Peak Load Forecasting in Summer Using Artificial Neural Network

Sung-jun Baek*, Ye-Ji Son, and Sung-guk Yoon†**

Abstract – To reliably operate power systems, not only hourly load forecasting but also peak load forecasting are required. Because of the heat, ventilation, and air conditioning (HVAC) loads, the annual peak load is generally occurred in summer or winter season, so the peak load heavily depends on the temperature. In this work, we propose an ANN-based peak load forecasting method in summer using ensemble average. The proposed method uses three different ANN structures which use different input parameters among the previous years' peak loads in summer, temperature, and gross domestic product (GDP), and the final result is obtained using the ensemble average of the three ANNs' results. Through the case study, it is confirmed that the proposed method shows good forecasting performance for the summer peak load.

Keywords: Artificial neural network (ANN) , Ensemble method, Load forecasting, Summer peak load

1. Introduction

The power system requires balancing between demand and supply in every instance. Therefore, load forecasting is the essential part of power system operation. In addition, load forecasting is used for dispatch control and load management. After a blackout in Korea in September 2011 [1], more accurate peak load forecasting was required for reliable operation of the power system. However, the peak load forecasting has become more difficult due to the increasing summer temperature caused by global warming and changing in electric rates [2].

The artificial neural network is one good technique in load forecasting area. It generates outputs through activation function and weights using several input variables based on historical data. Learning proceeds to minimize the error between the outputs of the artificial neural networks (ANN) and the target values [3].

Recent researches in peak load forecasting have investigated using machine learning techniques. Amin-Naseri et al. have proposed a hybrid neural network model for daily peak load forecasting by using unsupervised learning method [4]. In [5], the authors have proposed a daily peak load forecasting scheme using ANN which has hidden units with connecting weights between only one group of related input units. To forecast hourly peak loads, Kim et al. employed a Kohonen neural network model for clustering weekdays [6]. In [7], a neuro-evolutionary technique known as Cartesian Genetic Algorithm evolved Artificial Neural Network (CGPANN) has been deployed to forecast a day-ahead peak load.

In this paper, we propose a peak load forecasting scheme in summer using ANNs. Since the summer peak load is highly correlated with temperature, a representative temperature is firstly calculated using temperature data from top eight largest cities in Korea. The representative temperature is obtained by weighted average with respect population. To improve the forecasting performance, we design three ANN structures and the final forecasting result is the ensemble average of the three ANNs' result. Through case studies, i.e., summer peak load forecasting in Korea from 2013 to 2017, the forecasting performance and reliability of the proposed method is verified. The proposed forecasting scheme generally shows better forecast performance than single-ANN based forecasting schemes.

The rest of the paper is organized as follows. We analyze the peak load data of Korea in Section 2. Section 3 describes the proposed forecasting method including representative temperature. After evaluating the proposed method in Section 4, we conclude our paper in Section 5.

2. Data analysis

The daily summer peak load in Korea provided by Korea power exchange (KPX) is used in this paper.

The KPX publicly released full data sets of anonymized full data sets of the past few years. Data was collected since Jan 1, 2005, with installed capacity, availability, Peak Load, Supply capacity and Reserve [8]. Before devising a forecasting method, we analyzed the provided data set. Previous researches on the load forecasting have shown that electric load and temperature are highly correlated [9], [10]. The maximum power consumption in South Korea typically occurs in January or February in the winter months and in

† Corresponding Author : Dept. of Electrical Engineering, Soongsil University, Korea (sgyoon@ssu.ac.kr)
* Dept. of Electrical Engineering, Soongsil University, Korea (sjbaek@soongsil.ac.kr)
** Dept. of Electrical Engineering, Soongsil University, Korea (suonj1024@naver.com)

July or August in the summer months.

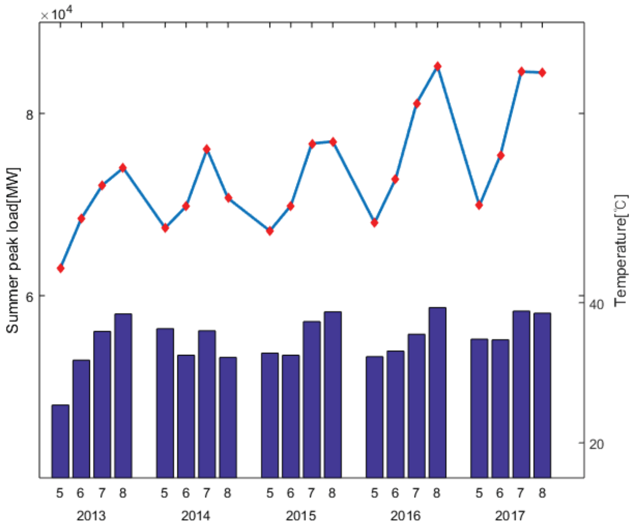

Fig. 1. Summer peak loads and maximum temperatures on the same day in 2013 – 2017. The numbers 5, 6, 7, and 8 represent May, June, July, and August, respectively

Fig. 1. shows peak loads in each month and the maximum temperature on the same day in summer from 2013 to 2017. As shown in Fig. 1, the summer peak load in a year occurs in July or August, which are the hottest season in Korea. The two data are proportional to each other, indicating that temperature and power are closely related in summer. Exceptionally, in May 2014, although the maximum temperature is very high, the peak load in May is low because HVAC is not fully working in May. Therefore, in this paper, temperature data are added as input data for summer peak power forecasting. The influence of temperature is generally high in general loads and household loads. On the contrary impact of industrial loads is low [11].

In Section 3.C, the representative temperature calculation method is described.

3. Forecasting methods

In this section, we describe the theoretical background of the ANN and propose an ANN-based summer peak load forecasting scheme.

A. Artificial Neural Network (ANN)

The ANN is an algorithm designed to imitate the human brain and to train the machine with various data. In ANN, a perceptron receives inputs from its child perceptron, and then the perceptron also sends its data to the next perceptron by using the activation function. Typical activation functions include step function, sigmoid function and ReLU function. In this paper, we choose the activation function as the sigmoid function. Multi-layer perceptron (MLP) has several hidden layers to solve nonlinear problems. To train the MLP, a backpropagation algorithm is used which adjusts weights until the error reached a threshold. The error is defined as the difference between the output of MLP and the target value. The output value obtained by the feedforward algorithm is calculated as

$$y = \varphi\left(\sum_{k=1}^{n} w_k x_k + b\right) \tag{1}$$

where x_k, w_k, and b, are the input of neuron k, the weight of neuron k, and the bias, respectively. φ is the activation function and it is a sigmoid function in this work. The error is defined by the difference between the output value y and the target value t, i is an index of output and target it is given by

$$E(w) = \frac{1}{N}\sum_{i=1}^{N} |y(x_i, w) - t_i|^2 \tag{2}$$

Training the ANN means that adjusting the weights to reduce $E(w)$. To this end, we use the backpropagation algorithm based on the gradient descent method. That is,

$$w(t+1) = w(t) - \eta\frac{dE}{dw} \tag{3}$$

η is learning rate, in this work η is 0.01.

B. Ensemble method

The load forecasting is a chaotic work. Small errors in the initial conditions of a forecast grow rapidly and affect accuracy. Ensemble methods combine the forecasting result of several base estimators built with a given learning algorithm in order to improve generalizability/robustness over a single estimator. One of ensemble method is averaging, where the main idea is to build several forecasting models independently and then to average their results. Examples of the ensemble average are the bagging methods and RF [12].

C. Representative Temperature

We choose the top largest eight cities in Korea based on population to create one representative temperature. Note that we choose the eight cities in terms of a population rather than load since the industrial load is not much sensitive temperature. The temperature weight formula for each city is

$$TW(i,y) = \frac{L(i,y)}{\sum_{i=1}^{m} L(i,y)} \tag{4}$$

where TW(i,y), m, and $L(i,y)$ denote the temperature weight of city i and year y, the total number of cities[1], and the sum of general load and housing load in the summer season (May, Jun, Jul, and Aug) of city i and year y.

[1] In this paper, m is 8.

After calculating temperature weights in each city. The city temperature is obtained by multiplying weights by the average, maximum and minimum daily temperatures of each city. Finally, the representative temperature is the sum of the city temperature. That is,

$$RT(i,d,y) = \sum_{i=1}^{m} TW(i,y) * T(i,d,y) \qquad (5)$$

where RT(i,d,y) and T(i,d,y) are the representative temperature and real temperature data of city i, day d, and year y.

C. Peak Forecasting model `

If we use a single ANN-based peak load forecasting scheme, the result might have a difficulty of overfitting problem resulting in poor forecast performance. In this case, the reliability of peak load forecasting is greatly reduced. In this paper, we use a bagging method (Bootstrap aggregating) among the ensemble methods described above, to improve robustness. The proposed method uses three ANN cases by using different period and each case creates four forecasting models by using different input parameters. The final forecasting result for target year is obtained by ensemble average of 12 forecasting results.

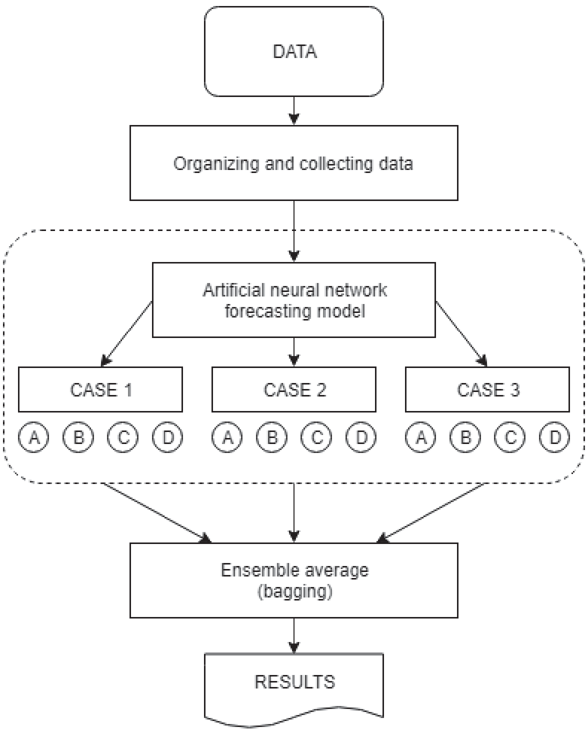

Fig. 2. Forecasting structure and algorithm

As shown in Fig. 2. we construct the learning model in three cases and each case has four forecasting models. The structure of the ANN is shown in Fig. 2. There is one hidden layer and the number of hidden neurons is about 1/3 of the sum of input and output neurons. It was obtained by empirically. In addition, the Levenberg-Marquardt approach was used to train the model.

Case 1: Using May, June data, to forecast July, August peak loads.

Input data -

- A: Weekly peak load data in May, June;
- B: Daily peak load data in May, June;
- C: Weekly peak load data in May, June; Temperature data for May, June (Avg, Min, Max);
- D: Daily peak load data in May, June; Temperature data for May, June (Avg, Min, Max);

Case 2: Using June, July data, to forecast August peak loads.

Input data -

- A: Weekly peak load data in June, July;
- B: Daily peak load data in June, July;
- C: Weekly peak load data in June, July; Temperature data for June, July (Avg, Min, Max);
- D: Daily peak load data in June, July; Temperature data for June, July (Avg, Min, Max);

Case 3: Using past July, August data, to forecast following July, August peak loads.

Input data -

- A: Weekly peak load data in July, August;
- B: Daily peak load data in July, August;
- C: Weekly peak load data in July, August; Temperature data for July, August (Avg, Min, Max); GDP;
- D: Daily peak load data in July, August; Temperature data for July, August (Avg, Min, Max); GDP;

The temperature in here means the representative temperature described in Section 3.C. Case 1 is an example of how to use the data for May and June to forecast the July and August peaks of the same year. Model A used weekly peak load and Model B used the daily peak load. The model C and D are added representative temperature data to model A and B. Case 2 forecasting the peak in August with data for June and July. Model A, B trained by the weekly, daily peak load for June and July. Model C, D added representative temperature data for same periods A and B. Case 3 train data for July and August to forecast the peak in July and August of next year. Model A and B used weekly peak load and daily peak load as inputs. Model C and D added representative temperature data and GDP growth to reflect the electric load change between last year and this year. The reason for forecasting July or August in all models is that the summer peak load always occurs during the period.

Each model's results are set for weekly or daily peak loads. Among them, the maximum load in each model is chosen as the peak load in this summer. Finally, using an ensemble average method called bagging (Bootstrap Aggregating) to obtain the summer peak load of the certain year.

4. Results and Discussion

In this section, we evaluate the proposed peak load forecasting method in terms of forecasting performance. To measure the forecasting performance, we use the Percentage Error. It is calculated by :

$$ERROR = \left|\frac{A_t - F_t}{A_t}\right| * 100 \qquad (6)$$

where F_t is forecasting peak and A_t is an actual peak. The training period of the forecasting model is set as 5 years before the forecasting year. In ANN the weighting matrices are sensitive to the choice of initial value, each model was estimated 200 times from random initial value. Then, the forecasting results were determined by minimizing the out-of-sample error.

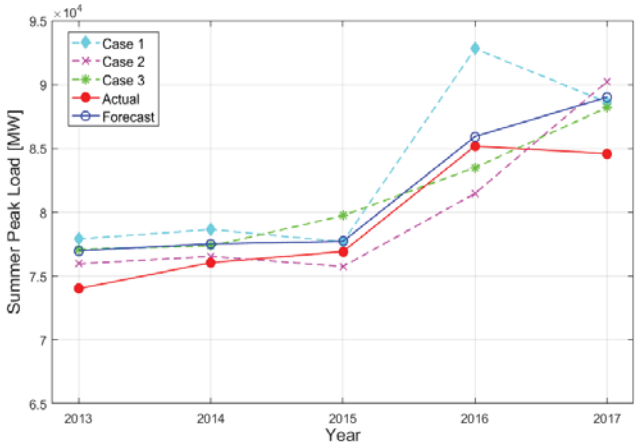

Fig. 3. Forecasting Results by Cases

Fig. 3 show forecasting results by cases. Final forecasting results is ensemble average of three cases. As shown in Fig. 3. Summer peak loads continue to increase throughout the year. Table 3 shows forecasting accuracy. In some years, forecasting accuracy of a single case is good, but on average, the ensemble method shows better performance than single models which means that the ensemble method give more stable result. From 2013 to 2017, generally forecasting results have small error. That indicate to proposed forecasting method is valid for peak forecasting. In 2013 and 2017, there are large errors compared to other years since the peak load of these year is lower than that of previous year.

Table 1. Forecasting Accuracy by Cases

Year	Case 1	Case 2	Case 3	Ensemble
2013	5.2584 %	2.6305 %	4.1600 %	4.0154 %
2014	3.4186 %	0.6469 %	1.7093 %	1.9248 %
2015	0.9907 %	1.5198 %	3.6624 %	1.0446 %
2016	8.9572 %	4.3553 %	1.9793 %	0.8739 %
2017	4.7537 %	6.6843 %	4.2797 %	5.2395 %
AVERAGE	4.6757 %	3.1674 %	3.1581 %	2.6196 %

Note that, when we use more data to train, the forecasting accuracy is improved. When we use 10-year data as training, the forecasting error of 2017 is 3.7651%.

5. Conclusion

In this paper, we propose a summer peak forecasting scheme based on ensemble average of ANNs. We use three ANNs of which input parameters are selected among the previous years' peak loads in summer, temperature, and gross domestic product (GDP). The proposed scheme consists of three stages: i) collect the peak load data and the temperature data, and then obtain the representative temperature and; ii) train three ANNs and obtain three forecasting loads from the three ANNs; iii) the final result is obtained by using ensemble average of the three results, i.e., bagging. The case study shows that the proposed ensemble average method generally performs better than single ANN method.

For the future work, we will use non-temperature meteorological elements and more power data to improve forecasting accuracy.

Acknowledgements

This work was supported in part by "Human Resources Program in Energy Technology" of the Korea Institute of Energy Technology Evaluation and Planning (KETEP), granted financial resource from the Ministry of Trade, Industry & Energy, Republic of Korea (No. 20164010201010), and in part by Korea Electric Power Corporation (Grant number: R17XA05-62).

References

[1] A. Goia; C. May; G. Fusai, "Fuctional clustering and linear regression for peak load forecasting", International Journal of Forecasting, vol.26, no.4, pp.700-711, 2010.

[2] KEPCO Electric Rates Table. Retrieved from https://home.kepco.co.kr/kepco/EN/F/htmlView/ENFBHP00101.do?menuCd=EN060201

[3] Stephen Marsland, (2014), *Machine Learning : An Algorithmic Perspective, Second Edition*, CRC press

[4] M. Amin-Naseri; A. Soroush, "Combined use of unsupervised and supervised learning for daily peak load forecasting", Energy Conversion and Management, vol.49, no.6, pp.1302-1308, 2008.

[5] T. Matsui; T. lizaka; Y. Fukuyama, "Peak load forecasting using analyzable structured neural network", IEEE Power Engineering Society Winter Meeting, vol.2, pp.405-410, 2002.

[6] Kim, C.; Yu, I.; Song Y.-H.; "Kohonen neural network and wavelet transform based approach to short-term load forecasting", Electric Power Systems Research, vol.63, no.3, pp.169-176, 2002.
[7] G. Khan; S. Khan; F. Ullah, "Short-Term daily peak load forecasting using fast learning neural network", ISDA 2011 International Conference, pp.843-848, 2012.
[8] KPX, Full-day power supply and demand. Retrieved from http://www.kpx.or.kr/www/contents.do?key=219
[9] L. Hernández; C. Baladrón; J. Aguiar; L. Calavia; B. Carro; A. Sánchez-esguevillas; D. Cook; D. Chinarro; J. Gómez, "A Study of the Relationship between Weather Variables and Electric Power Demand inside a Smart Grid/Smart World Framework.", Sensors(Basel), vol.12, no.9, pp.11571-11591, 2012.
[10] Hou, Y.-L;, Mu, H.-Z; Dong, G.-T; Shi, J, "Influences of Urban Temperature on the Electricity Consumption of Shanghai", Advances in Climate Change Research, vol.5, no.2, pp.74-80, 2014.
[11] Turan Gonen (2008), *Electric Power Distribution System Engineering : Second Edition* , CRC press
[12] S. Papadopoulos; L. Karakatsanis, "Short-term electricity load forecasting using time series and ensemble learning methods", IEEE PECI, 2015.