LOAD PREDICTION FOR OPTIMAL THERMAL STORAGE -COMPARISON OF THREE KINDS OF MODEL APPLICATION-

Nobuo Nakahara
Department of Architecture, Faculty of Engineering, Kanagawa University
Yokohama + 221-0802 - JAPAN
Mingjie Zheng, Song Pan and Yoshihiko Nishitani
Research Laboratory, SANKO Air Conditioning Co. Ltd.
Nagoya + 450-0003 - JAPAN

ABSTRACT

It is necessary to predict the load of the following day and hours to establish optimal thermal storage. In this paper, three kinds of load prediction models, the Kalman filter, GMDH and neural network are used and characteristics and usability of each method were compared. It has been shown that proper selection of input variables, method of preprocessing the measurement data and the form of prediction equation gave a large influence on the prediction accuracy, and that each of them could predict the cooling load for thermal storage operation with sufficient accuracy.

INTRODUCTION

The thermal storage HVAC system has the long history [1][2] in Japan, which has been used as a useful means of the economy, energy conservation, and effective use of energy. It is also paid attention as the effective means of electric power demand leveling by leveling peak cooling load for heat pumps and shifting the daily peak power demand to off-peak hours.

Improper operation of the thermal storage system often results additional heat loss from the storage tank and follow-up operation at the daytime, quite often at the very time of power peak demand. In order to get rid of these defects in storage operation, it is recommended to predict the heat load of the following day in total before beginning night-time operation and to know the hourly load after the present time till the end of operation of the day during HVAC operation.

Various load predicting models have been proposed until now. They are the linear presumption models using Kalman filter[3], to be cited as KF model hereafter, group method of data handling model[4], to be cited as GMDH model hereafter, linear regression model[5], time-series model[6] and neural network model[7], cited as NN model hereafter. Neither the performance nor the characteristics of each method, however, have been made clear. A committee in SHASE, the Society of Heating, Air-conditioning and Sanitary engineers of Japan, called as the committee for optimal thermal storage operation, planned to hold an open benchmark test on the heat load prediction, aiming at evaluating various predicting tools under the same conditions[8] concerning the number and kind of available data. The present paper reports authors' research, participating in the benchmark test, on the optimization of input variables and method of preprocessing data, applying three kinds of mathematical models, the KF, GMDH and NN models. The advantage, disadvantage, applicability and characteristics of each method have been compared. The effects due to the type of building and HVAC system and use of the weather forecast data are also discussed.

LOAD PREDICTION OBJECT

The objective buildings and HVAC systems for study, to be called as A and B hereafter, respectively, are a research laboratory and an office building. The bench mark test was executed twice in a series, Trial-1 for A building in Yokohama and Trial-2 for B building in Tokyo[8]. The HVAC system of the former is the single duct VAV air-handling unit on each floor with low temperature air supply distribution, of which data was for south-directed open office zone at the 4th to 10th typical floor. The HVAC system of the latter building is the single duct VAV air-handling unit on each floor, of which data is for office floors on the 2nd to the 9th floor, containing a bank and restaurants. Outlines of the two buildings and HVAC systems are shown in Table1 and Table2, respectively.

The data provided for the benchmark test is shown in Table3. The periods for learning and prediction of the heat load in Trial-1 are two months from June to July and August in 1996, respectively. In Trial-2, they are five months from June to October in 1995 and five months from June to October in 1996 respectively. Figure1 and Figure2 are the weather condition data provided during the Trial-1 period. The data between June 30 and July 2 are missing as shown in both figures.

Table 1 Outline of object building

	Building A	Building B
Location	Yokohama	Tokyo
space use	Research duty	Office
Structure, Scale	B1-11F, S+SRC	B1-9F, RC
Total floor area	28,841 m^2	5,404 m^2

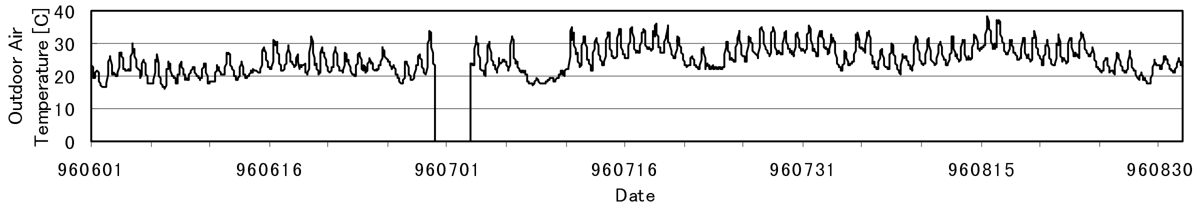


Figure 1 Outdoor air temperature in Trial-1

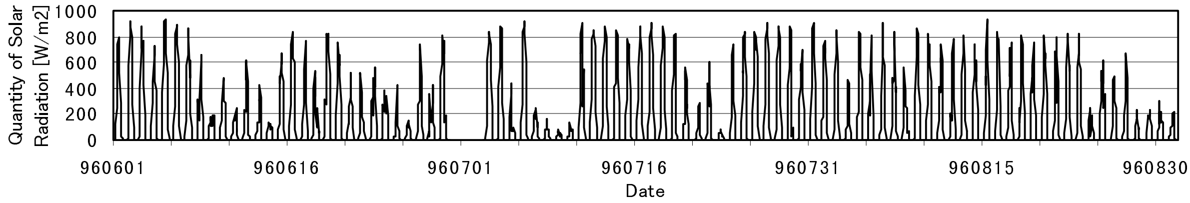


Figure 2 Quantity of solar radiation in Trial-1

Table 2 Outline of HVAC for objective buildings

item	Building A	Building B
Heat source equipment	a) Air source double bundle brine chiller/heat pump Cooling capacity 1,163 kW Heat storage capacity 721 kW Heating capacity 1,121 kW b)Water-cooled double bundle brine chiller/heat pump w/heat recovery Storage capacity 920 kW Heating capacity 1,140 kW Ice storage tank(Ice on coil type) 680 m^3 IPF 40％ Water thermal storage tank (Multi-connected complete mixing tank) 730 m^3	Two air source brine chiller/heat pump Cooling capacity 102 kW×２ Heating capacity 212 kW×２ Thermal storage 64 m^3 Summer: used as ice storage tank(Ice on coil type) IPF 8％ Winter: used as hot water tank for heating
HVAC System	Single duct VAV AHU system on each floor (low supply air temperature) Outside air intake control by CO2 concentration	Single duct VAV AHU system (partly CAV is applied) Air-handling unit + Fan coil units on each floor
Window control	Ventilation window with automatic blind control	None
Lighting control	Lighting control in perimeter due to daylighting	Ordinary lighting

Table 3 The data provided for benchmark test

item	Building A (June-August, 1996)	Building B (June-October, 1995 &1996)
Weather data	Outdoor air temperature of the day to be predicted [C] Relative humidity, ditto[%] Horizontal solar radiation, ditto [W/m^2] Wind velocity, ditto [m/s]	Forecast data for the highest and lowest outdoor air temperature at the day before [C] Forecast data for the lowest humidity, ditto[%] Horizontal solar radiation of the day before [W/m^2] Forecast weather data at that day before [degree of fine/cloudy/rainy]
The indoor environmental value	Average room temperature in each floor, ditto [C] Average humidity in each floor, ditto [%]	Average room temperature in each floor at the day before [C] Average humidity in each floor, ditto [%]
Schedule	On/off flag of AHU in each floor, ditto	On/off flag of AHU and FCU in each floor of the day to be predicted
Heat load	Heat extraction rate at AHU in each floor at the day before [kW]	Total heat extraction rate of all AHUs and FCUs at the day before [kW]

OUTLINE OF PREDICTION MODELS

(1) Neural network model, NN

Neural network is one of the pattern recognition methods by which nerve of human is imitated. Because this method has the ability to learn and to express a variety of mapping relation which includes the nonlinear one, researches on application to resolve various difficult problems have been developed up to now. Especially, as it is provided with the learning function even against a complicated system for which physical relations are not clearly defined, the application to prediction problem of the air-conditioning heating and cooling load, which is a non-linear property, are well expected.[10] The calculation algorithm of neural network used here introduces the threshold to excite the hidden layer, and the optimal correction based on the error back-propagation learning method was adopted to identify weighting vector linking neurons in the neural network. The structure of three layers of neural network is shown in Figure3. Principal equations of the NN algorithm are as follows.

a) Hidden layer output

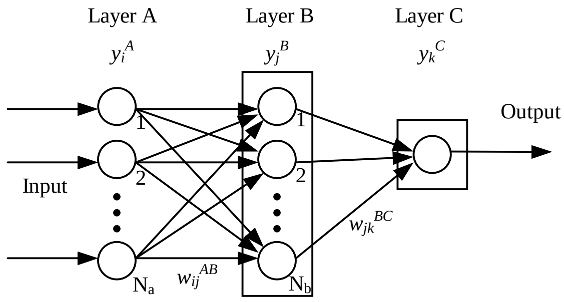


Figure 3 Structure of three layer neural network

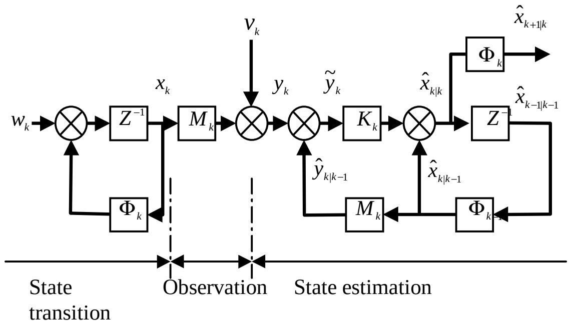


Figure 4 State estimation flow using Kalman Filter

$$y_j^B = f\left(\sum_i^{Na} w_{ij}^{AB} y_i^A - \theta_j\right) \qquad f(x) = \frac{1}{1+\exp(-x)} \qquad \text{---(1)}$$

where, θ_j is threshold value of exciting neuron and transmission function $f(x)$ in the hidden layer is a nonlinear function that is called the sigmoid function.

b) Output layer output

$$y_k^C = f\left(\sum_j^{Nb} w_{jk}^{BC} y_j^B\right) \qquad f(x) = mx \qquad \text{---(2)}$$

c) Judging convergence

The calculation ends when equation (3) is satisfied. Otherwise, the weighting vector is updated by using equation (4) to (7) and returns to equation (1).

$$\left| T_{fk} - y_k^C \right| < b \qquad \text{---(3)}$$

d) Calculating internal error at output layer

$$\delta_k = m(y_k^C - T_{fk}) \qquad \text{---(4)}$$

e) Calculating internal error at hidden layer

$$\delta_j = \delta_k w_{jk}^{BC} y_j^B (1 - y_j^B) \qquad \text{---(5)}$$

f) Updating weighting vectors at hidden and output layers

$$w_{jk}^{BC}(n) = w_{jk}^{BC}(n-1) - \alpha_{BC} \delta_k y_j^B \qquad \text{---(6)}$$

g) Updating weighting vectors at input and hidden layers

$$w_{ij}^{AB}(n) = w_{ij}^{AB}(n-1) - \alpha_{AB} \delta_j y_i^A \qquad \text{---(7)}$$

Where, T_{fk} is a teacher's signal, α_{AB} and α_{BC} are the learning coefficients. In the present calculation, the initial value of the weighting vector are arbitrarily given and the measured load of the day before was taken as the learning data, then the load of the day in concern was predicted using weighting vectors which were identified on-line through iterative calculation of seven equations shown above until the error reduces smaller than *b*. The value of weighting vectors α_{AB} and α_{BC} were identified as 10^{-8} and 1.2, respectively, as the optimal through results of prediction during one month of learning periods.

(2) Kalman Filter method, KF

The Kalman filter was developed based on the theory of the orthogonal projection by Kalman and Bucy in 1960's[11]. It is an online data processing algorithm to give the least square presumption in the state of the system value one after another by using the measured data in time series as shown in Figure 4.

The algorithm introduced hereafter is to use Kalman filter for identifying the hourly heating/cooling load as well as coefficients for variables which are chosen as significant to explain the load, simultaneously. The accuracy of this algorithm considerably depends on how to select variables, how to compose a load prediction equation, in addition to optimal selection of coefficients. Experiences of data analysis and actual application of this method by Nakahara [3] showed that the outside air-temperature greatly affects on the heating/cooling load based on the weather and that actual load of the precedent day or week before the day concerned contributes to explain either short term or weekly cycle load variations, which also explain the rate of internal heat generation based on human activities. Therefore, the weather forecast data such as the highest/lowest outside air temperature and the degree of cloudiness, actual air temperature, solar radiation and the previous day's load were taken into account. The load prediction equation is as follows.

$$\begin{aligned} l_{j+1}^* = {} & a_{j+1}^* \hat{l}_{j+1} + b_{j+1}^* g(\theta) + c_{j+1}^* K_\theta^* \\ & + d_{j+1}^* M_\theta^* + e_{j+1}^* l_j \\ & + f_{j+1}^* H_{i-1} + k_{j+1}^* H_{i-2} \end{aligned} \qquad \text{---(8)}$$

where,

$\hat{l}$: estimated load, l : observed load.

$g(\theta) = \left\| {}_D t_o + \alpha \right\|^m = \sum_{k=1}^{n} ({}_D t_o + \alpha)_{i-k}^m$:

α, m, n : constants

K_θ^* : the highest temperature of that day concerned by the weather forecast,

M_θ^* : the lowest temperature of that day concerned by the weather forecast

H : solar radiation

*: predicted value

^: estimated, or identified, value

i : time

j : day

When coefficients of load estimation equation and the load to be estimated is somehow correlated, an extended Kalman filter method in which the load and coefficients are estimated simultaneously is available. When the initial value of the coefficients a~k are assumed and the noises $\mathbf{v}, \mathbf{w}$ in the following equations (9) and (10) are assumed to follow the Gauss' distribution, the load estimation equation (8) is converted into a state equation(9) with transition matrix ϕ and an observation equation(10), abbreviating higher order terms than the first order of Taylor's expansion of the first term of equation (8).

$$\mathbf{x}_{j+1} = \phi_j \mathbf{x}_j + \mathbf{w}_j \qquad \text{---(9)}$$

$$\mathbf{y}_j = \mathbf{M}_j \mathbf{x}_j + \mathbf{v}_j \qquad \text{---(10)}$$

Where, $\mathbf{x}$: state vector of parameter

$$\mathbf{x}_j = \begin{bmatrix} \hat{l}_j & a_j & b_j & c_j & d_j & e_j & f_j & k_j \end{bmatrix}^T$$

ϕ : Transition matrix

$$\phi_j = \begin{bmatrix} a_j & \hat{l}_j & g(\theta) & K_\theta^* & M_\theta^* & l_j & H_{i-1} & H_{i-2} \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$\mathbf{M}_j = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$,observation matrix

$\mathbf{w}_j = \begin{bmatrix} w_{1j} & \cdots & w_{8j} \end{bmatrix}^T$, input noise vector,
with the mean value 0 and variance Q, in general

$\mathbf{v}$: observation noise scalar, with the mean value 0 and variance R, in general

The U-D factorization algorithm developed by Bierman-Thornton for Kalman filtering [13], which enables to shorten calculation time in a stable condition, was used to calculate equations (9) and (10).

First, initial values of $\mathbf{x}_0$ and $\mathbf{S}_0$ are given, then calculate the estimated vector $\mathbf{x}^*$ successively with the sequential equations as shown in (11)~(17) using observed, or measured, load and estimated load in sequence as shown in Figure 4, and finally the predicted load for coming day and/or hours l^* is obtained by equation (8). The [α, m, n] were decided as [30,2,1] for object A and [15,2,2] for object B, respectively, and two variances of noises, Q and R, are decided as 10^{-8} and 0.1, respectively, after several trials with measured data during leaning periods.

$$\mathbf{f} = \mathbf{S}^* \mathbf{M}^T \qquad \text{---(11)}$$

$$\beta = \mathbf{M}\mathbf{f} + R \qquad \text{---(12)}$$

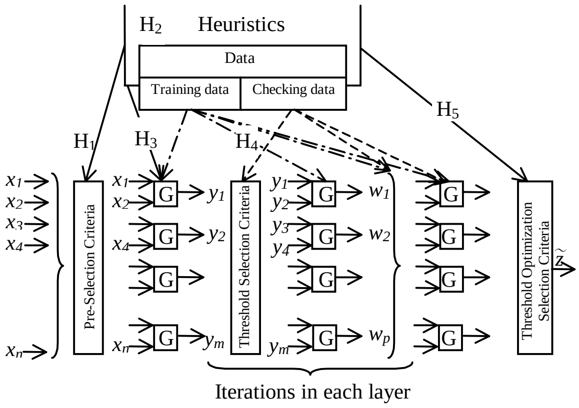
$$\mathbf{K} = \mathbf{f}/\beta \qquad \text{---(13)}$$

$$\hat{\mathbf{x}} = \mathbf{x}^* + \mathbf{K}\left[\mathbf{y} - \mathbf{M}\mathbf{x}^*\right] \qquad \text{---(14)}$$

$$\hat{\mathbf{S}} = \mathbf{S}^* - \mathbf{K}\mathbf{f} \qquad \text{---(15)}$$

$$\mathbf{x}^* = \phi\hat{\mathbf{x}} \qquad \text{---(16)}$$

$$\mathbf{S}^* = \phi\hat{\mathbf{S}}\phi^T + \mathbf{Q} \qquad \text{---(17)}$$



H1: Pre-selection of valuables (single correlation etc.)
H2: Regularization of data
(divide into two for training and checking)
H3: Partial expression structure (combination of two variables)
H4: Partial expression selection (least squares of mean error)
H5: Ending rule (beginning degradation)

Figure 5 Composition of self-organizing heuristic of GMDH

Where,
$\mathbf{S}$: Covariance of estimated error,
$\mathbf{K}$: Kalman gain

(3) Group Method of Data Handling, GMDH [14]

The GMDH based on the principle of heuristic self-Organization handles a complex system, which may be a function of quite a lot of variables and parameters of higher dimension, and the relations of each other are nonlinear, where cause and result relationships can be hardly mathematically described [15]. It is a method which combines multi-input variables and identify a definite non-linear equation consisting from significant variables selected from many candidates through self-organizing procedure. The GMDH is a concept which provides the frame of the calculation mechanism, and it has an extremely flexible structure where users can arbitrarily select basic function and uniting rule of variables, etc.[16]. It was originally developed by A.G.Ivakhnenko of former USSR to estimate grain productions based on the principle of species improvement.

The mechanism of GMDH is a repetitive calculation, which identifies the system structure and the significant parameters simultaneously as shown in Figure 5. The following procedures are necessary as the preparation phase. The Hi means the i-th stage heuristic.

a) Selection of initial input variable (H1): Calculate correlation of all variables which are thought related to output to be predicted, and choose *n* 'useful' variables highly correlated with output.

b) Normalization of data (H2): Divide data into two groups, one for training to establish model

Table 4 Input variables of each method and coefficient of correlation with measurement load

Input variables	Correlation		Neural Network		Kalman Filter		GMDH	
Object building	A	B	A	B	A	B	A	B
Heat load at the same time of the week before	0.8365	0.8364					O	O
Total number of AHU[AHU+ FCU] in operation at the same time of the day	0.8072	[0.8417]		[O]			O	[O]
Heat load at the same time of the day before	0.4588	0.6951	Teacher's signal		O		O	
Heat load at 1~9 hours before of the day			O					
Solar radiation at an hour before of the day	0.4319	0.3675					O	
Solar radiation at 1~9 hours before of the day			O					
Lowest (forecast) air temperature of the day	0.4064	(0.6621)		(O)	O	(O)		(O)
Solar radiation at two hours before of the day	0.3818	0.3597					O	
Average air temperature of the day before	0.3670	0.6683					O	O
Highest (forecast) air temperature of the day before	0.3565	(0.6422)					O	
Air temperature at the same time of the day before	0.3393	0.6090					O	O
Air temperature term of Eq.(8), g(θ)	0.2929	0.4700			O	O		
Highest (forecast) air temperature of the day	0.2788	(0.6441)		(O)	O	(O)	O	(O)
Air humidity at the same time of the day before	-0.2690	-0.2621						O
Highest (forecast) air temperature of the week before	-0.1951	(0.4660)					O	
Lowest (forecast) air temperature of the week before	-0.0538	(0.4129)						(O)
Numeric data of weather forecast at that day	-*	0.1413		O				
Estimated load at the same time of that day	-**	-**			O	O		

*: The data is not offered, **: It is not an independent variable to the prediction

Table 5 Data preprocessing for each prediction method

Method	Building	Preprocessing method	Grouping
Neural Network	A	Grouping by the day of the week	①Monday, ②Saturday/Sunday/ holiday, ③Tuesday to Friday
	B	Grouping by the total operating hours of AHU + FCU (product of number and hours of operation)	①Less than 50, ②50 to 100, ③100 to 150, ④more than 150
Kalman Filter	A	Grouping by the day of the week	①Monday, ②Saturday/Sunday/ holiday, ③Tuesday to Friday
	B	Grouping by the total operation time of AHU + FCU (product of number and hours of operation).	①less than 90, ②90 to 130, ③more than 130
GMDH	A,B	Normalization by dividing data into two groups, checking data and training data	①checking data: data of the same day of the week as to be predicted, ②training data: the rest

equation and the other for checking calculation accuracy. Several methods of normalization have been proposed.

c) Partial expression (H3): Arbitrarily chosen two input variables are combined using a basic function G(x) and obtain its output y as intermediate variable. A simple second order polynomial expression as shown in equation (18) was used as the basic function in the present study.

$$Y_k = G(x) = a_0 + a_1 x_i + a_2 x_j + a_3 x_i^2 + a_4 x_j^2 + a_5 x_i x_j \quad \text{---(18)}$$

d) Selection of intermediate variables (H4) : Put the intermediate variable y as the measured load, substitute x with the training data, obtain the normal equations by the method of least square and decide coefficients a for each variable. The checking data are then substituted into x of each partial expression and the mean square error between actual load and calculated one. The m variables (m<n) are then selected in the order of smaller error as the significant intermediate variables. This procedure is repeated until the last heuristic H5.

e) Obtaining complete expression (H5): The above procedure is ceased when the species improvement stops, that means when the minimum value of the least square exceeds the one at the precedent step, which is the beginning of degradation. The final and complete expression with the smallest least square error is the estimation, or prediction, function.

Different from the other two methods, GMDH requires many data. In the present calculation, The number of learning data which resulted in good results in estimating heat load for object A and object B was 35 days for Trial-1 84 days for Trail-2, respectively.

CONDITIONS AND ASSUMPTIONS

(1) The object buildings

Each method was applied to Building A and B.

(2) Learning and prediction Period

According to the data period decided by the test organizer, each method selected the most appropriate learning period to meet its character as the tool for load prediction. As described before, full data of the learning period was used in NN and KF application and a part of them was used for GMDH application.

(3) Input variables

The most appropriate input variables are not considered same for three methods. In order to know applicability and characteristics of each method, not only compare their precision of prediction, variables listed in Table 4 were selected to each trial, building and method as appropriate after several trials.

(4) Preprocessing

Because of the same reason as described above, preprocessing as listed in Table 6 was selected as appropriate. In addition, GMDH needs to separate data into two for normalization as described before.

(5) Criterion for evaluation

As the criterion of the prediction accuracy an expected error percentage, or EEP [7], was adopted to for relative evaluation in order to exclude the effect of building scale, type of building and HVAC system, etc. The EEP is defined in equation (19).

$$\text{EEP (\%)}=\frac{\sqrt{\dfrac{\sum_{t=1}^{n}\left(y_{pred,t}-y_{data,t}\right)^2}{n-1}}}{\left|y_{data,\max}\right|}\times 100 \quad \text{--(19)}$$

Where,

$y_{data,t}$: observed load at the time

$y_{pred,t}$: predicted load at the time,

$y_{data,\max}$: maximum load observed

n : number of load data observed

(6) Object of prediction

There are several object of load prediction for thermal storage as follows. In each case the real objects are to operate heat pumps in the highest efficiency and to maximize the daytime peak power demand shift to non-peak hours.

a) To decide operation schedule of heat pumps for thermal storage at 22 hours of previous night for the next day. Hourly load prediction is necessary, to be called day-prediction, hereafter.

b) The day-prediction can be corrected more precisely after HVAC operation begins and actual weather load data can be collected in order that any additional on-off operation of heat pumps are necessary, in comparing with remained heat stored. Hourly prediction is necessary for successive hours of the day concerned, to be called hour-prediction, hereafter.

(7) Weather forecast data

According to accommodate the data actually gathered in each case, weather forecast data was handled as follows. The weather forecast data had not been actually gathered in Trial-1 for the building A, so that the measured lowest and highest outside-air temperature of the day were considered as the forecasted lowest and highest one. On the other hand, the data was actually gathered in Trial-2 for the building B, so that the data were used for prediction, but the measured lowest and highest values were also used for comparison only, to measure availability of weather forecast.

Table 6 Comparison of predicted accuracy, EEP [%], of three kinds of methods

<table>
<tr><th></th><th>Kind of prediction</th><th>Unit for evaluation</th><th>NN</th><th>KF</th><th>GMDH</th></tr>
<tr><td rowspan="4">Building A (weather forecast unused*)</td><td rowspan="2">hour-prediction</td><td>Hourly</td><td>4.57</td><td>6.01</td><td>6.74</td></tr>
<tr><td>Dayly</td><td>7.31</td><td>6.58</td><td>6.72</td></tr>
<tr><td rowspan="2">day-prediction</td><td>Hourly</td><td>5.13</td><td>6.50</td><td>6.99</td></tr>
<tr><td>Dayly</td><td>8.76</td><td>8.05</td><td>7.04</td></tr>
<tr><td rowspan="4">Building B (weather forecast unused*)</td><td rowspan="2">hour-prediction</td><td>Hourly</td><td>5.83</td><td>6.58</td><td>7.47</td></tr>
<tr><td>Dayly</td><td>7.10</td><td>7.39</td><td>8.52</td></tr>
<tr><td rowspan="2">day-prediction</td><td>Hourly</td><td>5.83</td><td>7.29</td><td>7.77</td></tr>
<tr><td>Dayly</td><td>7.10</td><td>8.48</td><td>8.52</td></tr>
<tr><td rowspan="4">Building B (weather forecast used)</td><td rowspan="2">hour-prediction</td><td>Hourly</td><td>5.76</td><td>6.58</td><td>7.59</td></tr>
<tr><td>Dayly</td><td>7.10</td><td>7.39</td><td>8.81</td></tr>
<tr><td rowspan="2">day-prediction</td><td>Hourly</td><td>5.76</td><td>7.29</td><td>7.79</td></tr>
<tr><td>Dayly</td><td>7.10</td><td>8.49</td><td>7.66</td></tr>
</table>

* “unused” means that actual lowest/highest outside air temperature data was used on behalf of weather forecast data

CALCULATION RESULTS

Table 6 shows the result of load prediction with three methods. Figure 6 and Figure 7 show the comparison between calculation results and actual measurements of a representative week for building A and B, respectively, each includes the result with three prediction methods. The following facts are observed.

(1) Precision

The EEP is sufficiently small for practical use for optimal control/operation of thermal storage system. The difference of EEP among three methods is small, too. Detailed report of the benchmark test result by all parties can be referred to in a literature. [8]

(2) Hour-prediction and day-prediction

The hour-prediction has higher accuracy than the day-prediction, because the data used for the former includes newer and more correct data gathered an hour before at the shortest

(3) Effect of building type

Precision for building A is higher than that of building B for each calculation result. It is because load changing pattern in Building A is more steady and leaning is more effective than in building B.

(4) Effectiveness of weather forecast

The difference of calculation results between the case using weather forecast data and the case using actual weather data is small in each method, so that using weather forecast data can be said very useful.

(5) Effectiveness of preprocessing

Watching that Monday load is very well estimated and that Saturday and Sunday load, which is much

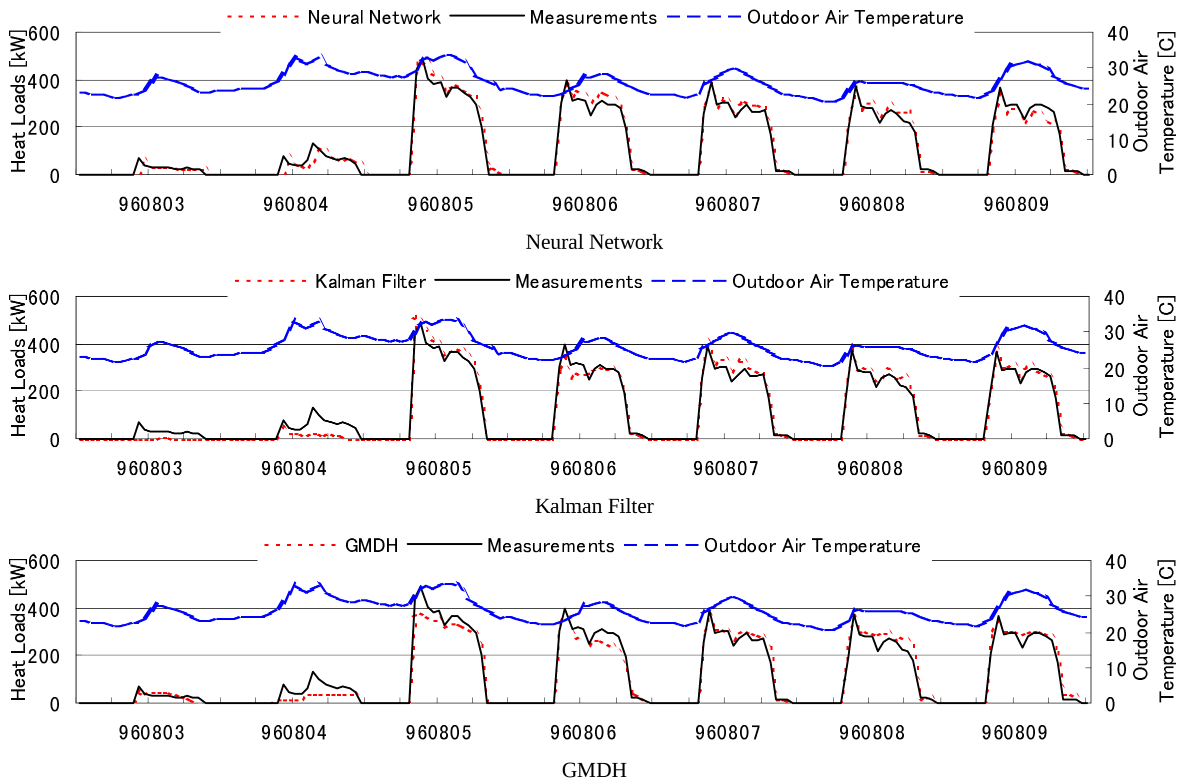


Figure 6 comparison of predicted values and measurements of three kind methods in building A

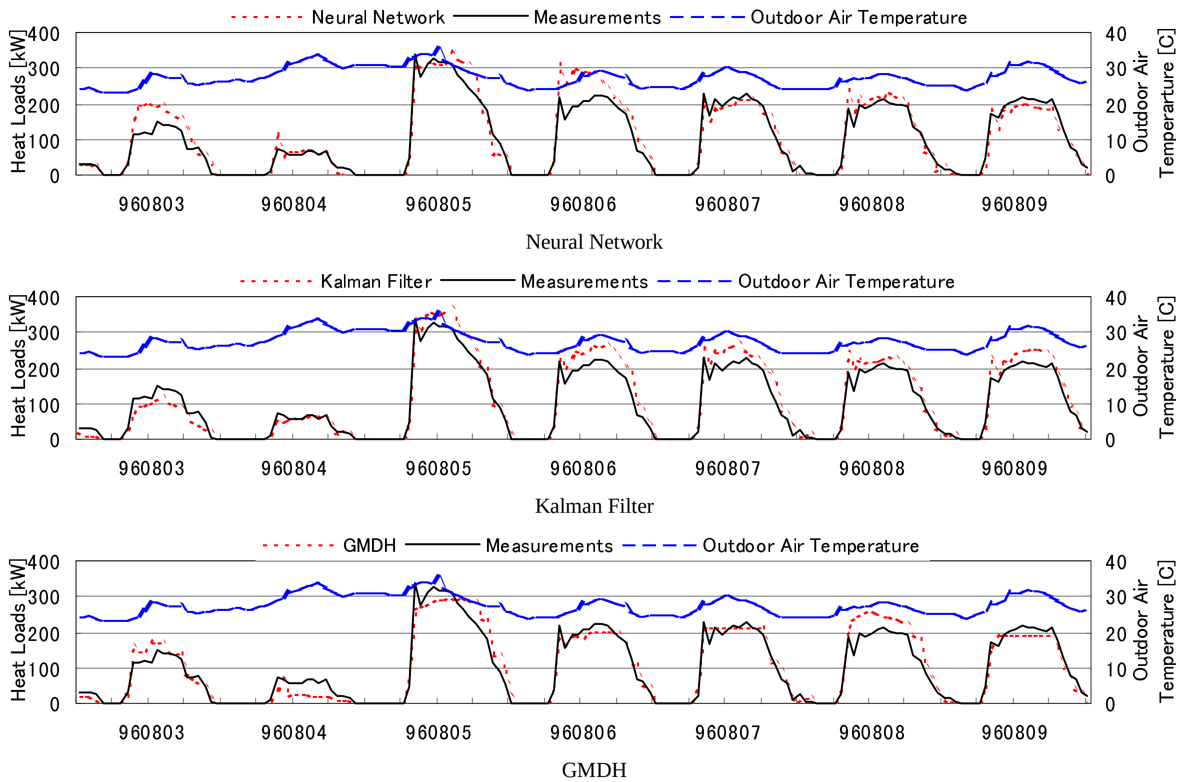


Figure 7 comparison of predicted values and measurements of three kind methods in building B

changeable, are considerably well estimated, pre-processing by grouping based on the day of the week, number of operating equipment and operating hours and normalization in GMDH as shown in Table 5 were quite effective, though any calculation results in contrary conditions are not presently shown.

DISCUSSIONS

(1) Inputs variables and preprocessing

Considering that calculation result worsens when the input variables and preprocessing way is replaced with each other, any prediction model will have its peculiar and optimal input variables and preprocessing way.

(2) Characteristics of each method

Because the NN can automatically correct the weighting vector of each layer by error back-propagation learning method, and because the GMDH can automatically optimalize selecting variables and uniting rule of each layer based on the principle of heuristic self-organization, good estimation results are expected, if appropriate initial input variables are selected. However, a physical meaning of the identified model is not clear. On the other hand, the KF, in the present application, the physical meaning of the predicting equation was clear, because the input variable and the dimension of state equation were decided beforehand according to the user's experience or know-how by user's knowledge.

Besides, when the data contains turbulence, the predicted accuracy might worsen in the NN and GMDH, because the prediction is executed using the measurement data as they are, while in the KF, because the influence of turbulence is deleted by filtering function, accuracy will not decrease so much.

(3) Flexibility of each method

As GMDH select input variables and identify coefficients using all the learning data based on the self-organizing heuristics, it can deals with changing load pattern flexibly, if the selection process is each time using newly added learning data. Formularization with Kalman filter also includes all the information in the past in its successive process though it identifies an optimal filter corresponding to the newest information. In the neural network model used here, it mainly learns the load pattern of the day before and has some unfavorable results when load pattern differently changed. Therefore, in case of Kalman filter and neural network methods, it seems necessary to group the load data according to their characteristics.

However, it should not be misunderstood that the present study is only an example of formularization of each model and it does not always fairly judge the advantage and disadvantage of three kinds of modeling. The authors feel that formularization of prediction equation and selecting way of input variables largely affects the accuracy of prediction.

SUMMARY

It is summarized that,

-According to the result of participating benchmark test on heating and cooling load prediction, the difference of the predicted accuracy among three methods was not so large, and the accuracy was satisfactory.

-It is important that formularization of load prediction equation and proper selection of input variables to meet the load characteristics of object building and HVAC system of the building presently concerned.

-It is necessary to choose the predicting model to suit the load characteristic and kinds of available data, and to choose an appropriate data preprocessing method.

ACKNOWLEDGMENTS

Data used in the present paper was prepared by the working team for the benchmark test with the kind offer of measured data by R&D Center, Tokyo Electric Power Co. and Toyo-netsukogyo Co. Authors express our gratitude to persons concerned with these projects and deeply appreciate large efforts made by Dr. Yoshida, Kyoto University and Dr. Kawashima, SHIMIZU Co. in execution of the benchmark test.

REFERENCE

1) Seinosuke Yanagimachi: Heat storage in Japan, history and evolutive development of its applied technology, and related technology to be further studied, Journal of SHASE Vol.50-9, 1976 (Japanese)

2) Yasutaka Nakajima: Historical progress in thermal storage theory and future view, Journal of SHASE, Vol.56-6, 1982 (Japanese)

3) Nobuo Nakahara: Research and development on optimizing control of Air-conditioning system, Part 4-Softwares and actual results on optimization thermal storage operation and selection of heat source, Transactions SHASE, No.18, pp.17-31, 1982 (Japanese)

4) Nobuo Nakahara, et al.: Optimizing method of HVAC and heat source system in Chita Citizen's Hospital, Part 7, Proceedings of SHASE Academic Meeting, Oct. 1984 (Japanese)

5) Yo Matsuo et al.: Basic research on heat load prediction, Proceedings of SHASE Academic Meeting, Oct. 1985 (Japanese)

6) Akiomi Kinbara et al.: Some problems on prediction for energy of HVAC systems, Proceedings of SHASE Academic Meeting, Oct. 1990 (Japanese)

7) Minoru Kawashima et al.: Study on real-time load prediction methods for HVAC system control, J. Archit. Plann. Environ. Eng., AIJ, No.484, pp.43-51, Jun. 1996 (Japanese)

8) Thermal Storage Optimization Committee of SHASE research report of 1997, Mar. 1998 (Japanese)

9) Ryuji Yanagihara et al.: HVAC system of TEPCO R&D Centers, Journal of SHASE, Vol.71-11, 1997 (Japanese)

10) Hiroshi Ichikawa: Hierarchical type neural network application to nonlinear problem analysis, Kyoritsu Shuppan, June 1993 (Japanese)

11) R.E.Kalman: A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, Mar. 1960

12) Noriyasu Sagara: A study of thermal load prediction algorithm for thermal storage system, J. Archit. Plann. Environ. Eng., AIJ, No.440, pp.11-20, Jun. 1992 (Japanese)

13) C.L.Thornton and G.J. Bierman: UDU^T covariance factorization for Kalman Filtering, Control and Dynamic Systems (C.T.Leondes, Ed.), Vol.16, pp. 177-248, 1980

14) Saburo Ikeda et al.: Base and application of GMDH Part 1-8, Systems and Control, Vol.23, 24, 1979,80 (Japanese)

15) Saburo Ikeda et al.: GMDH (Heuristic self-Organization), identification and prediction of complex system, Measurement and Control, Vol.14-2, pp.185-195, Feb. 1975 (Japanese)

16) Koji Hayashi et al.: Identification of nonlinear autoregressive model based on GMDH and its spectral analysis, Transactions SICE, Vol.28-10, pp.1216-1223, Oct. 1992 (Japanese)

17) Song Pan et al.: A study on load prediction for optimum operation of the storage system, Part 1 Investigation of load prediction method based on neural network, Proceedings of SHASE Academic Meeting, Aug. 1998 (Japanese)

18) Mingjie Zheng et al.: A study on load prediction for optimum operation of the storage system, Part 2 Investigation of three load prediction method, Proceedings of SHASE Academic Meeting, Aug. 1998 (Japanese)