The Optimal Combination of Neural Networks for Next Day Electric Peak Load Forecasting

Hiroyasu Konishi (1) **Masanori Izumida** (2) **Kenji Murakami** (2)
Shikoku Research Institute inc. (1) Ehime university (2)
Phone +81-89-947-43111 (1), Fax +81-89-947-4936 (1)
E-mail konishi@etl.ssken.co.jp (1)

Abstract: We introduce the forecasting method for a next day electric peak load that uses the optimal combination of two types of neural networks. First network uses learning data that are past 10days of the target day. We name the neural network Short Term Neural Network (STNN). Second network uses those of last year. We name the neural network Long Term Neural Network (LTNN). Then we get the forecasting results that are the linear combination of the forecasting results by STNN and the forecasting results by LTNN. We name the method Combination Forecasting Method (CFM). Then we discuss the optimal combination of STNN and LTNN.

Using CFM of the optimal combination of STNN and LTNN, we can reduce the forecasting error.

1. Introduction

A next day electric peak load forecasting is important for electric power companies. In case of making the plan of generation for power plants, an accurate next day electric peak load forecasting is important. And, It is important especially in summer, because the electric peak load and the change in load of this season are large. Therefore, we discuss a next day electric peak load forecasting in summer. Target days are weekdays in summer season of Shikoku district. As for Shikoku district, the north side faces Seto inland sea, and the south side faces the Pacific Ocean. Therefore, the forecasting is difficult, as the weather condition is different.

A neural network is used as a technique of peak load forecasting, and it is reported effective [1]-[5]. Therefore, we use a neural network for forecasting.

Usually, following learning data are used in the forecasting by a neural network.

(1) Learning data of past about 10days of a target day.

(2) Learning data of past years.

In our forecasting method, two neural networks are introduced to use those two learning data. First neural network uses the learning data (1), and we call the neural network Short Term Neural Network (STNN). Second neural network uses the learning data (2), and we name the neural network Long Term Neural Network (LTNN). Next, a linear combination of the forecasting result of STNN and LTNN is given as the forecasting result. We name the method Combination Forecasting Method (CFM). In addition, we use the proposed method [6] in STNN .

In this paper, we forecast a next day electric peak load using CFM, and we discuss the optimal combination of STNN and LTNN.

2. Definition of Neural Network

We use a three layer neural network shown in Fig. 2.1 and the back-propagation method. In Fig. 2.1, m is an input layer's unit number, n is a hidden layer's unit number, and an output unit number is one. Input features are selected in the following four features.

(1) Maximum temperature of a day (Tx).

(2) Average temperature of a day (Ta).

(3) Minimum temperature of a day (Tm).

(4) Electric peak load of two days ago a target day (P).

Desired output y_i and P are electric peak loads of general use of Shikoku district and we use the value that 10000Mw are normalized at 1.0.

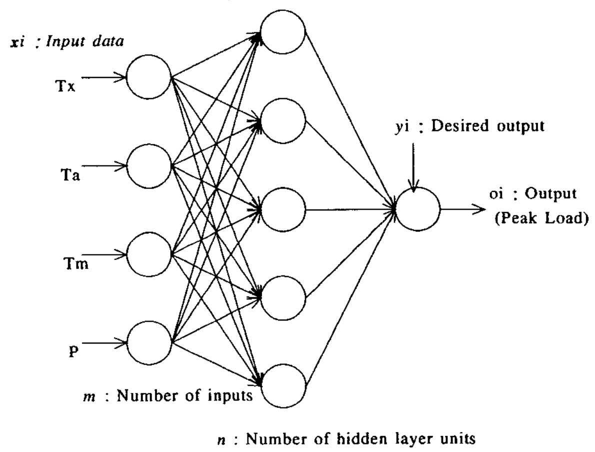

Fig. 2.1 Three layer neural network

Input data that have the input features (1), (2), (3) are weighted average data of actual values in the four cities, and the cities are with the prefectural offices in Shikoku district.

The ratios of weights are the electric peak load ratios of four cities. These data are not normalized, and Table 2.1 indicates the ratios of weights.

Table 2.1 The ratios of electric peak load

Prefecture	Tokushima	Kochi	Ehime	Kagawa
Ratio(%)	20.5	17.5	34.0	28.0

We assume each atmospheric data of four prefectures T_1, T_2, T_3, T_4, and input data T. Then, we get the following formula

$$T = \frac{20.5\,T_1 + 17.5\,T_2 + 34.0\,T_3 + 28.0\,T_4}{100}.$$

In case of learning, these atmospheric data and a desired output y_i are actual data. In case of forecasting in practice, atmospheric data should be forecasted data by a weather forecast. Nevertheless, we use actual data for a forecasting, not to take the influence of the error of the weather forecast. The error disturb us to verify the effect of the method and to find the suitable input features.

Target days are weekdays in the period from 1 July to 30 September.

3. Short Term Neural Network (STNN)

3.1 Parameters of STNN

The atmospheric conditions of the days near the target day tend to be similar to those of a target day. Therefore STNN uses learning data of past 10 days of a target day.

Forecasting by STNN, we make a parameter as follows. They are decided by preliminary experiments.

(1) Input features: All combinations of four input features. (Tx,Ta,Tm,P).

(2) Learning data: data of 10 days. (Before the target day - the past 10 days).

(3) Initial weight: Random numbers of uniform distribution (the range of -0.01~0.01).

(4) Number of hidden layer's unit: 20 .

(5) Output function:
   Hidden layer: sigmoidal function f (x) = 1/(1+exp(-x)).
   Output layer: linear function f (x) = x.

(6) Learning cycles: 3,000,000 cycles.

In addition, we use the learning method named Priority Based Learning Method (PBLM) described as follows.

3.2 Priority Based Learning Method

The Back-propagation learning method is a method decreasing the error. The error is sum of the errors for all learning data, so learning data are treated as same importance or priority. In some cases, the method is effective, because the average error is minimized. There are other cases that the importance or reliability of learning data is not same. If we use the unreliable learning data as same reliability as other reliable data, resulted neural networks may not be reliable. In general, it is difficult to say which data are trusted or not.

In peak load forecasting, the learning data far from the target day have less reliability, because the weather condition is different to the target day. In this case, the errors of learning data far from the target day have lower priority than learning data of yesterday.

Priority Based Learning Method (PBLM) is proposed for such a situation. In PBLM, errors of unreliable learning data are estimated small. This means that the neural network roughly approximate the learning data. On the other hand, errors of reliable data are estimated larger, so errors will be smaller. PBLM is given as follows.

The sum of square errors E at k sets of learning data $(\boldsymbol{x}_i,\ y_i)$ is given by

$$E = \sum_{i=1}^{k} (y_i - o_i)^2 = \sum_{i=1}^{k} \varepsilon_i,$$

where y_i is desired output; o_i is output, when $\boldsymbol{x}_i$ (input data of learning data) is inputted to the neural network; ε_i is a square error at input data $\boldsymbol{x}_i$.

In the same way, the sum of square error by PBLM E' is given by

$$E' = \sum_{i=1}^{k} \mu_i (y_i - o_i)^2 = \sum_{i=1}^{k} \mu_i \varepsilon_i,$$

where μ_i is the priority.

And the amount of correction of the weight of PBLM $\Delta \boldsymbol{w}'$ is given by

$$\Delta w' = \frac{\partial E'}{\partial w} = \sum_{i=1}^{k} \mu_i \frac{\partial \varepsilon_i}{\partial w},$$

where $\boldsymbol{w}$ is a weight vector. And $\Delta \boldsymbol{w}'$ is written as

$$\Delta w' = \sum_{i=1}^{k} \mu_i \Delta w_i$$

where

$$\Delta w_i = \frac{\partial \varepsilon_i}{\partial w}$$

and Δw_i is the amount of correction of the processing weight at x_i by normal learning method.

In case of PBLM, the amount of correction at input data x_i is $\mu_i \Delta w_i$. Being changed μ_i , an amount of correction can be set at arbitrary values. In this way, the amount of correction in learning is changed corresponding to the input data, and put a suitable priority for the learning data x_i .

3.3 Experiments by STNN

Using all combinations of four input features, we have had forecasting experiments. Then we have had the result that a combination of Ta and P gives the smallest average error of 2 year's average shown in Table 3.1.

Table 3.1 Forecasting error by STNN

Input Features	Average error (%)		
	1995	1996	2 years
Ta,P	2.02	2.03	2.03

The average error is given as

$$\text{Average error}(\%) = \frac{1}{d}\sum_{i=1}^{d}\frac{|o^f_i - o^{act}_i|}{o^{act}_i} \times 100\,(\%) \quad ,$$

where d is the number of forecasted days and d is about 60; o^f_i is forecasted output; o^{act}_i is actual electric peak load of a target day.

In summer, atmospheric conditions change in short period. STNN learns the data according to the changing of atmospheric conditions in short period near the target day. Though, STNN can forecast corresponding flexibly to the changing.

On the other hand, if the atmospheric condition of a target day is different from those of learning data, STNN cannot learn the data for forecasting. Therefore , forecasting errors become large in this case. Such the case is likely to happen, because STNN learns only 10days. In addition, if learning data contain irregular data forecasting errors become large.

4. Long Term Neural Network (LTNN)

LTNN uses learning data from 1 June to 30 September in a last year. And, it uses more data than those of STNN do. Therefore, even if the atmospheric condition of a target day is different from those of past 10 days, LTNN can learn the data in a last year. And, even if learning data contain the irregular data, the influence of the irregular data is small. In addition, we can remove the irregular data, because there are many data though the irregular data is removed. In this way, LTNN can give stable forecasting results.

On the other hand, LTNN cannot learn the data according to the changing of atmospheric conditions in short period near the target day. Therefore, LTNN cannot forecast corresponding flexibly to the changing.

Then we have forecasting experiments by LTNN. LTNN is the same neural network of STNN in Fig. 2.1, but some parameters are different as follows. These parameters are given by experiment. In addition, these are suitable for LTNN.

(1) Input features: All combinations of four input features (Tx, Ta, Tm, P).
(2) Learning data: data of weekdays of last year.
(3) Initial weight: Random numbers of uniform distribution (the range of −0.01～0.01).
(4) Number of hidden layer's unit: 2.
(5) Output function:
Hidden layer: sigmoidal function f(x) = 1/(1+exp(-x)).
Output layer: linear function f(x) = x.
(6) Learning Cycles: 20,000,000cycles.

In forecasting by LTNN, outputs must be corrected according to the increasing of electric peak load by a year, because LTNN uses data of last year. The rate is the average of daily electric peak load on April in a target year divided by the one of a last year. The values are 1.057 for 1995, 1.045 for 1996.

Using all combinations of four input features, we have forecasting experiments. Then we have had the result that a combination of Tx, Ta, Tm and P gives the smallest average errors of 2year's average shown in Table 3.1.

The average errors by LTNN are given by the same ways for STNN.

Table 4.1 Forecasting error by LTNN

Input Features	Average error (%)		
	1995	1996	2years
Tx,Ta,Tm,P	2.06	2.10	2.08

5. Combination Forecasting Method (CFM)

There is a difference between the distribution of the forecasting errors by STNN and that by LTNN, because STNN and LTNN use different learning data to each other. Therefore, we consider that a forecasting error can be reduced, combining those forecasting results suitably.

Therefore, we set the final forecasting result O_C as,

$$O_C = (1 - r) O_S + r O_L,$$
$$(0 \leqq r \leqq 1),$$

where O_S *is the forecasting result by STNN*, O_L is the forecasting result by LTNN, and r is the combination ratio changing from 0 to 1.

To find the optimal input features of STNN and LTNN for CFM, we have had the forecasting experiments by all combination of input features of STNN and LTNN. Table 5.1 shows the result of the experiments, and it shows the top six results. The best result is given, when input features of STNN are Ta and P, and those of LTNN are Tx,Tm,P, and r is set at 0.43.

Fig 5.1 shows the transition of forecasting errors when r is set from 0 to 1.0. When r is set at 0, an error is the error by STNN, and when r is set at 1, an error is the error by LTNN. Input features of STNN are Ta and P, and those of LTNN are Tx,Tm,P. Fig 5.1 shows that CFM gives smaller errors compared with the errors by STNN or LTNN. In case of the actual forecasting, we can't know the best value of r before the forecasting. But we can have nearly the smallest errors, when we set r from 0.4 to 0.5.

Table 5.1 Forecasting errors by CFM

Input features		Average Error (%)			r
STNN	LTNN	1995	1996	2years	
Ta,P	Tx,Tm,P	1.657	1.618	1.638	0.43
Ta,P	Tx,Ta,Tm,P	1.654	1.633	1.643	0.45
Ta,P	Tx,Tm	1.835	1.529	1.682	0.32
Ta	Tx,Tm,P	1.755	1.618	1.686	0.45
Ta,P	Tx,Ta,Tm	1.829	1.549	1.689	0.30
Ta	Tx,Ta,Tm,P	1.745	1.638	1.691	0.47

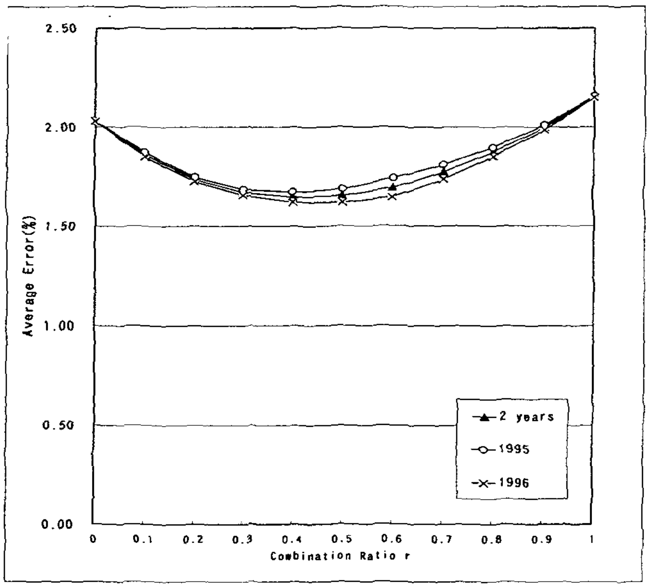

Fig. 5.1 Transition of Forecasting Errors

When we use optimized CFM, forecasting errors are smaller than the errors by STNN and the errors by LTNN shown in Table 3.1 and Table 4.1. We can reduce about 19% of average error compared with STNN or LTNN.

6. Conclusion

We have forecasted next day electric peak load of weekday in summer of Shikoku district. In the forecasting, we have proposed Combination Forecasting Method (CFM) that is the linear combination of Short Term Neural Network (STNN) and Long Term Neural Network (LTNN). And in forecasting by CFM, we discuss the optimal combination of STNN and LTNN. Finally, we have had 1.64% of average error when the input features of STNN are Ta, P and those of LTNN are Tx, Tm, P, and combination ratio r is set at 0.43. In addition, we find that we can have nearly the smallest errors, when we set r from 0.4 to 0.5. Using CFM, we can reduce about 19% of average error compared with STNN or LTNN. It should be concluded from above that CFM is effective in the forecasting.

A problem is how to set the combination ratio r. If it can be set flexibly corresponding to weather conditions, forecasting errors may be reduced more. In the future, we may discuss how to set the r, and have the forecasting experiments at the actual use.

References

[1] T.Onoda, E.Ohba, "Selection of the Optimal Neural network Architecture for Next Day Peak Load Forecasting" T.IEE Japan, Vol.118-B, No.5, pp.497-504, 1998.

[2] Y.Sugai, H.Horibe, T.Kawase, "Forecast of Daily Maximum Electric Load by Neural Network" T.IEE Japan, Vol.117-B, No 6, pp.872-879, 1997.

[3] K.Makino, T.Shimada, R.Ichikawa, M.Ono, T.Endo, "Daily Peak Load Forecasting Using an Artificial Neural Network and Its Improving Method of Reducing the Forecasting Errors" T.IEE Japan, Vol.115-B, No.11, pp.1304-1313, 1995.

[4] Y.Ueki, T.Matsui, H.Endo, T.Kato, R.Araya, "Development of Peak Load Forecasting System using Neural Network and Fuzzy Theory" T.IEE Japan, Vol.115-B, No.9, 1038-1045, 1995.

[5] T.Ishida, S.Tamaru, "Daily Electric Load forecasting Using Artificial Neural Network" T.IEE Japan, Vol.114-B, No. 11, pp.1109-1115, 1994.

[6] H.Konishi, M.izumida, K.Murakami "Next Day Electric Peak Load Forecasting Using Priority of Learning data" 1999 International Technical Conference on Circuits/Systems, Computers and Communications.