18th International Power System Conference

81-E-LEM-790

One Month Ahead Temperature Prediction for Mid-term Load Forecasting

Bahman Kermanshahi

Department of Electrical and Electronics Engineering
Tokyo University of Agriculture & Technology (JAPAN)

Keywords: Weather Conditions, Temperature Prediction, Mid-Term Load Forecasting,Artificial Neural Network (ANN), Neuro-Genetic Learnin

Abstract:

Temperature is the most important factor in load forecasting of power system analysis. Particularly in short-term load forecasting, it plays an important role in increase/decrease of energy consumption. For instance, according to the Japanese Power Industry announcement in 2001, increase of 1 degree Celsius will cause about 5GW increase in electric power consumption at the summer peak. This amount is as same as consumption power used by 1.6 million general households or amount of generated power by 5 large-scale utility power plants. On the other side, 1 degree Celsius of temperature change had caused about 1.85GW increase in power consumption in the winter of year 2000.

Basically, the short-term temperature (hourly up to 1 week ahead) is researched and predicted by environment agencies of every country. Therefore, it is easy to obtain the forecasted temperature data from those agencies, newspapers, TV news and so on. However, it is difficult to obtain the hourly temperature beyond 1 week. Although Japan Meteorological Agency (JMA), which uses the Numerical Weather Predictions (NWP), announces the forecasting data up to 1 or 2 months ahead, but they are expressed only as "high" or "low" which is compared with normal years. This means, we can only know that the temperature may goes up or comes down every day. In addition, super-computer processes it with lots of complex meteorological formulations. The applied data the ones which have observed by weather satellite all over the world.

However, if the temperature could be predicted for a longer period, it becomes even a useful factor for projecting a better resolution for the long-term load forecasting, prediction of fuel amount necessary for next couple

months of power plants and so forth. In this paper, some intelligent methodologies such as artificial neural network and a combined neuro-genetic algorithm have been used to predict the temperature up to one month ahead

1. Introduction

There are 3 typical weather forecasting methods, the numerical weather prediction, NWP, statistical method and intelligent method. They all have unique characteristics described as the following sections.

1.1-.*Forecasting Methods*

A. NWP [1][2]

This forecasting method simulates the atmosphere using meteorological equations known as numerical models. Forecasting is then made by solving very complex equations using the weather data observed at observatory stations, satellites, and so on.

The atmosphere is split into layers, stacked in columns, with thousands of these columns distributed over grid points spaced over the Earth's surface.

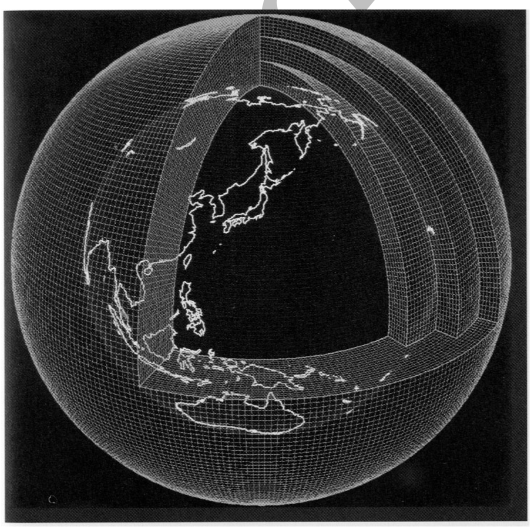

Fig. 1 Global grid points
(Provided by Japan Meteorological Agency: JMA)

Table 1NWP models

Model	Resolution		Lattices	Applied to
	Horizontal	Vertical		
Global Model	55 Km	30 layers	640 x 320	Longterm forecast
Limited Area Model	20 Km	36 layers	257 x 217	Shortterm forecast

Fig. 1 shows the global scale of grid points. Weather observations, for example wind speed, wind direction, temperature, etc., on each grid points are computed for any individual time. Then they are fed into the equations to simulate the atmospheric fluctuations. Accuracy of the simulation depends on distance between the grid points. Shorter distance projects better resolution and therefore fine divisional forecasting that considers landform is accomplished.

In short-range forecasting, few hours up to few days ahead, there aren't big affects by global atmospheric fluctuation. Therefore, in Japan, the data of Asian countries around Japan are often used for the forecasting. This is why a limited area model which has high resolution is considered for small regional atmospheric fluctuation. This model consists of about 2 million grid points, 36 layered, 257×217 lattices.

On the other hand, the global atmosphere fluctuates is considered during one week of period. That is why global model is used for long-range forecasting up to 7 days ahead. This model consists of about 6 million grid points, 30 layered, 640×320 lattices. Table 1 shows these NWP models.

There is 20Km distance between each grid points in horizontal on former and 55Km distance on latter. If the resolution of latter were same as former, there are 45million grid points, 1760×880 lattices. It should be noted that there a plenty of equations for forecasting and it is obvious that longer forecasting period need more equations. It is very difficult to obtain a solution fast with such big equations by even today's supercomputer.

On the other hand, in NWP, grid points are used to express a continuous atmosphere. However, it is impossible to recognize the weather by only one point information. For example, one point may show a cloudy weather, but actually there is a possibility that the weather around the points become sunny. There are limited to express the quintuple scale weather if the distance between grid points is 20Km.

In such a reason, it is said that it is impossible to simulate the atmospheric fluctuation completely. Also, the longer forecasting period the worse accuracy you obtain by the NWP alone. Therefore, some other methodology and/or technology should be investigated.

B. Statistical method

Statistical method becomes one of the choices if and only if there are lots of past observed weather data available. This method forecasts the weather by comparison with past-observed data. Fig.2 shows a general idea of this method. If a present pattern of weather resemble with past pattern that is in a database, the data becomes a base of forecasting result. This method does not need any mathematical equations and is only depended on the forecasting place.

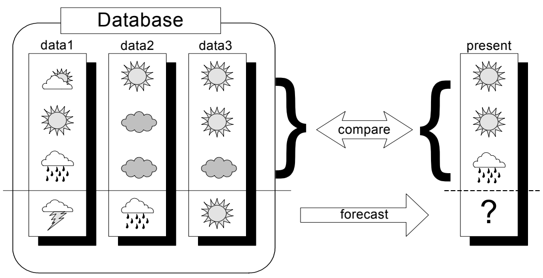

Fig.2 Statistical method

In addition, a huge database of past data has to be prepared. Accuracy of this method is low as this method does not relate to physics.

C. Intelligent method

Intelligent method is constructed by using intelligent techniques like Artificial Neural Network (ANN), Genetic Algorithm (GA), and Chaos and fractal, etc. If these systems are given some input and output data, they can learn the correlations among them and produce the forecast output(s). However, enough past observed weather data is required. However, the number of required data are much smaller than that required by statistical and NWP methods. The unique characteristic of this system is the learning behavior. Fig.3 shows a general idea about this method.

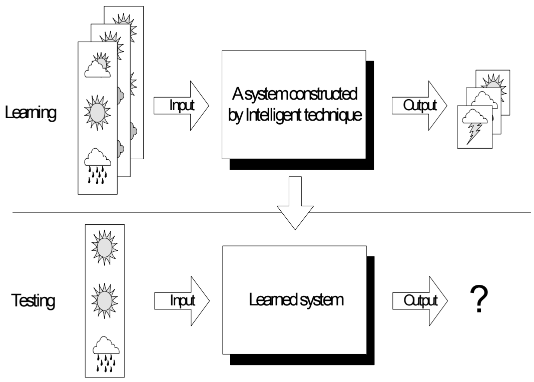

Fig.3 Intelligent method

1.2 Comparisons

Based on the comparisons on the characteristics and uniqueness of the above mentioned methods, it is decided to use the intelligent method. The prime objectives of this study using the intelligent system is to forecast every hour temperature up to 1 month ahead easily and construct forecasting system that substitute for the mid-term load forecasting.

As mentioned earlier, statistical and NWP methods needs more data than intelligent method. Particularly, the intelligent method is more flexible than others and can handle the non-linear problems easily

2. Construction of Temperature Forecasting System

2.1 Choice of forecasting technique

In this study, the intelligent method is chosen for forecasting the temperature up to 1 month ahead. There are many intelligent techniques like ANN, GA, Chaos, etc, and also there are many researches on these techniques, such as studies on ANN[3], Chaos-Neuron and so on. The author had already used the ANN[4] for temperature forecasting up to 2 weeks ahead. In this paper, not only stand-alone ANN is used, but also a combination of ANN and GA, known as Neuro-Genetic Learning (NGL) is chosen for the forecasting system.

2.2Neuro-Genetic Learning (NGL)

There are many common points between ANN and GA. First, they are both systems that are used for leaning or adaptability. Second, they are imitation of living things, ANN for nerve systems and GA for evolutions. Third, they have parallel processing behavior. On the other hand, their biggest difference is that ANN looks for individual answer and GA looks for adaptability of a group.

The relation between GA and ANN is similar to the relation between evolution and brain of animals. A framework of nerve system is almost determined by the genetic information among the evolutions. It should be noted that animals also learn by their nerve systems which is very much depended on their living environment. Based on this reason, a fusion of ANN and GA had been used by many researchers for constructing the basic architecture of artificial life. The NGL is one of those propositions. In this technique, GA prevents ANN to stack at local minima of the solution space, and therefore the efficiency of the ANN's learning increases. Fig. 4 shows the difference of ANN and GA from the view point of searching for global minima in an error surface.

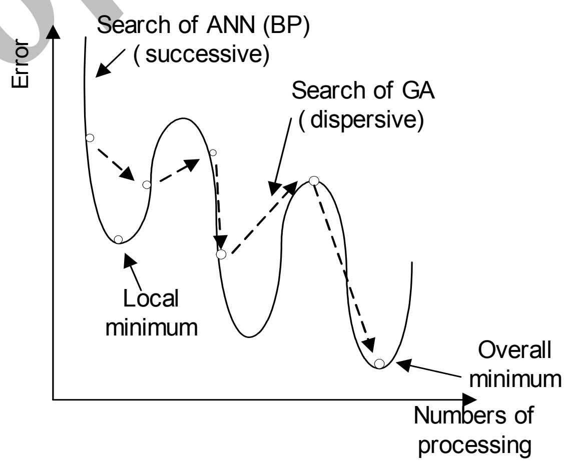

Fig. 4 The solution search technique of ANN and GA

2.3 Flowchart of the system

Fig. 5 shows the flowchart of this temperature forecasting system. This system consists of 3 stages. The processes of each stage are shown as the following sections.

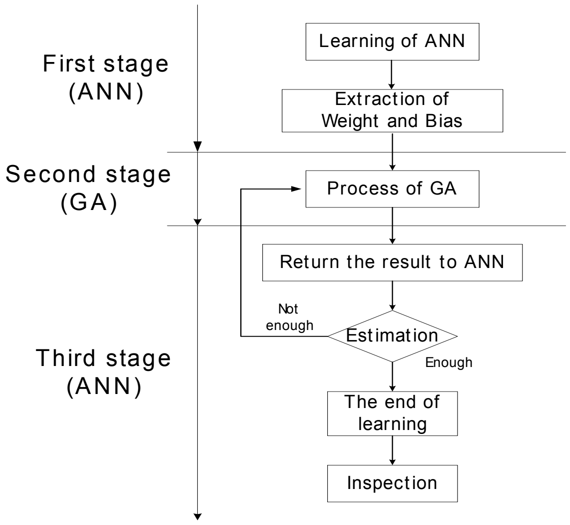

Fig. 5 System flowchart

(1) The first stage (ANN)

In this stage, the ANN learns correlation between inputs and outputs data using Back Propagation (BP) technique, which is one of the learning methods of ANN. Fig. 6 shows a 3 layered ANN that is used for this stage. In BP, the errors between target data and system output are propagated back to input layer and the weight coefficients, which is the correlation between inputs and outputs, the biases, which is a supplementary term of weight coefficients, are adjusted in order to lower the error.

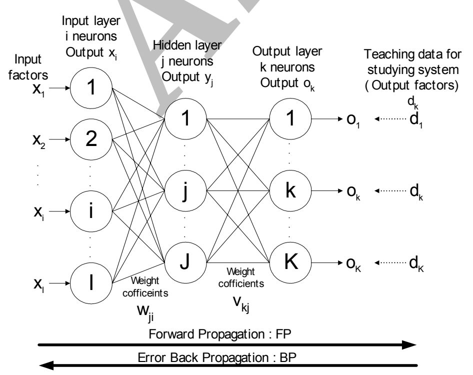

Fig. 6 An ANN model (back-propagation model)

ANN has 3 typical system parameters, learning rate, momentum and numbers of hidden neurons. Generally, they are adjusted by trial and error before learning. In this study, it is also considered to determine the optimal values for these parameters by GA. However, this may make the problem of temperature forecasting more complex, so this time; it is evaded to do it. In this system, learning rate and momentum are automatically changed wit the range of 0.1 from 0 to 1 respectively. Each of them is processed for fixed times and become candidacies for the forecasting system.

(2) The second stage (GA)

Here, the new combinations of weight coefficients and the biases are created to rearrange the ANN. On the other hand, the GA creates new combinations by reproduction and selection, crossover and mutation, which are the process of GA. This is similar to the process of living life. Fig. 7 shows the concept of GA process. One-point crossover and Multi-point crossover is shown in this Figure. In this study, one-point crossover is selected.

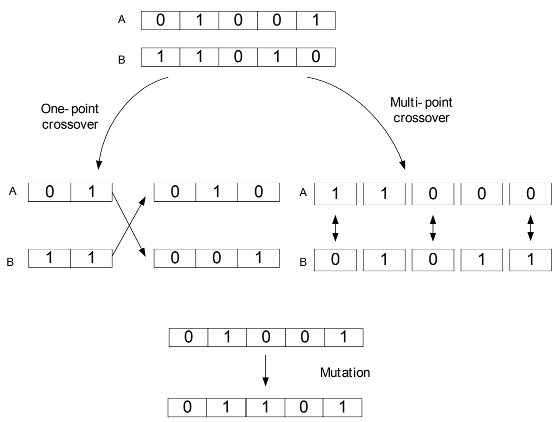

Fig. 7 The concept of GA process

(3) The third stage (ANN)

Finally in this stage, the new individuals, that created at the second stage are sent to this stage and estimated. It is intended to recognize whether they are fit to an environment or not. The environment is defined as "fitness function" and if the individuals fit to the function, they survive and become candidates for the generation. In this study, Forward Propagation (FP), shown in Fig. 6, is used for the fitness function. If the error between target data and system output reduced, the individual seems to fit the environment. This routine is repeated until an arranged pre-set time is reached or the error becomes lower than a pre-set level.

If they meet the conditions, the learning and constructing of the system is terminated. In this case, it is said that the optimal weight coefficients are obtained and the ANN is trained successfully.

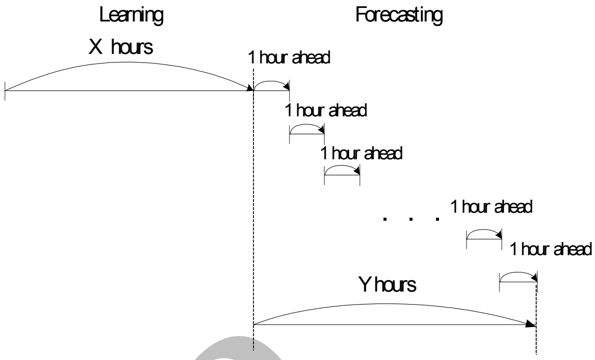

Fig. 8 direct forecasting

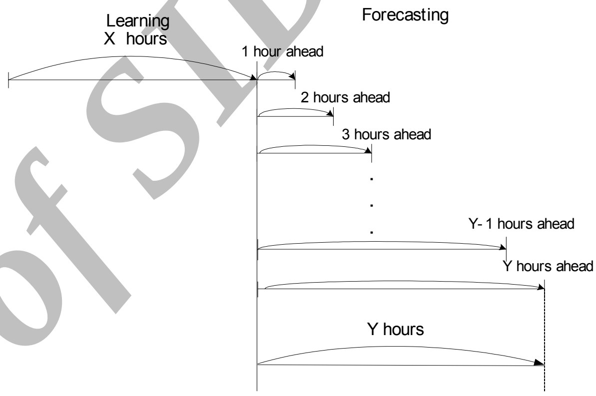

Fig. 9 Renewing forecasting

2.4The prediction method

Every hour temperatures up to one month ahead are forecasted in this study. However, still there are some problems regarding "how to forecast the desired temperature successfully." The typical problems are shown as below.

1. What is the best way to forecast?
2. What kind of data should be used for the forecast?

First, the first problem is considered. It seems that there are 2 methods to forecast temperature up to 1 month ahead. One is to forecast the term directly. The concept of this method is shown in Fig. 8. On the other hand, there is another forecasting way that repeats to forecast 1 hour ahead and accumulate the results until it becomes 1 month. This concept is shown in Fig. 9. In this study, the latter way is chosen.

Previously, the author constructed a system for the maximum and minimum temperature forecasting system up to 2 weeks ahead. It became clear that we need to have 28 networks (1 system for 1 forecasting). Therefore the previously developed network is not applicable for the present study. On the other hand, the latter forecasting system needs only one system for the forecasting. Although this forecasting method, the latter one, is useful, it still has a problem. It repeats forecasting with this idea that the forecasted data is used as for the next input data. Therefore, input factors which are useful on the next forecasting should be chosen. So, inputs related to temperature and time became candidate inputs. In the pre-simulation, 52 factors are selected as for inputs of the forecasting system. They are shown in Table 2.

Table 2 Input factors for the forecasting system

Input Factors (52 factors)	• Temperature at I, I-1, I-2 (1 factor respectively) • Difference between temperature I and I-1 (1 factor) • Difference between temperature I-1 and I-2 (1 factor) • Difference between temperature I-2 and I-3 (1 factor) • Month (4 binary factors) • Day (5 binary factors) • Hour (24 factors) • Ranges of temperature at I (10 factors) • Fluctuation of temperature at I against I-1 (1 factor)

3. Temperature Forecasting Simulation

3.1 The definition of forecasting term

Fig. 10 and Table 3 show the average temperature of every month observed at Otemachi weather station (considered as center of Tokyo) during 1995 to 2000. There is a similar trend for each year and it is clear that there are almost 20 degrees difference in temperature between February and August. In this study, an ANN is used for constructing the forecasting system. All input and output data are normalized into 0 to 1 and the learning process is done on the ANN. This is why 1 year is divided into 4 seasons to make the differences of temperature lower.

Table 3 Average and rounded temperatures of a year

Month	Average Temperature (C)	Rounded Temperature (C)
January	6.6	7
February	6.4	6
March	9.7	10
April	14.8	15
May	19.4	19
June	22.1	22
July	26.4	26
August	27.8	28
September	24.2	24
October	19.1	19
November	13.6	14
December	8.8	9

If the average temperatures, which are rounded off to the nearest place of decimal, are above 20 degrees Celsius, it is defined as summer, and if it is below 10 degrees Celsius, it is defined as winter. Others are defined as spring and fall respectively. This is shown in Table 4.

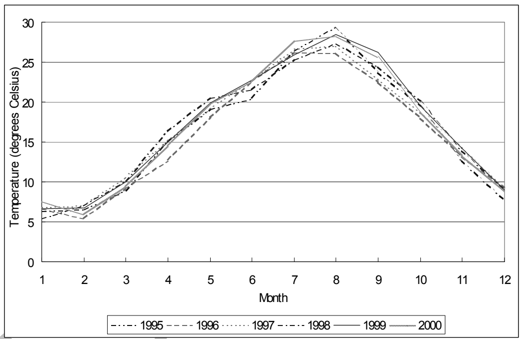

Fig. 10 The change of average temperature(Observed at Otemachi weather station, center of Tokyo)

Table 4 Definition of seasons

Season	Month
Spring	March - May
Summer	June - September
Fall	October - November
Winter	December - February

3.2 Data for constructing system and simulation

The definition of constructing (Learning) and simulating terms are shown in Fig. 11. Based on this figure, 80% of data is used for learning and 20% for temperature forecasting simulation. Table 5 shows the details of the data used for learning and simulating.

In the learning stage, the system is learned as to forecast temperature 1 hour ahead, and in the simulating stage, the system repeats forecasting of every hour temperatures up to 1 month ahead. As mentioned earlier, the forecasted data is then used as input for the forecasting of next hour temperature.

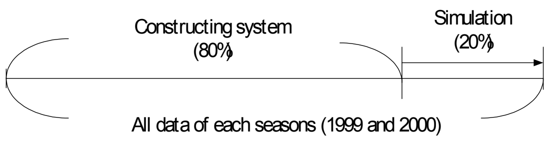

Fig. 11 Learning and simulating terms

Table 5 Learning and simulating term

	Learning Term	Simulating Term
Spring	From 1/Mar/'99 0:00 to 31//May/'99 23:00andFrom 1/Mar/'00 0:00 to 24//Apr/'00 5:00 (3533) hours	From 24/Apr/'00 6:00 to 31/May/'00 23:00 (881 hours)
Summer	From 1/Jun/'99 0:00 to 1/Sep/'99 23:00andFrom 1/Jun/'00 0:00 to 12/Aug/'00 3:00 (4683hours)	From 12/Aug/'00 4:00 to 22/Sep/'00 19:00 (1000hours)
Fall	From 1/Oct/'99 0:00 to 30/Nov/'99 23:00andFrom 1/Oct/'00 0:00 to 5/Nov/'00 14:00 (2342hours)	From 5/Nov/'00 15:00 to 30/Nov/'00 23:00 (584hours)
Winter	From 1/Dec/'99 0:00 to 29/Feb/'99 23:00andFrom 1/Dec/'99 0:00 to 23/Jan/'00 22:00 (3475hours)	From 23/Jan/'00 23:00 to 28/Feb/'00 23:00 (865hours)

3.3. Numbers of hidden neurons

Successful learning of the ANN depends on the numbers of hidden neurons (shown in Fig. 6). Therefore it should be investigated first. Here, three patterns, situations at the generations (process times) of GA are set at 1, 5, 10, and tried to obtain optimal numbers of hidden neurons. They are investigated with numbers of 1 to 20, 25, 30 and 50. Table 6 shows the conditions of the investigations. In this table, "Learning epochs" is considered in which the system learns all data once. This simulation is done for every season (Table 7).

Table 6 Conditions of the investigations

ANN	Number of hidden neurons	1-20, 25, 30, 50
	Learning epochs	10
GA	Generations	1, 5, 10
	Individuals	50
	Crossover rate (%)	50
	Mutation rate (%)	2

Table 7 Investigated conditions regarding 4 seasons

GA Generations	Number of hidden neurons			
	Spring	Summer	Fall	Winter
1	12	16	7	6
5	19	16	6	6
10	19	12	6	6
Decided number of hidden neurons	19	16	6	6

In Table 7, three simulations for each season do not show the same numbers of hidden neurons. But 2 of the results of spring, summer and winter show the same numbers. In addition, the defined neurons for fall take the middle number of the simulation.

3.4. Selection of GA generations

Next, the generations of GA are investigated. Generations of 1 to 10, 15 and every 10 from 20 to 100 are tried for this investigation. Table 8 shows the conditions of this investigation.

Table 8 Conditions of the investigation

ANN	Number of hidden neurons	Spring-19, Summer-16, Fall and Winter-6
	Learning epochs	10
GA	Generations	1-10, 15, every 10 from 20 to 100
	Individuals	50
	Crossover rate (%)	50
	Mutation rate (%)	2

As the results of this simulation, it is clear that there is no big difference among all simulations. They all showed almost 2 degrees Celsius of minimum absolute average error (Table 9), and the differences between the minimum and the maximum are less than 0.3 degrees Celsius. Also it was shown that there are no big effects that depend on the generations of GA. Some typical simulated results for 4 seasons are shown through Figs. 12-15 (red line: forecasted curve, blue line: actual curve).

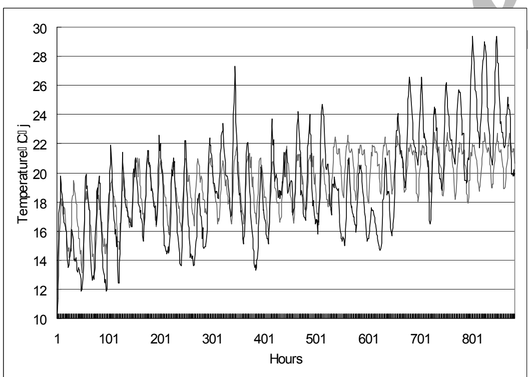

Fig. 12 A typical simulation result of spring

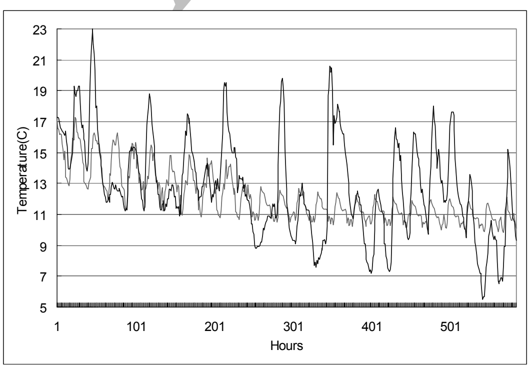

Fig 13 A typical simulation result of summer

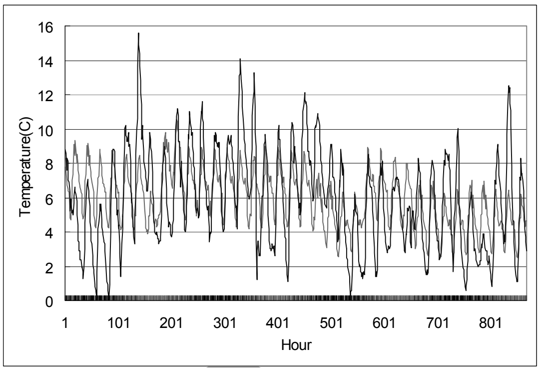

Fig. 14 A typical simulation result of fall

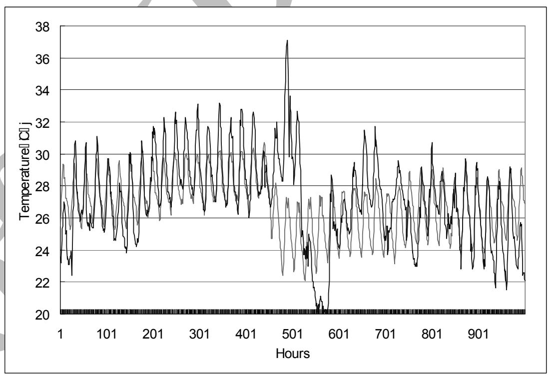

Fig. 15 A typical simulation result of winter

Still there is some inaccurate temperature forecasting results obtained from this study. They seem to be caused by the characteristics of ANN. Long learning with so many data (or learning epochs) will lower the redundancy and then it becomes difficult to forecast the flexibly. And if the learning data are not enough, it also shows inaccurate results. Therefore, in the future an algorithm is required to choose an effective amount of data for learning.

The accuracy of fall simulation looks worse than others. This is caused by the low amounts of learning data. Fall that is defined on this study is the shortest of all, and it has less than 1000 data.

However, trends of the observed temperatures that depend on the change of season seem to be followed by simulated result.

In Table 9, the averages are almost 2 degrees Celsius respectively and their average is 2.02 degrees Celsius. Table 10 shows that almost 30% of the results were below 1 degree Celsius error and almost 55% of the results were below 2 degrees Celsius error. By take a look at these results, it becomes clear that this temperature forecasting system could forecast every hour temperature up to 1 month ahead with almost 2 degrees Celsius difference.

Table 9 The rates of the error

Season	Absolute Average Error
Spring	2.11
Summer	1.80
Fall	2.15
Winter	2.02
Average	2.02

Table 10 Absolute average error of temperatures predicted in 4 seasons

	Rate (%)				
Absolute Average Error (C)	Spring	Summer	Fall	Winter	Average
0.0 – 1.0	28.1	37.5	30.1	30.5	31.6
0.0 – 2.0	54.5	66.7	55.2	56.4	58.2
0.0 – 2.5	65.4	76.3	65.2	66.0	68.2
0.0 – 5.0	94.8	94.7	93.3	96.0	94.7

4. Conclusion

in this study, it is shown that the average absolute errors of forecasted data are almost 2 degrees Celsius using a temperature forecasting system that constructed by NGL with 52 input factors. As shown above, the simulated results show some characters of this system. This system can follow the changes of temperature that depends on seasons. This makes it possible to know tendencies of temperature changes. And if the data for constructing system are less than 3000, it is difficult to catch the trend of the temperature. On the other side, still there are some points for modification of the forecasting system which may lead to accurate temperature forecasting. If these problems overcome in the future, it becomes possible to forecast temperature semi-permanently. This study shows a possibility of the long-term temperature forecasting useful for mid-term load forecasting by intelligent method.

5. References

[1] Testuhiko Takatsuka: "Weather Condition Forecasting", Seito Publishing Company, Tokyo, chapter 4, 1999.

[2] Takehiko Furukawa: "Introduction to Weather Forecasting Techniques", Ohm Publishing Company, Tokyo, chapter 3 and 4, 1998.

[3] A. Khotanzad, et al., "An Artificial Neural Network Hourly Temperature Forecaster with Applications in Load Forecasting", IEEE Trans. PWRS, Vol. 11, No.2, pp.870-876, May 1996.

[4] B. Kermanshahi, et al., "Temperature Forecasting using Artificial Neural Networks", Proceedings of ICEE2K, pp.429-432, July 2000.Trans.On Power Systems. Vol.9 .1994.