Cooperative game-based method to determine the weights of load forecasting solution incorporated with various algorithms

VietCuong Ngo¹, Wenchuan Wu¹, Yi Yang², Boming Zhang¹

¹Department of Electrical Engineering, Tsinghua University, Beijing, People’s Republic of China
²State Grid Jiangsu Electric Power Research Institute, Nanjing, People’s Republic of China

E-mail: langtu_cvp@yahoo.com

Published in The Journal of Engineering; Received on 9th October 2017; Accepted on 2nd November 2017

Abstract: Ultra-short-term load forecasting is one of the core techniques for distribution networks operation. Since individual load in distribution networks is small and with strong randomness, single prediction model cannot guarantee the forecasting accuracy and combined forecasting solution is usually used. In such forecasting scheme, how to weight the different forecasting model is a challenge. By introducing the concept of cooperative game, the weight of every algorithm can be determined according to the error distribution. The proposed method has been tested for a real distribution networks and the results show this scheme has good prediction precision, and can control the error distribution at the inflection points.

1 Introduction

In real practice, ultra-short-term load forecasting is one of the main sources of pseudo-measurement for distribution state estimation (DSE) [1, 2]. In distribution networks, the measurements are mainly from automatic meter reading system (AMR). However, the failures or time delay of AMR may lead to data missing. Therefore, load forecasting is needed to generate pseudo-measurement to mend these missing data for DSE. The accuracy of pseudo-measurement will significantly affect the results of DSE.

Till now, many practical prediction models have been proposed. For the published works, most predictive models are single prediction methods, and the core idea is the process of making predictions of the future based on past and present data and most commonly by analysis of trends [3, 4]. However, 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 [5]. Such combination model overcomes the defects of the single prediction model, and brings more choices [6, 7].

Since the weights of different algorithms in the combination prediction model influence the prediction accuracy, how to optimise these parameter is a key problem [8, 9].

Recently, many solutions were proposed to determine the weight of combination forecasting [10–16]: equal weighting method [10], Bayesian estimation [11], regression method [12], joint adaptive training [13], rough sets theory [14], OWA or IOWA operation, and so forth [15, 16].

These methods can be roughly divided into two categories: the first one is to formulate a optimisation problem with certain constraints, so as to obtain the weighting coefficient of the combined forecasting method; for the other idea, the weight coefficient of each individual forecasting model is selected inversely proportional to its variance of the prediction error. Since different time series have different dimensions, the sum of squares of the absolute errors is not comparable, so the criterion of the sum of squares of errors may not accurately reflect the validity of the prediction method, the results of the first class of weight solving methods may not be optimal in this situation. Furthermore, the computation burden of this method is very heavy. For the second type schemes, a simple weight calculation formula cannot guarantee the weight positive.

In this paper, the game theory is used to determine the weights of various prediction algorithms in the combination model. Each individual forecasting model is regarded as the cooperative player in the combination forecasting model, the sum of square error of combination forecasting is considered as a total payoff, then the weights to each single forecasting model are allocated using the Shapley value method. The forecasting results for a real distribution network show that this method has good prediction accuracy and can control the error distribution.

The remainder of this paper is organised as follows. In Section 2, the characteristics of a Chinese city’s load series are explained. Section 3 introduces short-term load forecasting using combination forecasting model. Section 4 introduces the Shapley value. Section 5 presents a cooperative game method to determine the weights of combination forecasting model. The numerical test results are given in Section 6, and follows conclusions in Section 7.

2 Load series

In this section, the load series of a China’s urban distribution networks is analysed (named city A load series). The load series consist of all 15 min observations in 2006, the data of the first 3 months was used to determine the model parameters, and the remaining months of data is used to evaluate the post-sample forecast accuracy. The data series are shown in Fig. 1, the load data of distribution network changes significantly, and the data of some days are absent and abnormal. This situation will affect the prediction accuracy and error distribution, and it is the main problem to be solved in this paper.

Fig. 2 shows the load profile data of a typical week in city A. This figure emphasises the differences between the load patterns of different days within a week. Fig. 3 shows the load series of typical week and weekend days in city A. These two series both possess strong information seasonality with lower demand in the weekend.
A combination model, incorporated with different forecasting methods in an appropriate way, can provide comprehensive information and improve the prediction accuracy. The basic form of combination forecasting model can be formulated as

$$f = k_1f_1 + k_2f_2 + \ldots + k_m f_m = \sum_{i=1}^{m} k_i f_i$$ (1)

where $k_i$ is the weight of the $i$th prediction algorithm, $f_i$ the $i$th predictive value, and $m$ the number of prediction algorithms. The formula to optimise the weights is

$$\min w = \min \sum_{i=1}^{n} (y_i - f_i)^2$$

$$= \min \sum_{i=1}^{n} (y_i - \sum_{i=1}^{m} k_i f_i)^2$$ (2)

The above model is a quadratic problem, and the existing method can find the optimal solution, but it cannot control the error distribution. For example, we should guarantee the prediction accuracy of the inflection points instead of minimising the overall sum square of the prediction errors.

### 4 Shapley value

In the combination forecast solution, the single forecasting model set is denoted by $N = \{1,2,\ldots,n\}$. Then each single forecasting model can be considered as a player, which compose a set of game players in combination forecasting. Set $2^m$ of all coalitions (subsets of $N$), these subsets form grant coalition. The coalitional form of an $n$-person game is given by the pair $G = (N,v)$, where $v$ is the revenue function, called the characteristic function of the game.

Let $v(i)$ be the characteristic function of each model, and $v(N)$ be the characteristic function of the combined forecasting. It has the following characteristics:

- standardisation: $v(\emptyset) = 0$;
- non-negative: $v(C) \geq 0 \forall C \subseteq V$;
- monotonicity: $v(C) \geq v(N)$;
- super additive: $v(C) + v(D) \geq v(C \cup D) \cup C, D \subseteq N$.

Where $v(C)$ denotes the variance of the combined forecasting from the coalitional of $C$.

We use the Shapley value (in the coalition, the proceeds to which the $i$th participant should be assigned):

$$\varphi_i(v) = \frac{1}{n!} \sum_{C \subseteq N, |C| = i} (|C|! - |C|)! (v(N) - v(C) - v(C \setminus \{i\}))$$ (3)

An allocation is individually rational and efficient if $\sum_{i=1}^{m} \varphi_i(v) = v(N), \varphi_i(v) \leq v(i)$ for all $i$.

The meaning of (3) is all participants join in the alliance in sequence, and the possible sequences of alliance $S$ have $|S|$! possibilities; then, $i$ joins in the alliance, and brings an extra income of $v(SU\{i\}) - v(S)$; the other $(N - |S| - 1)$ participants can join in the alliance in $(N - |S| - 1)!$ different sequences, and hence, the alliance totally has $|S|!(N - |S| - 1)!$ different sequences. Therefore, $|S|!(N - |S| - 1)!$ in (3) denotes the weight of contribution of $i$ in $S\rightarrow S\cup\{i\}\rightarrow N$; the possible sequences of the final alliance are $N!$, and we can average the contribution of all $i$ to formulate (3).
5 Cooperative game-based method for determining weights of combination forecasting

The combination forecasting model combines the results of different models, and the essence of the integrated model is the weight of each single method in the final prediction of the proportion of the results. Then if there are some single methods in the combination no contribution in the combined model, we can remove it first. So that model does not participate in the process of combination forecasting, which is the formation of the basic idea of forecasting.

Case 1: if $\varphi_i(v) \approx \varphi_i(v)((\varphi_i(v) - \varphi_i(v))/\sqrt{N}) < \varepsilon$, then the contribution of these two models is the same, one of which can be removed.

Case 2: if $\varphi_i(v) > \sqrt{N}$, then this model causes loss in combination forecasting, it is removed.

The determination strategy of weights for combination forecasting model is:

Step 1: Assume that the absolute error produced by each single forecasting model at time $t$ is $e_i(t) = y - f_i$. We calculate the variance of a single model:

$$v(|i|) = \text{Var}(e_i(t)) \quad (4)$$

Step 2: Assume that the total return in the combined forecast is:

$$v(N) = \frac{1}{m} \sum_{t=1}^{m} \text{Var}(e_i(t)) \quad (5)$$

Step 3: Applying the Shapley value assign to each forecasting algorithm, it can fairly and fully reflect the contribution of each prediction model to the total prediction:

$$\varphi_i(v) = \frac{1}{m} \sum_{C \subseteq S, i \in C} ((|C| - 1)!/(|C|)(v(C) - v(C\{i\})) \quad (6)$$

Step 4: Calculate the weight of combination forecasting (we used variance analysis method [17]):

$$k_i = \frac{1}{\varphi_i(v)} \sum_{j=1}^{m} (1/\varphi_j(v)) \quad (7)$$

Obviously, $k_i$ satisfy condition (2).

The flowchart of this solution can be described in Fig. 4:

6 Numerical test

6.1 Performance of single prediction model

Exponential smoothing (ES) and trend extrapolation methods (TE) are widely used in ultra-short-term load forecasting. In this paper, we mainly focus on these two methods and naive method [18], and now very efficient Holt–Winter [19] model to verify the practicality of this paper method. The weights of these method are limited between 0 and 1. In order to calculate the parameters, we use OPTI package – non-linear least square algorithm method. To compare the improvement of combination solution, the accuracy of each single prediction model is listed in Table 1.

Table 1 Forecasting results of a single forecasting model

| Forecasting method | Average error (mape) | Error distribution % |
|--------------------|---------------------|----------------------|
|                   | Error < 2% | Error < 4% | Error < 5% |
| naive method       | 3.4442    | 40.56     | 56.04     | 77.30     |
| trend extrapolation| 1.4454    | 73.79     | 89.80     | 98.86     |
| exponential smoothing| 1.4422  | 73.89     | 89.87     | 98.88     |
| H–W method         | 1.0363    | 87.98     | 97.72     | 99.90     |

Table 2 Prediction results of the combination forecasting model with different weight determination methods

| Forecasting method | Average error (mape) | Error distribution % |
|--------------------|---------------------|----------------------|
|                   | Error < 2% | Error < 4% | Error < 5% |
| variance analysis  | 1.1012    | 85.19     | 96.52     | 99.88     |
| NLS                | 1.0158    | 88.37     | 97.75     | 99.92     |
| proposed method    | 1.0183    | 88.24     | 97.80     | 99.94     |

6.2 Comparison of the weight determining methods

Besides the proposed method, variance analysis method [17] and NLS method of OPTI Software Package are also used to solve the weights from comparison.

Table 2 shows that the error distribution of the proposed method is smaller than the traditional method, so it has the best prediction performance. Although the average error of this method is bigger than that of the NLS method, but its error distribution is better than the NLS method. This is because the NLS method is to minimise the sum square of error, but the proposed method is to control error distribution.

From Table 3, we can see accuracy of the proposed method is significantly better than the other two methods in holidays. Fig. 5
shows that most of the blue lines are covered by red lines and black lines; this indicates error distribution of this paper is significantly better than the other two methods. Furthermore, Fig. 6 shows the errors at the inflection points for these two methods. In Fig. 6, it can be seen that the proposed method has a much high precision than the NLS method at the inflection points. This is the most important advantage of the proposed method.

Assuming the prediction error distributions are subject to a normal distribution, Table 4 lists the interval width of the predictions results under a given confidence level of 90%. It shows that NLS and proposed method have the smallest width which indicates that the combination model has the best performance. NLS has smaller width than the proposed method, but the proposed method has high prediction precision at the inflection points. The large prediction error at the inflection points may lead to more conservative dispatch strategies. Therefore, the proposed method is more useful for real practice.

### Table 3 Comparison accuracy for holidays

| Forecasting method | Average error (mape) | Error distribution % |
|-------------------|---------------------|----------------------|
|                   |                     | Error < 2% | Error < 4% | Error < 5% |
| variance analysis method | 1.2101 | 82.60 | 99.17 | 100 |
| NLS               | 1.1769 | 83.85 | 99.17 | 100 |
| proposed method   | 1.1718 | 83.23 | 99.27 | 100 |

### Table 4 Interval width of the prediction results under a confidence level of 90%

| Index | ES method | TE method | Naïve method | H-W method | Variance analysis method | NLS proposed method |
|-------|-----------|-----------|--------------|------------|-------------------------|---------------------|
| interval width | 138.18 | 138.48 | 344.29 | 99.30 | 105.23 | 97.25 | 97.51 |

### 7 Conclusions

This paper proposes a weight determination method for combined forecasting model in a distribution network. This method can improve the prediction accuracy and control error distribution. The proposed method is especially suitable to forecast the loads of small values and strong randomness. Moreover, it usually has a good accuracy at the inflection points where many other published methods may produce notable forecasting errors.

### 8 Acknowledgments

This work was supported in part by the National Key Research and Development Plan of China (grant 2016YFB0900400), in part by the State Grid Jiangsu Electric Power Company under grant number J2016015 (Research on friendly coordination technology of large-scale customers and active distribution networks).

### 9 References

[1] Sinha A.K., Mandal J.K.: ‘Dynamic state estimation using ANN based bus load prediction’, *IEEE Trans. Power Syst.*, 1999, 14 (11), pp. 1219–1225.
[2] Zhinong W., Tieming X., Guoqiang S.: ‘Linear dynamic state estimation based on mixed measurements using ultra-short term load prediction’, *Proc. CSEE*, 2010, 30 (1), pp. 47–51.
[3] Liu K., Subbarayan S., Shoultz R.R., et al.: ‘Comparison of very short-term load forecasting’, *IEEE Trans. Power Syst.*, 1996, 11, pp. 877–882.
[4] Srivastava A.K., Pandey A.S., Singh D.: ‘Short-term load forecasting methods: a review’. 2016 Int. Conf. on Emerging Trends in Electrical Electronics & Sustainable Energy Systems (ICETESES), 2016, pp. 130–138.
[5] Bate J.M., Granger C.W.J.: ‘The combination of forecasts’, *Oper. Res. Q.*, 1969, 20 (4), pp. 451–468.
[6] Thordarson F.O., Madsen H., Nielsen H.A., et al.: ‘Conditional weighted combination of wind power forecasts’, *Wind Energy*, 2010, 13 (8), pp. 751–763.
[7] Germi M.B., Mirjavadi M., Namin A.S.S., et al.: ‘A hybrid model for daily peak load power forecasting based on SAMBA and neural network’, *J. Intel. Fuzzy. Syst.*, 2014, 27 (2), pp. 913–920.
[8] Huang C.-M., Huang C.-J., Wang M.-L.: ‘A particle swarm optimization to identifying the ARMAX model for short-term load forecasting’, *IEEE Trans. Power Syst.*, 2005, 20 (2), pp. 1126–1133.
[9] Xu C., Chongqing K., Qing X., et al.: ‘Integrated model of short-term load forecasting’, *Autom. Electr. Power Syst.*, 2000, 24 (9), pp. 42–44.
[10] Wang L., Libert G., Liu B.: ‘Combining forecasts using recursive equal weighting and linear programming’, *Proc. of the 31st IEEE Conf. on Decision and Control*, 1992, pp. 3705–3706.
[11] Kuhner M.K.: ‘LAMARC 2.0: maximum likelihood and Bayesian estimation of population parameters’, *Bioinformatics*, 2006, 22 (6), pp. 768–770.
[12] Zhang F., Zhou X.: ‘Gray-regression variable weight combination model for load forecasting, IRCRMEM 08’. Int. Conf. on Risk Management & Engineering Management, 2008, no. 2, pp. 311–316.
[13] Bao Y., Huang M., Zhu Z., et al.: ‘Application of combination forecasting based on rough sets theory on electric power system’. Proc. of the 6th World Congress on Intelligent Control and Automation, June 21–23, 2006, Dalian, China, pp. 1745–1748.
[14] Ye L., Chongqing K., Xinyu C.: ‘Joint adaptive training of parameters in integrated models and related single forecasting methods’, *Autom. Electr. Power Syst. China*, 2010, 34 (22), pp. 36–40.
[15] Li F., Qian J., Yan Q., et al.: ‘A combination forecasting model based on IOWA operator for PV generation’. Int. Conf. on Renewable Power Generation (RPG 2015), pp. 1–4.
[16] Kaur G., Dhar J., Guha R.K.: ‘Minimal variability OWA operator combining ANFIS and fuzzy c-means for forecasting BSE index’, *Math. Comput. Simul.*, 2016, 122, (122), pp. 69–80.
[17] Chongqing K., Qing X.: ‘Power systems load forecasting’ (China Electric Power Press, 2007).
[18] Hong T.: ‘Short term electric load forecasting’. PhD dissertation, North Carolina State University, 2010.
[19] Taylor J.W.: ‘An evaluation of methods for very short-term load forecasting using minute-by-minute British data’, *Int. J. Forecast.*, 2008, 24, pp. 645–658.