International Conference on Applied Energy 2021
Nov. 29 - Dec. 5, 2021, Thailand/Virtual
Paper ID: 756

A GAN-CNN Based Model for Short-Term Load Forecasting

Xiangya Bu[1], Qiuwei Wu[2*], Jian Chen[1], Jinyong Dong[1]
1 School of Electrical Engineering, Shandong University, Jinan, China, 250100
2 Centre for for Electric Power and Energy, Department of Electrical Engineering, Technical University of Denmark, Kgs. Lyngby, Denmark 2800

ABSTRACT

This paper proposes a power system load forecasting method based on generative adversarial network and convolutional neural network (GAN-CNN), and applies it to the short-term load prediction. In this model, the generation layer and the discrimination layer form a maximum-minimum game and finally reach a Nash equilibrium. The data feature extraction method is integrated with the CNN convolution operation and variational mode decomposition (VMD) technology, which improves the quality of sample generation and reduces the prediction error. This paper provides a new and effective method for selecting similar days and forming the input matrix of the model. Finally, the real data were used to demonstrate the superior performance of the proposed method.

Keywords: convolutional neural network, generative adversarial network, input matrix, load forecasting

1. INTRODUCTION

Short-term load forecasting plays an important role in power system scheduling and operation. High-precision short-term load forecasting is beneficial to improve the utilization rate of power generation equipment and the safety and effectiveness of power system operation.

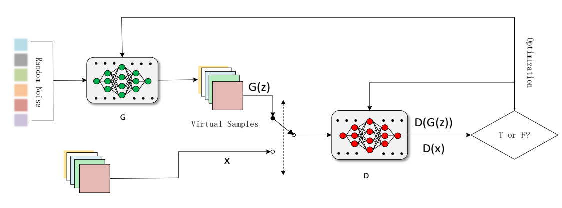
At present, the researches on load forecasting can be classified into two categories: mathematical statistical methods and machine learning algorithms. Most mathematical statistical methods [1] [2] rely on simplified linear function modeling, which is difficult to simulate the complexity and variability of load value. The prediction method based on machine learning algorithm [3] [4] [5] [6] can achieve more accurate prediction effect with the construction of neural network. However, the tranditional machine learning algorithms usually ignore the correlation of attributes in the data set and cause overfitting problem.

In this paper, a hybrid forecasting model is proposed based on generative adversarial network (GAN) and convolutional neural network (CNN) for the short-term day-ahead load forecasting. The hybrid algorithm exhibits a strong capability to generate virtual data similar to real data. A new method based on grey correlation degree is adoped to select the similar day and form the input matrix. A more accurate forecasting value could be obtained with the proposed input matrix compared with the traditional ones.

2. THE FRAMEWORK OF GAN-CNN MODEL

2.1 Generative Adversarial Network(GAN)

Inspired by the idea of zero-sum game, Goodflow first proposed a GAN structure including generative model and discriminative model [7]. The basic architecture of GAN is shown in figure 1.



$$\min_G \max_D V(D,G) = E \bullet \log[D(x)] + E \bullet \log[1 - D(G(Z))]$$

Fig. 1. GAN model

Generative Adversarial Network network is mainly composed of generative layer(G) and discriminative layer(D). The G of GAN can generate virtual data samples similar to the real value. The D of GAN accepts two inputs: the generated virtual samples and the real samples, and return a cluster of error discrimination results based on cross entropy. The error is fed back to the G and the D through the back propagation algorithm. The parameters of the model are updated based on the error with the Adam algorithm. In general, the dual objective optimization function is the most basic objective function for successful GAN training, as follows:

$$\min_{G}\max_{D} V(D,G) = E\bullet\log[D(x)] + E\bullet\log[1-D(G(Z))] \tag{1}$$

where E represents expected parameters; x and z represent real samples and random noise, respectively; D function and G function represent the forward effect of the two layers.

2.2 GAN combined with CNN

As shown in figure.3 and figure.4, CNN is combined with the GAN network. The main goal of CNN consists of extracting the local trend from adjacent values in the input arrary, by using the convolution operation [8]. It helps the model to achieve a stronger capability of extracting intrinsic nonlinear from the input data[9].

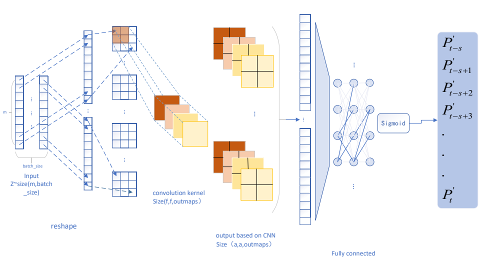


Fig.2. the structure of G

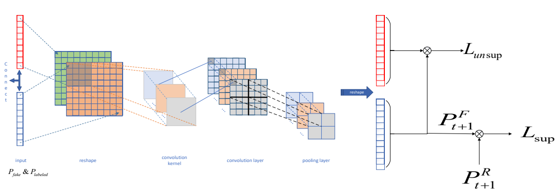


Fig.3. the structure of D

During training, only the weights of kernels are updated for the convolutional layers, thus the training of CNN is generally faster and more memory efficient to most other neural networks. In this paper, we adopt a multi-dimensional convolution operation, which has a deep significance for the overall performance improvement of the generative model.

The convolutional layer may be succeeded by a pooling layer. With this layer, the pooling operation is applied. It uses a sliding specific kernel to perform further data processing on the convolutional data, and returns the maximum or average value in the receptive field area to obtain the downlink data sample. Although the number of data dimensions are reduced, its main features have been successfully extracted, thereby reducing the complexity of the data.

3. CONSTRUCT THE INPUT MATRIX

Short term load forecasting is different from ultra-short-term load forecasting. Many factors should be taken into account, which undoubtedly increases the dimension of data preprocessing and the difficulty of forming network input arrary. In this chapter, considering the influence of comprehensive factors, a method, based on a variety of mathematical statistics methods, is used to construct the input matrix of the hybrid forecasting model.

3.1 Analysis of comprehensive actors acting on load forecasting:

(1) The impact of single meteorological condition;
(2) The impact of comprehensive meteorological index;
(3) The impact of cumulative effect generation;
(4) The impact of date type;
(5) The impact of similar day.

3.2 Input element

3.2.1 Temperature-humidity index (THI)

Temperature-humidity index is a comprehensive meteorological index established according to the temperature and humidity value of one day. It is conducive to analyze the comprehensive effect of these two meteorological factors on power system load. Among the various calculation methods of THI, the method recommended by the U.S. Meteorological Administration is adopted in this paper, which can be expressed as follows,

$$THI = \mathrm{Temp}_c + \frac{1450.8(\mathrm{Temp}_c + 235)}{4030 - (\mathrm{Temp}_c + 235)\ln Hmd} - 43.4 \tag{2}$$

Where, Temp_c is the temperature in Celsius; Hmd is the percentage humidity.

3.2.2 Date distance

Date distance refers to the distance between the historical day and the predicted day in the time dimension. The quantified date distance is formulated as:

$$D = \begin{cases} a^k , & a^k \geq b \\ b \end{cases} \tag{3}$$

where, D is the quantified value of date distance; α is the attenuation coefficient; b is the threshold; k is the number of days between historical and predicted days. In this paper, $\alpha = 0.9$, $\mathrm{b} = 0.1$.

3.2.3 Date type

Since power load value and week type have different units, the date type is needed to map onto another space to adap the load value.

Table.1 Date type mapping function table

Day of the week	Value before mapping	Value after mapping
Monday	1	0.1
Tuesday	2	0.2
Wednesday	3	0.3
Thursday	4	0.4
Friday	5	0.5
Saturday	6	3.2
Sunday	7	3.5

3.3 Selection of similar days based on weighted grey correlation degree

Similar day is an important concept in power system load forecasting. This section provides a new and effective method for selecting similar days. The calculation steps are given as follows:

Step 1: Calculate correlation coefficient (between historical days and predicted days) and the Pearson correlation coefficient (between each characteristic quantity and the load value). Calculate grey correlation degree with the above two sets of data.

Step 2: Calculate the weight of grey correlation degree.

Step 3: Calculate the load value of the weighted similar days.

3.3.1 Pearson correlation coefficient

Pearson correlation coefficient is a common mathematical calculation method to measure the linear correlation degree of two random variables X and Y. The absolute value of correlation coefficient reflects the linear correlation degree of two variables [10].

3.3.2 Weighted grey correlation degree

Grey correlation degree is an analysis method proposed by Mr. Deng Julong to measure the correlation degree of two factors based on the similarity of the development trend between the comparison curves. In this paper, the improved grey correlation theory is used for selecting similar days, which can be expressed as follows,

$$X_0 = \{x_0(1), x_0(2), ..., x_0(n)\} \tag{4}$$

$$X_i = \{x_i(1), x_i(2), ..., x_i(n)\} \tag{5}$$

$$\xi_i(k) = \sum_{k=1}^{n} \frac{\min_i \min_k |x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \rho \max_i \max_k |x_0(k) - x_i(k)|} \tag{6}$$

$$\omega_k = \frac{|R_k|}{\sum_{k=1}^{n} |R_k|} \tag{7}$$

$$G_i = \sum_{k=1}^{n} \xi_i(k)\omega_k \tag{8}$$

Where, X_0 represents the feature vector of the day to be predicted, including the predicted daily temperature, humidity, precipitation, temperature and humidity index, date type, etc. ; X_i represents the daily feature vector of the days before the predicted day; $\xi_i(k)$ represents the correlation coefficient of the kth eigenvector factor; n represents the total number of eigenvectors; ω_k is the corresponding weight of the correlation coefficient; R_k is the Pearson correlation coefficient between the k-th eigenvector factor and the daily load eigenvalue; G represents the weighted grey correlation degree between the final predicted day and the corresponding historical day feature vector.

3.3.3 Weighted similar days

In this paper, the selection range of similar days is set within the last three weeks. The weighted grey correlation degrees between the historical days（the last 21 days）and the predicted day are sorted in descending order. The three historical days with the largest correlation degree are selected as the benchmark for the selection of weighted similar days, and the daily load value of weighted similar days is calculated as follows,

$$\omega_k = \frac{|G_k|}{\sum_{k=1}^{n} |G_k|}$$

$$L_{similar} = \sum_{k=1}^{n} L_k \omega_k \tag{9}$$

Where, ω_k represents the weight of grey correlation degree ; L_k represents the historical load values with the largest three grey correlation degree, n=3; $L_{similar}$ represents the power load value of the weighted similar day.

3.4 Form input matrix

The final input matrix of the proposed forecasting model is presented in Table 1,

Table.2 the input matrix

Serial number	eigenvector	Serial number	eigenvector	Serial number	eigenvector
1	The highest temperature of the previous day	8	The lowest temperature of the day to be predicted	15	The electric load value at that time two days ago
2	The lowest temperature of the previous day	9	average temperature of the day to be predicted	16	The electric load value at the time t-3 of the previous day
3	Average temperature of the previous day	10	Relative humidity of the day to be predicted	17	The electric load value at the time t-2 of the previous day
4	Relative humidity of the previous day	11	Daily rainfall of the day to be predicted	18	The electric load value at the time t-1 of the previous day
5	Rainfall of the previous day	12	The THI of the day to be predicted	19	The electric load value at the time t of the previous day
6	The THI of the previous day	13	Date type mapping of the date to be predicted	20	The electric load value at the time t+1 of the previous day
7	The highest temperature of the day to be predicted	14	The electric load value at this time a week ago	21	The power load value at that time of the weighted similar day

4. LOAD FORECASTING BASED ON GAN-CNN

The number of output of the neural network can be changed. There will be 96 load power points in a day with a data resolution of 15 min. If the multi-output mode is adopted to generate 96-time prediction output at one time, it is bound to increase the complexity and calculation time of the neural network. In this paper, a model with 96 parallel neural networks is built. The prediction process of each time corresponds to one neural network. This construction not only simplifies the network architecture, but also improves the prediction performance of the model.

A single prediction model is introduced as follows,

4.1 Variational mode decomposition

Input data often show nonlinearity and irregularity. These futures will interfere with the actual prediction effect. In this paper, a signal decomposition technology, Variational mode decomposition(VMD), is used to decompose the original input data into several intrinsic mode functions (IMFs), which helps to weaken the interference of noise and outliers on signal prediction [11] .

4.2 Data preparation

Step1: The normalization method can effectively eliminate the adverse effects caused by different data forms and sources, and improve the training speed and performance.

The normalized calculation is defined as,

$$x_i^{'}(t) = (y_{\max} - y_{\min}) \times \frac{x_i(t) - x_{i,\min}}{x_{i,\max} - x_{i,\min}} + y_{\min} (y_{\min} = 0, y_{\max} = 1) \tag{10}$$

Where, $x_i(t)$ is the true value of the power load data at time t, $x_i^{'}(t)$ is the corresponding normalized value, $x_{i,\max}$ and $x_{i,\min}$ are the maximum and minimum values of IMFs respectively, and $y_{\max}$ and $y_{\min}$ are the maximum and minimum values of the normalized mapping space.

Step2: Set the label of the real power load data as 1. Then we get a set of real load data with their labels as $P_{labeled}$.

4.3 Training of GAN

Step1: Given a GAN model including generative layer(G) and discriminative layer(D).

Set the relevant parameters of the GAN model, including: (1) the batch size of the training data; (2) the epoch number of model training; (3) the learning rate.

Step2: Parameter update algorithm:

(a) The input data goes through the generation layer to obtain the marked virtual data sample P_{fake} ;

(b)Input training data samples and virtual data samples into the discriminant layer of the GAN model;

(c)Several groups of IMFs short-term forecasting results of power load are obtained by semi-supervised regression domain.

(d)The prediction error loss functions are expressed as follows,

$$L_G = \| E \bullet f(P_{fake}) - E \bullet f(P_{labeled}) \|^2 \quad (11)$$

$$L_D = L_{un\sup} + L_{\sup} \quad (12)$$

$$L_{un\sup} = -\| E \bullet f(P_{fake}) - E \bullet f(P_{labeled}) \|^2 \quad (13)$$

$$L_{\sup} = E \bullet (P_{t+1}^F - P_{t+1}^R)^2 \quad (14)$$

Where, $f(P_{fake})$ and $f(P_{labeled})$ represent the output of a layer in the middle of the discrimination model, P_{t+1}^F and P_{t+1}^R represent the predicted value and real value of power at the next time respectively.

The GAN model is optimized by using the method based on the Eq.(11)-(14) form the minimum and maximum game of the generation model and the discrimination model, and the parameters in the generation layer and the discrimination layer are adjusted and updated iteratively by using the alternating training method.

(e) Use alternating training method to iteratively adjust and update the parameters in the generative layer and discriminative layer of the GAN model.

Step3: Collect the next batch of training samples and repeat training progress until all the original training samples are traversed once.

Step4: When the number of iterations reaches the preset value or the prediction error is less than a certain threshold, the training process should be ended immediately.

4.4 Short-term Load Forecasting

Step1: Input the input-matrix ;

Step2: Several groups of IMFs short-term forecasting results of power load are obtained by semi-supervised regression domain;

Step3: The rolling adaptive algorithm is used to realize the multi-step forecasting of power load. According to the different proportion of predicted value and historical data, it can be divided into one-step prediction and multi-step prediction. Its structure is shown in figure 4.

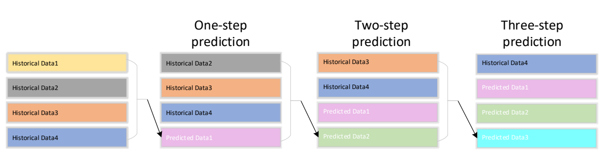


Fig. 4. Rolling prediction diagram

Step4: Add the short-term forecast results of IMFs to the power load forecast value.

5. CASE STUDIES

In this study, a real-world dataset from a certain region in China in 2012 has been applied for the forecasting model evaluation. The following results were implemented in matlab2020a.

The power load value of weighted similar day and the traditional similar day (day1) are shown in figure.5. It can be seen that the proposed method reached a better curve anastomosis rate.

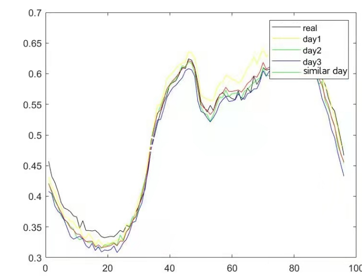


Fig. 5. weighted similar days

The one-step prediction and the multi-step prediction results are illustrated in figure.6. The prediction results from figure.6 show that the proposed method can be used in the multi-step load forecasting. It solves the practical problem of short-term load forecasting under the condition of lack of historical data.

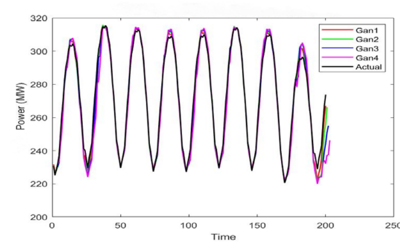


Fig. 6. one-step prediction and multi-step prediction results

The forecasting results of the GAN-CNN based forecasting method are graphically presented in figure.7. The data shown in figure.6 is selected randomly from the predicted data in this region in 2012.

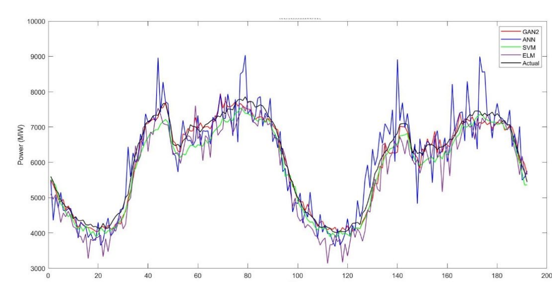


Fig. 7. prediction results

It can be found that, compared with other methods, the proposed method exhibits a superior performance. It can be clearly seen from the figure that the prediction value of Artificial Neural Network (ANN) in the data trough is lower than others. The Extreme Learning Machine (ELM) method is the most unstable in the presented prediction process. The results of ELM shows a large jagged change.

In this paper, three metrics, including mean absolute error (MAE),mean absolute percentage error (MAPE) and root mean square error(RMSE), are used for evaluating the load forecasting performance. The comparison of results is shown in table.2. The low error rate of GAN-CAN is obvious. Compared with ELM, it is close to reducing the error by almost half.

Table.2 evaluation parameters

	GAN	ANN	SVM	ELM
MAE	144.3292	357.1262	189.9727	329.3691
RMSE	176.6637	484.4415	234.2083	408.7274
MAPE	2.33%	5.77%	3.07%	5.32%

6. CONCLUSION

This paper combines the development status of artificial intelligence neural network, and proposes a power system load forecasting method based on GAN-CNN artificial intelligence neural network. The individual input matrix including weighted similar days applied in this model greatly improves the ability of data feature extraction. Combined with the application of VMD technology , the accuracy of load prediction could be improved dramatically. In addition, the rolling prediction algorithm provides a new way to solve the problem of data shortage. In a word, this paper provides a new effective method for short-term load forecasting of power system.

REFERENCE

[1] Niu Dongxiao, Cao Shuhua, Zhao lei, et al. Power Load Forecasting Technology and Its Application. Beijing: China Electric Power Press, 2001

[2] T.Masters. Neural Novel & Hybird Algorithms for Time Series Prediction[M]. John Wiley & Sons. Inc, 1995.

[3] Dong-Xiao Niu, Qiang Wanq and Jin-Chao Li, "Short term load forecasting model using support vector machine based on artificial neural network," 2005 International Conference on Machine Learning and Cybernetics, 2005, pp. 4260-4265 Vol. 7.

[4] S. Kumar, S. Mishra and S. Gupta, "Short Term Load Forecasting Using ANN and Multiple Linear Regression," 2016 Second International Conference on Computational Intelligence & Communication Technology (CICT), 2016, pp. 184-186.

[5] S. Luo, Y. Rao, J. Chen, H. Wang and Z. Wang, "Short-Term Load Forecasting Model of Distribution Transformer Based on CNN and LSTM," 2020 IEEE International Conference on High Voltage Engineering and Application (ICHVE), 2020, pp. 1-4.

[6] P. Yi, Z. Jianyong, Y. Yun, Z. Rui, Z. Cheng and S. Tian, "An Electricity Load Forecasting Approach Combining DBN-Based Deep Neural Network and NAR Model for the Integrated Energy Systems," 2019 IEEE International Conference on Big Data and Smart Computing (BigComp), 2019, pp. 1-4.

[7] Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets,International Conference on Neural Information Processing Systems, Kuching,Malaysia, 2014, pp. 2672-2680.

[8] J. Goodfellow, Y. Bengio and A. Courville, Deep Learning., MIT Press, 2016.

[9] A. M. Tudose, D. O. Sidea, I. I. Picioroaga, V. A. Boicea and C. Bulac, "A CNN Based Model for Short-Term Load Forecasting: A Real Case Study on the Romanian Power System," 2020 55th International Universities Power Engineering Conference (UPEC), 2020, pp. 1-6, doi: 10.1109/UPEC49904.2020.9209768.

[10] Benesty J., Chen J., Huang Y., Cohen I. (2009) Pearson Correlation Coefficient. In: Noise Reduction in Speech Processing. Springer Topics in Signal Processing, vol 2. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-00296-0_5

[11] Zhang J, Wei Y, Tan Z. An adaptive hybrid model for short term wind speed forecasting. Energy 2020. https://doi.org/10.1016/j.energy.2019.06.132.