SOLIMAN AND CHRISTENSEN ALGORITHM TO DETERMINE THE WEIGHT OF THE COMBINATION LOAD FORECASTING MODEL

THUẬT TOÁN SOLIMAN VÀ CHRISTENSEN ĐỂ XÁC ĐỊNH THAM SỐ CỦA MÔ HÌNH KẾT HỢP DỰ BÁO PHỤ TẢI

Ngo Viet Cuong [1], Nguyen Ngoc Tuan[1,*], Duong Ngoc Khang[1], Tran Xuan Tinh[2]

ABSTRACT

Ultra-short-term load forecasting is one of the core techniques for distribution networks operation. Actually, combined forecasting solution is usually used with strong randomness of loads. So how to weight the different forecasting model is a challenge. In such cases, Soliman and Christensen algorithm is used to determine the weight of the combined load forecasting model in this paper. This algorithm has been tested for a real distribution networks and the results show this algorithm has good prediction precision.

Keywords: *Soliman and Christensen algorithm, short term load forecasting, combination forecasting model.*

TÓM TẮT

Dự báo phụ tải siêu ngắn hạn là một trong những kỹ thuật cốt lõi để vận hành lưới điện phân phối. Trên thực tế giải pháp dự báo kết hợp thường được sử dụng với phụ tải mang đặc tính ngẫu nhiên cao. Vì vậy, làm thế nào để cân nhắc các mô hình dự báo khác nhau trong mô hình dự báo kết hợp là một thách thức. Bài báo trình bày phương pháp xác định tham số trong mô hình kết hợp dự báo phụ tải bằng thuật toán Soliman and Christensen. Thuật toán này đã được thử nghiệm cho lưới điện phân phối thực tế và kết quả cho thấy thuật toán này có độ chính xác dự đoán tốt

Từ khóa: *Thuật toán Soliman and Christensen, dự báo phụ tải ngắn hạn, mô hình dự báo kết hợp.*

[1]Faculty of Control Engineering, Le Quy Don Technical University
[2]Faculty of Fundamental technics, Air Defence - Air Force Academy
[*]Email: ngoctuanhvhn@gmail.com
Received: 20/9/2021
Revised: 02/10/2021
Accepted: 15/11/2021

1. INTRODUCTION

Renewable energy will make the distribution network automation system have higher requirements. In the actual operating state of the power system, a modern dispatch automation system should have two major characteristics: real-time and accuracy. However, due to the insufficient configuration of real-time measurement devices in the distribution network, it has severely restricted the realization of the controllable goals of the distribution network [1]. In such case, we need to form pseudo-measurements through ultra-short-term forecasts to make up for missing data [2]. But the trend of distribution network load is influenced by many random factors, which leads to a very large load random fluctuation. Therefore, a single model cannot guarantee its prediction accuracy. The combination forecasting model, incorporated with various forecasting algorithms, can significantly improve the prediction precision [3, 4]. Such combination model overcomes the defects of the single prediction model, and brings more choices [5, 6].

At present, according the advantages of a single forecasting model, many combined forecasting schemes have been proposed, such as: Reference [7] uses a neural network that combines the two most popular algorithms (Boosting and Bagging) to improve load forecasting in New England, compared with ordinary fully connected networks, the combined network can achieve better accuracy; reference [8] established combined prediction model with Prophet and LSTM (Long short-term memory) model, and used the least square method to determine the weights of combined model . The results show that the combined prediction model can obtain higher prediction accuracy than a single prediction model. Moreover, considering that LSTM is one of the most widely used recurrent neural networks. It has a strong ability to model time series [9], but it has a poor effect when the feature is discontinuous, so, a series of combined prediction models based on the LSTM algorithm have been proposed: LSTM-RF-SVM (Long short-term memory-RF-Support vector machine) combined model [10], CNN-LSTM (Convolutional neural networks- Long short-term memory) hybrid neural network [11, 12], CEEMD (Complementary ensemble empirical mode decomposition) – IPSO (Improve particle swarm optimization) - LSTM combined model [13].

It can be seen that most of the existing methods are based on a core model to establish a combined forecasting models, the possible consequence of this random combination is that the combined prediction accuracy is lower than that of the core model. Therefor, reference [14] based on the Cooperative Game theory proposed a method for selecting the optimal combination. Although this method improves the accuracy of prediction in some special scenes, it still has a certain degree of subjectivity. In such cases, Soliman and Christensen (SC) algorithm is used to determine the weight of the combined forecasting model in this paper [15]. This method can not only overcome the subjectivity of reference [14], but also has the characteristics of identifying discrete data. The results show that this method has high accuracy.

2. SOLIMAN AND CHRISTENSEN ALIGORITHM FOR DETERMINING WEIGHTS OF COMBINATION FORECASTING

2.1. Combined forecasting model

The basic form of combination forecasting model can be formulated as:

$$\begin{cases} f = k_1 f_1 + k_2 f_2 + ...k_m f_m = \sum_{i=1}^{m} k_i f_i \\ \sum_{i=1}^{m} k_i = 1 \\ k_i \geq 0; i = 1,2,\ldots,m \end{cases} \quad (1)$$

Then, (1) can be written as:

$$\mathbf{F} = \mathbf{Hk} + \mathbf{v} \quad (2)$$

Where, **F** predictive value matrix, **k** weight matrix, **H** Jacobian matrix, **v** residuals matrix.

Obviously, (2) is overdetermined equations. Various scholars have given different solutions. The SC algorithm is used to solve the problem in this paper, and the solving steps are in 2.2 section.

2.2. SC (Soliman and Christensen) algorithm

According to the reference [15], the steps of the SC algorithm to solve the equations (2) are as follows.

Step 1: Use the following formula to calculate the weight of the model.

$$\mathbf{k} = [\mathbf{H}^T\mathbf{H}]^{-1}\mathbf{H}^T\mathbf{F} \quad (3)$$

Step 2: Calculate least error squares residuals from this solution

$$\mathbf{v} = \mathbf{F} - \mathbf{Hk} \quad (4)$$

Step 3: Calculate the standard deviation σ

Step 4: If $v_i < \sigma$, reject data values with residuals v_i, go to Step 1.

If all $v_i < \sigma$, go to Step 5.

Step 5: Select the n measurements that correspond to the smallest least error squares residual and form the corresponding $\mathbf{H}^*$, $\mathbf{F}^*$. Go to Step 6.

Step 6: Calculate the weight of the combined model

$$\mathbf{k}^* = [\mathbf{H}^*]^{-1}\mathbf{F}^* \quad (5)$$

If $k_i < 0$ ($k_i \in \mathbf{k}^*$), show that the i^th^ model is not suitable for combining with other models, remove this model and go to Step 1, recalculate.

Step 7: Use the weights of the combined model to make predictions.

The flow chart of this solution can be descripted in Figure 1. This flow chart is basically the same as the reference [14]. The only difference is that this paper uses the SC algorithm instead of the cooperative games theory detection process. In order to show the advantages of the SC algorithm, this paper further replaces the cooperative game method of step 4 with the SC algorithm (CG+SC algorithm). This method will be compared with the SC algorithm in 3.2 section.

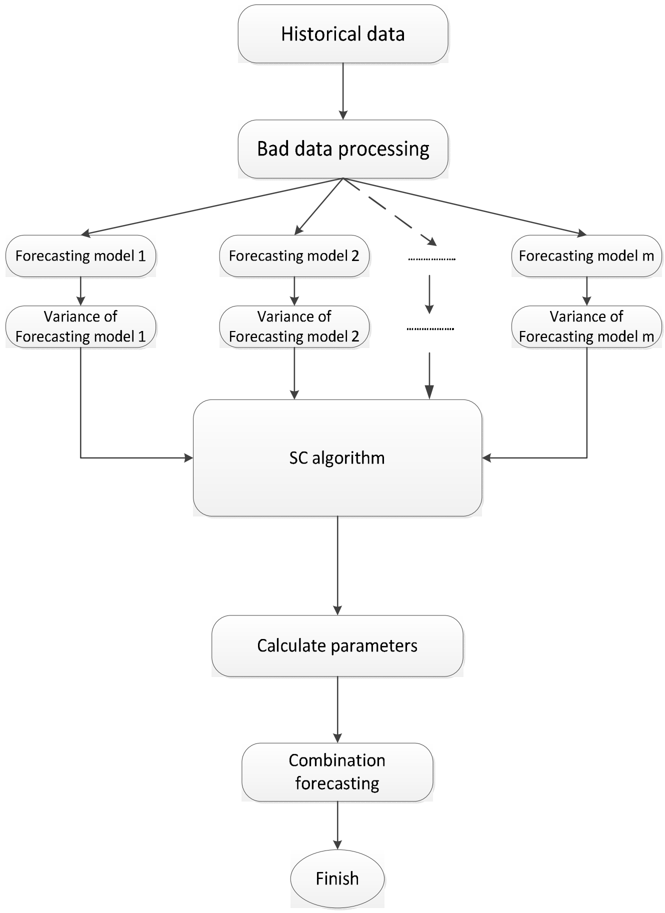

Figure 1. Experiments

3. RESULTS AND DISCUSSION

3.1. Load series

This paper adopts the load curve of reference [14]. The data series are shown fig.1 in reference [14]. The load data of distribution network changes significantly, and the data of some days are absent and abnormal. Generally, there are many ways to solve missing data, such as: average algorithm, interpolation algorithm, etc. In such case, these processed data may not meet the accuracy requirements, which may affect the accuracy of a single model and this data has a great influence on the weight of the combined forecasting model. Reference [14] also pointed out that this load will change influenced by various factors. In this case, the combined forecasting model is the best option. Considering that the method in the reference [14] will be subjective judgment to these changes and affect the influence of the best combination choice. Therefore, this paper will further discuss.

3.2. Combination forecast selection

Here this paper uses the four methods of reference [14] to combine: ES method, TS method, Naïve method, H-W method. The best combination of SC algorithm and Cooperative game theory is shown in Table 1 (0: rejected, 1: selected). It can be seen from the Table 1 that the SC algorithm rejects the ES method, while the method in [14] rejects the TS method and Naive method.

This paper compares predictions accuracy for that two case. For specific comparison index, please refer to 3.3 section.

Table 1. The best combination with various algorithms

Method	ES	TS	Naive	H-W
SC algorithm	0	1	1	1
Cooperative game theory [14]	1	0	0	1

3.3. Comparison index

This paper adopts the commonly used in ultra-short-term power load forecasting evaluation index: Root man squared error (RMSE), Mean absolute percentage error (MAPE). The calculation formula as follows:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{t=1}^{N}(y_t - f_t)}{N}}$$

$$\mathrm{MAPE} = \sum_{t=1}^{N}\left|\frac{y_t - f_t}{y_t}\right| \times \frac{100}{N}.$$

3.4. Combination model performance

The prediction accuracy of SC algorithm on two case is shown in Table 2. It can be seen from the table that the MAPE of CG+SC is reduced by 0.27% compared with the reference [14], moreover, the MAPE of the SC algorithm is 0.49% lower than the reference [14]. This shows that the SC algorithm has high prediction accuracy. Moreover, when the SC algorithm uses the check of CG theory to select the best combination, the MAPE of its solution is 0.22% higher than the SC algorithm itself. This is because the check of CG theory is subjective, so its method believes that the naive method does not contribute to the combination, but in fact its method has its own advantages in combination. In addition, the Table 2 also shows that the RMSE of the SC algorithm is the best among various prediction methods.

In order to further prove the superiority of the SC algorithm, this article gives the load curve of the SC algorithm on a certain day (Figure 2). It can be seen that the SC algorithm is basically consistent with the real curve.

Figure 3 shows the error distribution of the SC algorithm. From the error distribution curve, that the error distribution of the SC algorithm is about 4%, indicating that the accuracy of this method meets the requirements of the actual operation of the power grid.

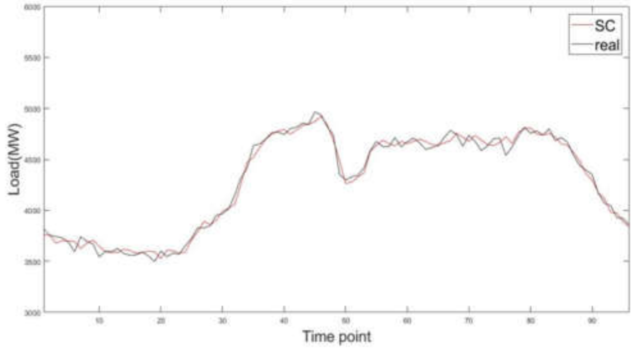

Figure 2. Load forecasting curve of SC algorithm

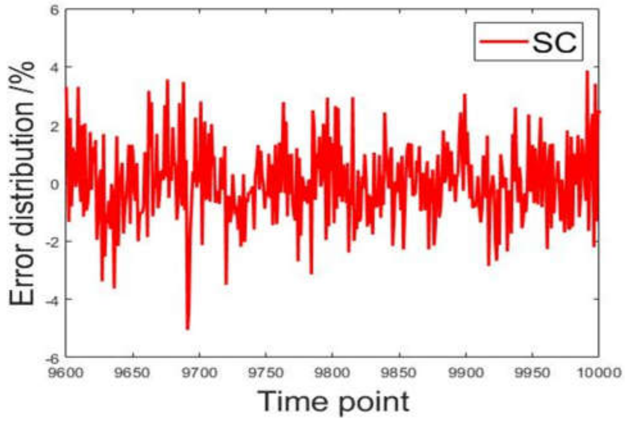

Figure 3. Error distribution of the SC algorithm

Table 2. Comparison of various forecasting methods

Method	RMSE	MAPE
Naive	223.08	3.4442
ES	84.69	1.4422
TS	84.90	1.4454
H-W	59.72	1.0133
[14] method	58.70	1.0183
SC with cooperative game theory	58.53	1.0155
SC	58.49	1.0133

4. CONCLUSION

SC algorithm is used to calculate the combination forecast weight in this paper. This algorithm has certain advantages in selecting the best combination. The results show that the SC algorithm can avoid the subjectivity of the reference [14] method model. In addition, no matter prediction accuracy or error distribution control, the SC algorithm can meet the requirements of actual system operation. Therefore, this algorithm is expected to be applied in the actual system.

REFERENCES

[1]. Primadianto A., Lu C N., 2017. *A review on distribution system state estimation*. IEEE Trans. Power Syst, 32/5, 3875-3883.

[2]. Sinha A K., Mandal J K.,1999. *Dynamic state estimation using ANN based bus load prediction*. IEEE Trans. Power Syst, 14/11/, 1219-1225.

[3]. Bate J.M., Granger C.W.J., 1969. *The Combination of Forecasts.* Operational research Quarterly, 20/4, 451-468.

[4]. Robert L. Winkler, Spyros Makridakis, 1983. *The Combination of Forecasts*. Journal of the Royal Statistical Society. Series A (General), 146/2, 150–157.

[5]. Thordarson F O., Madsen H., Nielsen H A., et al., 2010. *Conditional weighted combination of wind power forecasts*. Wind Energy, 13/8, 751-763.

[6]. Meng FS., Li MY., 2013. *Application of combined model in energy demand prediction*. J Harbin Inst Technol, 45/11, 106–111.

[7]. A.S. Khwaja, X. Zhang, A. Anpalagan, et al., 2017. *Boosted neural networks for improved short-term electric load forecasting*. Electric Power Systems Research, 143, 431-437.

[8]. Peng Pai, Liu Min, 2021. *Short-term load forecasting method based on Prophet-LSTM combined model*. Proceedings of the CSU-EPSA, China.

[9]. Zhang Yuhang, Qiu Caiming, He Xing, Ling Zenan, Shi Xin, 2017. *A short-term load forecasting method based on LSTM neural network*. Power information and communication technology, 15/09, 19-25.

[10]. Wei Wei, Jing Huang,2020. *Short-term load forecasting based on LSTM-RF-SVM combined model*. Journal of Physics: Conference Series.

[11]. Chuan Shang, Junwei Gao, Huabo Liu, Fuzheng Liu, 2021. *Short-term Load Forecasting Based on PSO-KFCM Daily Load Curve Clustering and CNN-LSTM Mode*l. IEEE Access, 9.

[12]. Xifeng Guo, Qiannan Zhao, el al., 2020. *A short-term load forecasting model of multi-scale CNN-LSTM hybrid neural network considering the real-time electricity price*. Energy Reports, 6, 1046-1053.

[13]. Yuan Pei, Lei Zhenglin, Zeng Qinghui, et al., 2021. *Load forecasting of refrigerated display cabinet based on CEEMD–IPSO–LSTM combined model*. Open Physics, 19, 360–374.

[14]. Viet Cuong Ngo, Wenchuan Wu, et al., 2017. *Cooperative game-based method to determine the weights of load forecasting solution incorporated with various algorithms*. The Journal of Engineering, November.

[15]. S.A. Soliman, Ahmad Mohammad Al-Kandari, 2010. *Modeling and Model Construction-Modeling and Model Construction*. Butterworth-Heinemann, Book, April.

THÔNG TIN TÁC GIẢ

Ngô Việt Cường[1], Nguyễn Ngọc Tuấn[1], Dương Ngọc Khang[1], Trần Xuân Tình[2]

[1]Khoa Kỹ thuật điều khiển, Học viện Kỹ thuật Quân sự

[2]Khoa Kỹ thuật cơ sở, Học viện Phòng không - Không quân