subsequent RBF networks. Building upon our previous research and addressing these issues, this paper proposes a short-term power load forecasting method using a BP neural network based on maximum deviation similarity criterion (MDSCBPNN). The MDSCBPNN approach first clusters load curves using the MDSC algorithm, grouping curves with similar shapes and load values within a certain range to obtain stable load data curves and reduce instability. It then uses the average load, maximum load, minimum load, and daily electricity consumption from the 100 days preceding the forecast day as BP network training data to obtain these four metrics for the forecast day, forming the forecast day' s load feature vector V. For similar day selection, the method determines the forecast day' s similar category by minimizing the distance between V and the cluster centers, enabling more accurate identification. Finally, using the stable similar-day load data as BP neural network training data yields accurate 96-point load curves for the forecast day. Experimental results demonstrate high prediction accuracy, fast training speed, and excellent performance, confirming the method' s practicality and rationality.

1 Maximum Deviation Similarity Criterion BP Neural Network Model

1.1 Maximum Deviation Similarity Criterion Algorithm

The MDSC algorithm clusters power load data based on morphological similarity according to the maximum deviation similarity criterion. However, this approach suffers from two drawbacks: large numerical differences among load curves within the same cluster, and failure to account for potential jump peaks at certain points. Therefore, this paper modifies the MDSC algorithm for power load data clustering by incorporating deviation restriction intervals and using original load data to ensure accurate reflection of enterprise user consumption characteristics and demand response potential.

For n m-dimensional power load data points, where the i-th load data is xi =

(xi_1, xi_2, ···, xim), with xik representing the load value at the k-th time point of the i-th load data (i = 1, 2, ···, n; k = 1, 2, ···, m):

a) Let sijk be the absolute difference between xi and xj at corresponding time points (i, j = 1, 2, ···, n; k = 1, 2, ···, m), calculated as:

$$s_{ijk} = |x_{ik} - x_{jk}|$$

b) Let nij be the number of sijk values satisfying sijk $\leq \gamma$, called the similar time points between xi and xj. Let mij be the maximum number of consecutive sijk values satisfying $\gamma <$ sijk $< \delta$, called the maximum consecutive deviation time points between xi and xj (i, j = 1, 2, ···, n; k = 1, 2, ···, m). Here, γ ($0 \leq \gamma \leq 1$) is a preset constant called the maximum deviation threshold, measuring similarity between corresponding time point load values; δ is the maximum allowable deviation for outlier points. When sijk $\leq \gamma$, xik and xjk are considered similar; otherwise, they are dissimilar. If sijk $> \delta$, the two curves are considered dissimilar. The expression for mij is:

$$m_{ij} = \max\{k_2 - k_1 + 1 | \gamma < s_{ijk} < \delta, \forall k \in [k_1, k_2]\}$$

In this paper, to minimize load data instability, γ is set to 0.10 and δ to 0.25.

Using load data xi as the comparison center, calculate nij and mij between xj and xi (i, j = 1, 2, ···, n). If nij and mij simultaneously satisfy:

(a) nij $\geq$ n_0, where n_0 = [α × m], α ($0 \leq \alpha \leq 1$) is a preset constant called the similarity degree.

(b) mij $\leq$ m_0, where m_0 = [β × m], β ($0 \leq \beta \leq 1$ - α) is a preset constant called the consecutive deviation degree.

These conditions (a) and (b) constitute the maximum deviation similarity criterion. Load data xj is considered similar to xi, and all xj similar to xi are regarded as similar curves.

To cluster curves with similar morphology and low deviation, this paper sets α to 0.80 and β to 0.20.

c) For all i, j = 1, 2, ···, n with i $\leq$ j, using xi as the comparison center, compare nij and mij with n_0 and m_0 respectively. All xj satisfying the maximum deviation similarity criterion are assigned to S(xi), i.e., S(xi) = S(xi) {xj}, and xj is removed from the original load data set U, where S(xi) is the set of curves similar to xi.

d) Calculate D(xi) for each S(xi) using:

$$D(x_i) = \sum_{x_j \in S(x_i)} d(x_i, x_j)$$

where the distance d(xi, xj) is calculated as:

$$d(x_i, x_j) = \sqrt{\sum_{k=1}^{m} (x_{ik} - x_{jk})^2}$$

If xi yields the minimum D(xi), then xi is the class center of S(xi).

1.2 BP Neural Network

A three-layer BP neural network model is employed, comprising an input layer, hidden layer, and output layer. The number of hidden layer neurons is determined by:

$$H = \sqrt{M + N} + a$$

where H is the number of hidden layer neurons, M is the number of input layer neurons, N is the number of output layer neurons, and a is a constant in [0, 10].

The hidden layer transfer function is:

$$g(x) = \frac{1}{1 + e^{-x}}$$

The forecast day' s load level feature vector V is constructed from its average load v_1, maximum load v_2, minimum load v_3, and daily electricity consumption v_4:

$$V = [v_1, v_2, v_3, v_4]^T$$

For the i-th class load curve S(xi) with class center xi, its average load is vxi_1, maximum load vxi_2, minimum load vxi_3, and daily electricity consumption vxi_4. The distance between the forecast day' s load level feature vector V and the class center xi is:

$$d(x_i, V) = \sqrt{\sum_{j=1}^{4} \omega_j (v_{xij} - v_j)^2}$$

where ωj are weight coefficients. In this paper, $\omega_1 = \omega_2 = \omega_3 = 0.3$ and $\omega_4 = 0.1$.

The forecast day' s similar category Z is determined by the class center xi that minimizes the distance:

$$Z = \arg\min_i \{d(x_i, V)\}, \quad i = 1, 2, 3, ..., k$$

where k is the total number of historical load curve categories.

1.3 MDSC-BP Neural Network Forecasting Process

The forecasting procedure based on the MDSC-BP neural network model consists of the following steps:

a) For all load curves xi = (xi$_1$, xi$_2$, ···, xim), i = 1, 2, ···, n, calculate the similarity nij and deviation mij between each pair of load curves xi and xj (j = 1, 2, ···, n, j $\neq$ i) using equations (1) and (2). Apply the maximum deviation similarity criterion to determine whether xi and xj belong to the same class. Repeat until all load curves are processed to obtain clustering result R with r categories.

b) Calculate D(xi) for each class S(xi), i = 1, 2, ···, r using equations (3) and (4), and identify the class center xi that minimizes D(xi).

c) Initialize the BP network.

d) Normalize historical load data using equation (10):

$$x'_{ik} = \frac{x_{ik} - x_{\min}}{x_{\max} - x_{\min}}$$

where xik is the load value at time k on day i, and xmax and xmin are the maximum and minimum values of the power load data, respectively. This normalization maps load data to [0, 1] to accelerate neural network convergence.

e) Use the average load v$_1$, maximum load v$_2$, minimum load v$_3$, and daily electricity consumption v$_4$ from the 100 days preceding the forecast day as input data to obtain the forecast day' s load level feature vector V.

f) For all classes S(xi) and their corresponding class centers xi, i = 1, 2, ···, r obtained in steps a) and b), calculate the distance between the forecast day' s load level feature vector V and each class center xi using equation (8).

g) Determine the forecast day' s similar category Z using equation (9).

h) Use load data from class Z as training data for the BP neural network. Calculate the network output error .

i) If the BP network output error meets the precision requirement, proceed to step j); otherwise, return to step h).

j) Use the trained BP network for load forecasting, outputting 96-point load values for the forecast day.

2 Experiments and Analysis

This study uses historical power load data from an enterprise spanning January 1, 2016 to May 15, 2017 as training samples to predict the power load curves for May 16-18, 2017. The load data consists of 96 sampling points per day, taken every 15 minutes from 00:00 to 23:45.

2.1 Clustering Results and Analysis

First, the MDSC algorithm clusters the 501 days of historical power load data from January 1, 2016 to May 15, 2017. The clustering results in 130 categories, with the first 16 categories shown in . The first six categories contain 365 days, accounting for nearly 72% of historical data and effectively reflecting the enterprise' s electricity consumption patterns. Categories 7 through 130 contain very few data points each, with most being single-day categories (categories 16-130 each contain only one day). Therefore, categories 7-130 are treated as outlier days. The load clustering diagrams for the first six categories are shown in [Figure 1: see original paper]. Notably, Figure 1: see original paper shows nearly zero load values, representing rest days or holidays.

Further analysis uses the BP neural network to obtain the forecast day' s load level feature vector V. Using the average load, maximum load, minimum load, and daily electricity consumption from the preceding 100 days as inputs, the network outputs V. Based on the clustering analysis, categories 7-130 represent outlier data, so only distances between V and the centers of the first six categories are computed. The category Z minimizing the distance is identified as the forecast day' s similar category. The distance calculations between V and the six class centers are shown in , revealing that the smallest distance occurs with category 1, which is therefore selected as the training data input for the BP neural network.

The MDSC-BP neural network model predicts 96-point load values for May 16-18, 2017. The prediction results are shown in [Figure 2: see original paper], with comparative results presented in [Figure 3: see original paper] and . The error analysis employs three metrics calculated for each time point:

The point-wise error is calculated as:

$$e_{ik} = \hat{y}_{ik} - y_{ik}$$

where $\hat{y}_{ik}$ is the predicted load and y_{ik} is the actual load for day i (i = 1, 2, 3) at time k (k = 1, 2, ···, 96).

The maximum error, minimum error, and average error are:

$$e_{\max} = \max_i \{e_i\}, \quad e_{\min} = \min_i \{e_i\}, \quad e_{\text{av}} = \frac{1}{96} \sum_{i=1}^{96} e_i$$

The results demonstrate that the MDSC-BP neural network predictions closely match actual values, following the true load variation trends without the extreme peaks or valleys seen in single BP network predictions. The prediction curve is smoother and more accurate. As shown in [Figure 3: see original paper], prediction errors remain around 10%, with most points below this threshold. The maximum error decreases by nearly 70%, while the average error is reduced by almost half. The minimum error reaches zero for most days, indicating exact matches at some points. confirms significant improvements across all error metrics. Additionally, the proposed model reduces neural network training time by nearly half, enhancing prediction speed and efficiency.

The clustering analysis before forecasting groups unstable, irregular load data according to enterprise-specific consumption characteristics, resulting in stable data with similar morphology and numerical ranges. This stability prevents extreme peaks and valleys, enabling more accurate BP neural network predictions that are less affected by instability and spike values.

3 Conclusion

This paper proposes a short-term daily load forecasting method based on the MDSC-BP neural network model. The approach first applies MDSC algorithm to cluster historical load data, grouping morphologically similar curves with load deviations within a certain range. This reduces instability and eliminates irregular spikes and valleys while preserving enterprise consumption characteristics. The method then uses average load, maximum load, minimum load, and daily electricity consumption from the preceding 100 days as BP network inputs to obtain the forecast day' s load level feature vector V. The similar day category is determined by minimizing the distance between V and cluster centers. Finally, the stable similar-day load data serves as BP network training data to generate 96-point load curves for three consecutive days starting from the forecast day.

Experimental results for three days of short-term daily load forecasting demonstrate significant error reductions at each time point, with maximum error decreasing by nearly 70% and average error reduced by almost half. Neural network training time is also shortened by nearly half compared to single neural networks, improving efficiency and speed. The results validate that the proposed method achieves higher accuracy and shorter training time, demonstrating strong practicality and rationality.

References

[1] Huang Helie, Cai Yanguang, Cai Hao, et al. Traffic flow clustering algorithm based on maximum deviation similarity criterion [J/OL]. Application Research of Computers, 2018, 35(8). (2017-07-21) [2018-04-25]. http://kns.cnki.net/kcms/detail/51.1196.TP.20170721.1352.046.html.

[2] Run Huaguang, Chen Songsong, Zhong Ming, et al. Research and design of power demand side energy efficiency management and demand response system [J]. Power System Technology, 2015, 39(1): 42-47.

[3] Tian Shiming, Wang Beibei, Zhang Jing. Key technology of demand response under the condition of smart grid [J]. Proceedings of the CSEE, 2014, 34(22): 3576-3589.

[4] Yang Xuying, Zhou Ming, Li Gengying. Overview of demand response mech-

anism analysis and modeling in Smart Grid [J]. Power System Technology, 2016, 40(1): 220-226.

[5] Wang Beibei. Review of user demand response characteristics and capabilities for smart grid [J]. Proceedings of the CSEE, 2014, 34(22): 3654-3663.

[6] Sheng Wanxing, Shi Changkai, Sun Junping, et al. Characteristics and research framework of automatic demand response in smart electricity consumption [J]. Automation of Electric Power Systems, 2013, 37(23): 1-7.

[7] Zhang Yuseng, Kong Xiangyu, Sun Jiwei, et al. Multi time scale family energy management optimization strategy based on power demand response [J]. Power System Technology, 2018(6): 1811-1819.

[8] Lei Ming, Wei WuQing, Zeng Jinhui, et al. Influence of load control on reliability of power supply considering demand response [J]. Automation of Electric Power Systems, 2018, 42(10): 53-59.

[9] Zhang Zhisheng, Yu Daolin. RBF-NN short-term load forecasting model considering the comprehensive influence factors of demand response [J]. Proceedings of the CSEE, 2018, 38(6): 1631-1638, 1899.

[10] Wan Qiang, Wang Qingliang, Wang Ruihao, et al. Short term load forecasting based on support vector machine for a regional power grid [J]. Power System and Clean Energy, 2016, 32(12): 14-20.

[11] Song K B, Baek Y S, Hong D H, et al. Short-term load forecasting for the holidays using fuzzy linear regression method [J]. IEEE Trans on Power Systems, 2005, 20(1): 96-101.

[12] Liu Y, Zhai M Y, Lu L P, et al. An improved method for short-term power load forecasting based on fractal extrapolation [J]. Applied Mechanics & Materials, 2012, 229: 1077-1080.

[13] He Yaoyao, Liu Rui, Li Haiyan, et al. Short-term power load probability density forecasting method using kernel-based support vector quantile regression and Copula theory [J]. Applied Energy, 2017, 185(1): 254-266.

[14] Short M, Crosbie T, Dawood M, et al. Load forecasting and dispatch optimisation for decentralised co-generation plant with dual energy storage [J]. Applied Energy, 2017, 186(3): 304-320.

[15] Dahl M, Brun A, Andresen G B. Using ensemble weather predictions in district heating operation and load forecasting [J]. Applied Energy, 2017, 193(1): 455-465.

[16] Lusis P, Khalilpour K R, Andrew L, et al. Short-term residential load forecasting: Impact of calendar effects and forecast granularity [J]. Applied Energy, 2017, 205: 654-669.

[17] Liu Yang, Wang Wei, Ghadimi N. Electricity load forecasting by an improved forecast engine for building level consumers [J]. Energy, 2017, 139: 18-

30.

[18] Vogler-Finck P J C, Bacher P, et al. Online short-term forecast of greenhouse heat load using a weather forecast service [J]. Applied Energy, 2017, 205: 1298-1310.

[19] Qiu Xueheng, 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: 182-196.

[20] Brodowski S, Bielecki A, Filocha M. A hybrid system for forecasting 24h power load profile for Polish electric grid [J]. Applied Soft Computing, 2017, 58: 527-539.

[21] Zhang Xiaobo, Wang Jianzhou, Zhang Kequan. Short-term electric load forecasting based on singular spectrum analysis and support vector machine optimized by Cuckoo search algorithm [J]. Electric Power Systems Research, 2017, 146: 270-285.

[22] Yaslan Y, Bican B. Empirical mode decomposition based denoising method with support vector regression for time series prediction: a case study for electricity load forecasting [J]. Measurement, 2017, 103: 52-61.

[23] Panda S K, Jagadev A K, Mohanty S N. Forecasting methods in electric power sector [J]. International Journal of Energy Optimization and Engineering, 2018, 7(1): 1-21.

[24] Kuster C, Rezgui Y, Mourshed M. Electrical load forecasting models: a critical systematic review [J]. Sustainable Cities and Society, 2017, 35: 257-270.

[25] Shepero M, Van Der Meer D, Munkhammar J, et al. Residential probabilistic load forecasting: A method using Gaussian process designed for electric load data [J]. Applied Energy, 2018, 218: 159-172.

[26] Fan Guofeng, Peng Liling, Hong Weichiang. Short term load forecasting based on phase space reconstruction algorithm and bi-square kernel regression model [J]. Applied Energy, 2018, 224: 13-33.

[27] Chen Gang, Zhou Jie, Zhang Xuejun, et al. Daily load forecasting based on BP and RBF cascaded neural network [J]. Power System Technology, 2009, 33(12): 101-105.

[28] Zhang Ping, Pan Xueping, Xue Wenchao. Short term load forecasting based on wavelet decomposition, fuzzy grey clustering and BP neural network [J]. Electric Power Automation Equipment, 2012, 32(11): 121-125+141.

[29] Bai Yang, Zhao Guan, Dou Jingyan, et al. Short term load forecasting based on parallel membrane computing [J]. Power System Protection and Control, 2017, 45(7): 35-42.

Note: Figure translations are in progress. See original paper for figures.