Data Driven Heating Energy Load Forecast Modeling Enhanced by Nonlinear Autoregressive Exogenous Neural Networks

Jeong-A Ryu and Seongju Chang
Korea Advanced Institute of Science and Technology/Civil and Environmental Engineering, Daejeon, South Korea
Email: jaryu@kaist.ac.kr, schang@kaist.ac.kr

Abstract—**As the building sector consumes considerable portion of energy worldwide, effective management of building energy is of great importance. In this regard, forecasting building energy consumption is essential to use and manage the energy efficiently. This paper describes hourly heating energy load forecasting method with the load dataset of National Renewable Energy Laboratory (NREL)'s Research Support Facility (RSF) in the United States using both typical Artificial Neural Network and Nonlinear Autoregressive with Exogenous Inputs (NARX) Neural Network. The accuracy of the model is evaluated by MBE (Mean Bias Error) and CvRMSE (Coefficient of Variation of the Root Mean Square Error). The NARX neural network model showed a better performance than typical ANN model and it is confirmed that the model satisfies the acceptable error range proposed by ASHRAE guideline 14. This research explored a way to build a better performing neural network model for heating energy load prediction based on accumulated dataset.**

Index Terms—**Building Energy, Heating Load Forecasting, Artificial Neural Network (ANN), Nonlinear Autoregressive with Exogenous Inputs (NARX) Neural Network**

I. Introduction

Ref. [1] According to the report by the International Energy Agency (IEA), buildings account for about one-third of global primary energy consumption and about one-third of total direct and indirect energy-related greenhouse gas emissions. Therefore, efficient management of building energy is required, and the introduction of building energy management system is gradually expanding in the building sector for efficient use and management of energy. The core function of the building energy management system is to enable in-depth analysis of the various data generated in the building, and the effective linkage of individual building components, plant systems and the overall operation of the building with the analysis results. Accurate load forecasting plays an important role in building energy management systems. It enables faster and more accurate energy analysis and an effective energy utilization plan. It can affect contingency planning, load shedding, management strategies and also commercialization strategies. Ref. [2,3] Accordingly, various models for precise load prediction are actively developed and applied. Those models can be mainly divided into the statistical approaches such as ARIMA, SARIMA, and ARMAX and the artificial intelligence approaches such as ANN, SVM, and fuzzy logic. Ref. [4] Due to the ease of use and adaptability to find the optimal solutions in a rapid manner, the artificial intelligence based approaches have gained popularity in recent years.

In this paper, we developed a neural network approach using nonlinear autoregressive with exogenous input model to forecast the heating energy load. Using the power consumption data of the Research Support Facility in the US National Renewable Energy Laboratory, the artificial neural network model and the NARX neural network model were developed and compared. The accuracy of the model was evaluated by MBE and CvRMSE tolerance proposed by ASHRAE guideline14.

The outline of the paper is structured as follows. Section II describes the methodology of ANN and NARX neural network. In Section III, it presents the model development procedure and the results. Future work is discussed in Section IV and the last section concludes the paper.

II. Methodology

A Artificial Neural Network

Ref. [5, 6] ANN, which was introduced by McCulloh and Pitts, is a biologically inspired technique that models the nonlinear relationships. Neural networks have ability to solve complex relationships, adaptive control, decision making under uncertainty, and predictive patterns.

Typical ANN consists of input layer, hidden layer, and output layer, each of which is made up of neurons interconnected between different layers. Neurons are connected by the weights and the activation function converts the sum of the weighted inputs to the output. Each weight is updated minimizing the error between the generated output and the desired output mapping inputs closer to the target with a learning algorithm. Fig.1 shows the basic ANN architecture. The output of neuron can be simply expressed as (1):

$$y_i = f\left(\sum_{j=1}^{n} x_j \cdot w_{ij}\right) \tag{1}$$

Manuscript received July 26, 2018; revised April 20, 2019.

doi: 10.18178/ijscer.8.3.246-252

where i is the index of the neuron in the layer, j is the input index in the ANN, x_j is the input vector, w_{ij} is the weight vector, and f is the activation function.

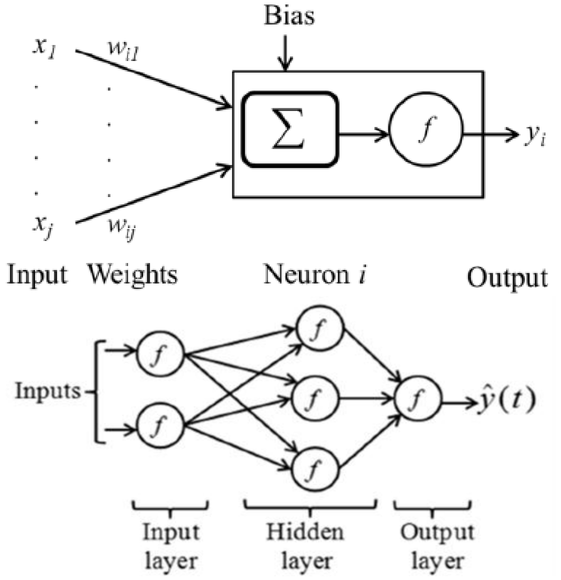

Figure 1. Ref. [7] Basic ANN Structure

B NARX(Nonlinear Autoregressive with Exogenous Inputs) Neural Network

Ref. [8] NARX, which was proposed by Leontaritis and billings, is a nonlinear model which estimates the future values of the time series based on its outputs and exogenous input. Ref. [9] It can represent wide variety of nonlinear dynamic behaviors and have been used in various applications. NARX model can be formulated as (2):

$$y(t+1) = f\begin{pmatrix} y(t), y(t-1), \dots, y(t-n_y) \\ x(t), x(t-1), \dots, x(t-n_x) \end{pmatrix} \quad (2)$$

where x is input of the network at time t, y is output of the network at time t, n_x is input delay and n_y is output delay. The next value of the dependent output signal y is estimated from previous values of the output signal and independent input signal. Ref. [9] Depending on how the function f is represented and parameterized, different NARX model structures and algorithms are derived, and neural network is used for this purpose in this study.

NARX network has two configurations which are series-parallel architecture (open-loop form) as shown in the left of Fig. 2 and parallel architecture (closed-loop form) as shown in the right of Fig. 2. In the series-parallel architecture, the true past values are used instead of feeding back the estimated output values to predict the future value of the time series. On the other hand, in the parallel architecture, output values of the NARX network are fed back to the input of the network to predict the future value. During the training phase, the series-parallel architecture is used and then after the training phase, parallel architecture is used for multistep-ahead prediction.

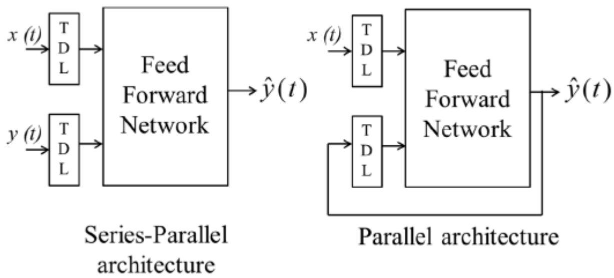

Figure 2. Ref. [7] NARX Neural Network Architecture

III. Experiments and Results

A Dataset Description

In this paper, the power consumption of the National Renewable Energy Laboratory (NREL)'s Research Support Facility (RSF) in 2011 was used as dataset for short-term load prediction. Ref. [10] The data provided by the National Renewable Energy Laboratory is based on the hourly power consumption measured by individual instruments for the purposes of heating, cooling, lighting, etc. from January 1 to December 31, 2011.

In this experiment, we used the heating energy demand which is the largest proportion of the power consumption and we used the part of the data of the winter period in which heating equipment is frequently used. As training data, the heating energy consumption, weather data (dry bulb temperature, average global irradiation) on working day (weekday) of 11/1 ~ 12/9 were used which consists of a total of 696sets and we developed separate model for one day-ahead forecasting model and one week(five days for working day)-ahead forecasting model.

Using the input processing function “mapminmax” in Matlab software, inputs and targets were normalized and scaled to be in the range [-1,1]. Then, the trained network provided outputs in the range [-1, 1] which were then reverse-processed back into the same units as the original targets. All simulations are performed in a Matlab software environment.

B Artificial Neural Network Model Development

The composition of the ANN prediction model is shown in Table I. The input variables consists of the dry bulb temperature, average global irradiation and hour of the day. Various configurations were tested by varying the number of hidden neurons (1~20) and the number of hidden layers (1~10) to find a network with optimal performance. The tangent-sigmoid and pure linear transfer functions were employed as the transfer functions for the hidden and output neurons, respectively. The feedforward network with Levenberg-Marquardt back-propagation method was used for training the developed neural network adapting weights between neurons minimizing the error between the outputs of the network and the targets. Cross-validation technique was performed to control the training process.

TABLE I. COMPOSITION OF ANN MODEL

Model Components		Contents
Structure	Input Layer	Number of neurons:3 (1) dry bulb temperature (2) average global irradiation (3) hour of the day(0~23)
	Hidden Layer	Number of layers:1~10 Number of neurons:1~20
	Output Layer	Number of neurons:1 (1) heating consumption
Transfer Function	Hidden Neurons	Tangent Sigmoid
	Output Neurons	Pure Linear
Training Method	Algorithm	Levenberg-Marquardt

To evaluate the performance of the models, following metrics are used: CvRMSE(Coefficient of Variation of the Root Mean Square Error), MBE(Mean Bias Error) are formulated as (3) and (5), respectively.

$$\mathrm{Cv}RMSE = \frac{RMSE}{M_{avg}} \times 100 \tag{3}$$

$$RMSE = \sqrt{\frac{\sum(S-M)^2}{N}} \tag{4}$$

$$MBE = \frac{\sum(S-M)}{\sum M} \times 100 \tag{5}$$

where S is estimated output, M is actual target data, N is total number of data and M_{avg} is average of M.

To find a network with optimal performance, trial-error procedure is conducted using a parametrical optimization process. The number of hidden neurons and hidden layers were sequentially optimized. After finding the optimal value for the first component (hidden neurons), the next component (hidden layers) was tested with the fixed first component as the optimal value. When finding the optimal value of the first component, hidden layer was fixed to one layer. Separate models were developed for one day-ahead forecasting model and one week-ahead forecasting model, respectively. Table II and Table III summarizes the parametrical values and the performance of the network. The former is for one day-ahead forecasting model and the latter is for one week-ahead forecasting model.

TABLE II. ANN PARAMETRICAL OPTIMIZATION(1DAY)

Hidden Neurons (Hidden layer:1)		Hidden Layers (Hidden neurons:12,optimal result)			
Hidden neurons	Cv RMSE	Hidden neurons	Hidden layers	Cv RMSE	MBE
1	50.04	**12**	**1**	**28.49**	**-2.22**
2	38.12	12	2	30.21	8.58
3	39.40	12	3	31.46	6.10
4	34.87	12	4	31.09	11.16
5	38.67	12	5	39.79	10.95
6	35.09	12	6	29.67	3.50
7	40.61	12	7	45.88	14.78
8	28.87	12	8	35.53	-0.14
9	35.38	12	9	40.27	6.17
10	34.98	12	10	31.63	-3.01
11	32.42				
12	**28.49**				
13	31.68				
14	37.48				
15	30.60				
16	30.74				
17	32.97				
18	33.44				
19	33.86				
20	33.86				

TABLE III. ANN PARAMETRICAL OPTIMIZATION(1WEEK)

Hidden Neurons (Hidden layer:1)		Hidden Layers (Hidden neurons:8,optimal result)			
Hidden neurons	Cv RMSE	Hidden neurons	Hidden layers	Cv RMSE	MBE
1	54.28	**8**	**1**	**29.70**	**-11.91**
2	40.80	8	2	36.97	-11.81
3	40.27	8	3	35.01	-8.95
4	36.21	8	4	34.60	-9.01
5	43.31	8	5	37.84	-21.41
6	35.85	8	6	34.59	-7.15
7	39.55	8	7	34.60	-13.61
8	**29.70**	8	8	37.65	-7.45
9	35.70	8	9	36.81	-13.01
10	35.59	8	10	35.96	-7.81
11	30.86				
12	31.13				
13	34.05				
14	36.38				
15	30.66				
16	33.93				
17	34.84				
18	35.30				
19	35.99				
20	33.44				

As shown in Table II and Table III, optimal number of hidden neurons and hidden layers can be found which produced the least CvRMSE between the predicted and the target value. The least value was obtained when the ANN model employed twelve hidden neurons and one hidden layer for one day-ahead forecasting model and eight hidden neurons and one hidden layer for one week-ahead forecasting model. Fig. 3 shows the example of the configuration of the developed ANN model. The CvRMSE between the predicted and the target value was 28.49 for the one day-ahead forecasting model and 29.70 for the one week-ahead forecasting model.

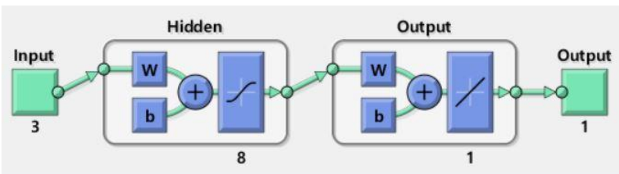

Figure 3. Matlab ANN Configuration

C NARX Neural Network Model Development

The composition of the NARX neural network prediction model is shown in Table IV. Series-parallel architecture (open-loop) is used for training phase which means that the actual target values are feedback to the network, and parallel architecture (closed-loop) is used for testing phase which means that the estimated outputs are fed back to the network. Various configurations were tested by varying the number of hidden neurons (1~20) and the number of delays (12~240) to find a network with optimal performance. Delay parameter concerns the number of hours the model is going to use to perform the prediction.

TABLE IV. COMPOSITION OF NARX NEURAL NETWORK MODEL

Model Components		Contents
Structure	Input Layer	Number of neurons:3 (1) dry bulb temperature (2) average global irradiation (3) hour of the day(0~23)
	Hidden Layer	Number of layers:1 Number of neurons:1~20
	Output Layer	Number of neurons:1 (1) heating consumption
	Delay	12~120 hours
Transfer Function	Hidden Neurons	Tangent Sigmoid
	Output Neurons	Pure Linear
Training Method	Algorithm	Levenberg-Marquardt

To find a network with optimal performance, the number of hidden neurons and delays were sequentially optimized with a parametrical optimization process. After finding the optimal value for the first component (hidden neurons), the next component (delay) was tested with the fixed first component as the optimal value. When finding the optimal value of the first component, delay parameter was fixed to 60 hours. Separate models were developed for one day-ahead forecasting model and one week-ahead forecasting model, respectively. Table V and Table VI summarizes the parametrical values and the performance of the network. The former is for one day-ahead forecasting model and the latter is for one week-ahead forecasting model.

TABLE V. NARX NEURAL NETWORK PARAMETRICAL OPTIMIZATION(1DAY)

Hidden Neuron (Hidden layer:1, Delay:60)		Delays (Hidden neurons:12,optimal result)			
Hidden neurons	Cv RMSE	Hidden neurons	Delays	Cv RMSE	MBE
1	35.73	12	12	36.54	10.75
2	32.20	12	24	28.92	13.92
3	42.70	12	36	25.51	-13.97
4	39.65	12	48	40.23	9.69
5	40.74	**12**	**60**	**18.87**	**-0.53**
6	61.57	12	72	30.95	7.24
7	49.09	12	84	37.63	-3.93
8	50.38	12	96	26.85	2.06
9	61.14	12	108	64.96	-10.43
10	37.56	12	120	24.34	-4.18
11	34.34				
12	**18.87**				
13	28.38				
14	28.55				
15	30.25				
16	31.30				
17	34.26				
18	33.79				
19	29.19				
20	40.07				

TABLE VI. NARX NEURAL NETWORK PARAMETRICAL OPTIMIZATION(1WEEK)

Hidden Neuron (Hidden layer:1, Delay:60)		Delays (Hidden neurons:12,optimal result)			
Hidden neurons	Cv RMSE	Hidden neurons	Delays	Cv RMSE	MBE
1	37.62	12	12	46.73	23.29
2	44.21	12	24	44.86	35.57
3	73.55	**12**	**36**	**27.68**	**-3.85**
4	60.46	12	48	41.74	15.68
5	43.87	12	60	37.52	14.73
6	65.68	12	72	37.30	16.20
7	60.22	12	84	74.57	40.93
8	57.76	12	96	49.09	29.30
9	113.46	12	108	66.43	-4.83
10	49.85	12	120	32.45	1.77
11	43.68				
12	**37.52**				
13	37.56				
14	46.12				
15	58.43				
16	46.73				
17	54.92				
18	50.27				
19	45.24				
20	44.90				

As shown in Table V and Table VI, the least CvRMSE value was obtained when the NARX neural network model employed twelve hidden neurons and 60 hours delay for one day-ahead forecasting model and twelve hidden neurons and 36 hours delay for one week-ahead forecasting model. Fig. 4 and Fig. 5 show the example of the configuration of the developed NARX neural network model for training phase and testing phase. The CvRMSE between the predicted and the target value was 18.87 for the one day-ahead forecasting model and 27.68 for the one week-ahead forecasting model.

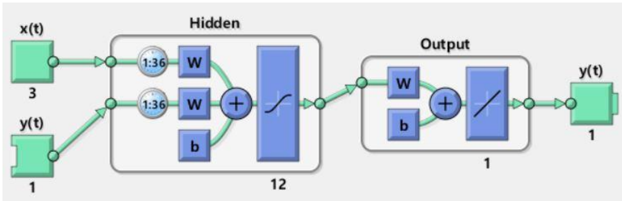

Figure 4. Matlab NARX Neural Network Series-Parallel Configuration(Training)

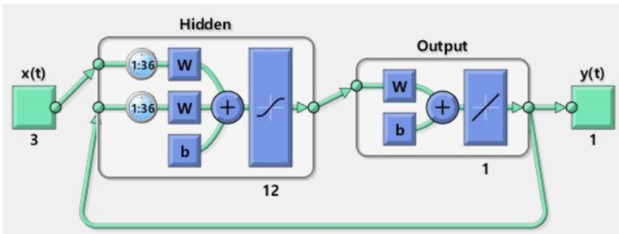

Figure 5. Matlab NARX Neural Network Pararell Configuration(Testing)

D Results

The results of the load forecasting of this study using the ANN and NARX neural network are shown in the following Table VII and Fig. 6, Fig. 7, Fig. 8, Fig. 9. The developed ANN model employed twelve hidden neurons and one hidden layer for one day-ahead forecasting model and eight hidden neurons and one hidden layer for one week-ahead forecasting model. The developed NARX neural network model employed twelve hidden neurons and 60 hours delay for one day-ahead forecasting model and twelve hidden neurons and 36 hours delay for one week-ahead forecasting model. Table VII summarizes CvRMSE and MBE of each model.

TABLE VII. RESULTS OF CALIBRATION TOLERANCE INDEX

Forecasting Period (Working Day)	Index	Artificial Neural Network	NARX Neural Network
1 day (24 hours)	CvRMSE	28.49	18.87
	MBE	-2.22	-0.53
1 week (120 hours)	CvRMSE	29.70	27.68
	MBE	-11.91	-3.85

Ref. [11] The accuracy of the model was evaluated by MBE and CvRMSE proposed by ASHRAE(American Society of Heating, Refrigerating and Air-Conditioning Engineers) as shown in Table VIII. According to the ASHRAE guideline14 (2002), the recommended values of MBE and CvRMSE for hourly measurements are within ±10% and less than 30%, respectively.

TABLE VIII. TOLERANCES BY ASHRAE GUIDELINE 14

Calibration Type	Index	Acceptable Value
Hourly	MBE	±10%
	CvRMSE	30%

For one day-ahead forecasting model, both ANN model (CvRMSE: 28.49, MBE: -2.22) and NARX neural network model (CvRMSE: 18.87, MBE: -0.53) satisfied the error range specified in Table VIII proposed by ASHRAE guideline14. Therefore, developed model can be regarded as reliable models. When comparing both models, the NARX neural network model showed a better performance than ANN model. It is easy to see that the actual and predicted data in Fig. 7 fit better than in Fig. 6.

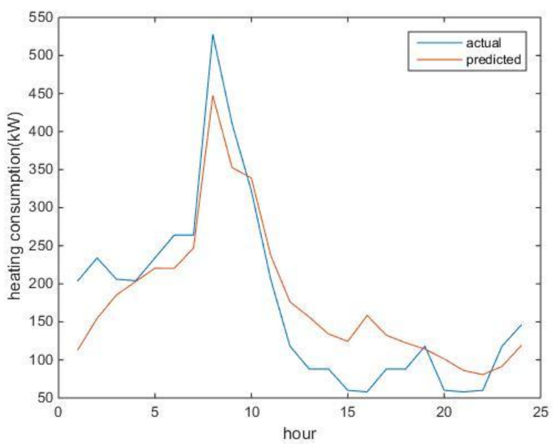

Figure 6. ANN Forecasting Results(one day-ahead)

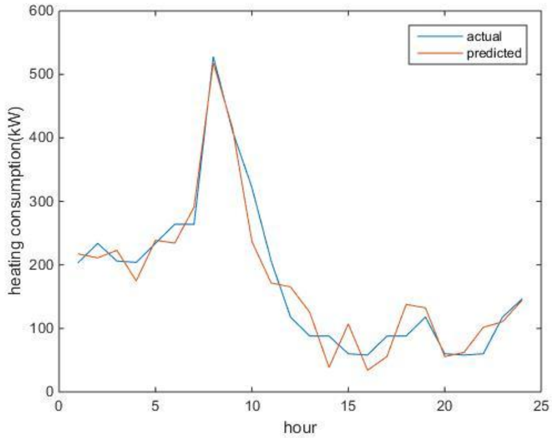

Figure 7. NARX Forecasting Results(one day-ahead)

For one week-ahead forecasting model, only CvRMSE (29.70) of ANN model satisfied the error range specified in Table VIII. However, both CvRMSE (27.68) and MBE(-3.85) of the NARX neural network model satisfied the error range proposed by ASHRAE guideline14. Therefore, developed NARX neural network model can be regarded as reliable models. When comparing both models, the NARX neural network model showed a better performance than ANN model also for one-week ahead forecasting model.

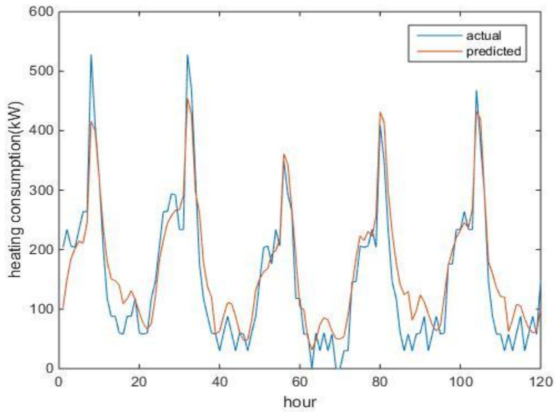

Figure 8. ANN Forecasting Results(one week-ahead)

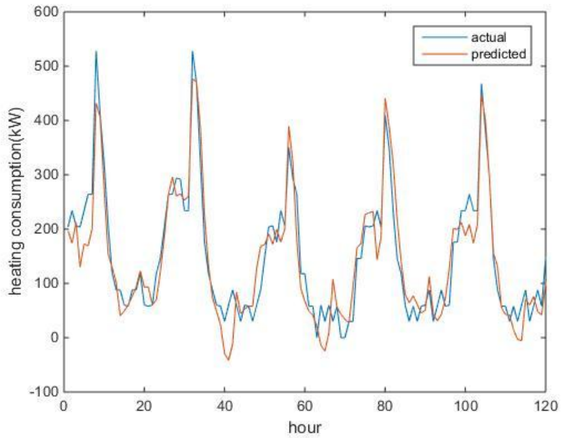

Figure 9. NARX Forecasting Results(one week-ahead)

IV. DISCUSSION

We have focused on developing a better performing neural network model for heating energy load prediction using both typical ANN and NARX neural network model. Trial-error procedure is conducted using a parametrical optimization process to determine the optimal number of hidden neurons, hidden layers and delays. The optimized parameter values were different for each model by the forecast period(one day-ahead or one week-ahead) and the developed model for the shorter term prediction showed a better performance which means that one day-ahead forecasting model (CvRMSE: 28.49/18.87) is more accurate than one week-ahead forecasting model (CvRMSE: 29.70/27.68). When comparing the two neural networks, as shown in Table VII, the NARX neural network model provide better prediction than typical ANN model as the NARX neural network model has the additional information contained in the previous values of target value which is fed back to the input.

To develop better load forecasting model, future works can be considered as follows.

1. The evolutionary optimization techniques can be considered to determine the optimal network architecture as the trial-error procedure cannot cover all cases.
2. The impact of various training algorithms on predictive models can be investigated to select the best suited algorithm.
3. Additional input factors can be explored which have great influence on the load consumption to enhance the performance of the forecast model.
4. As the size of training data has a great influence on the accuracy, optimal training data size for each forecasting period need to be explored and the best interval to update the model can be derived.
5. Ensemble models can be constructed with the combination of various AI based models compensating the weaknesses of each single model.

V. CONCLUSION

In this paper, we developed hourly heating energy load forecasting model using both typical ANN and NARX neural network. We used the heating energy consumption of the Research Support Facility(RSF) in 2011 provided by the National Renewable Energy Laboratory(NREL) in the United States. For training data, total 696sets which are the hourly data on working day(weekday) of 11/1 ~ 12/9 is used. We developed separate model for one day-ahead forecasting model and one week(five days for working day)-ahead forecasting model, respectively. Dry bulb temperature, average global irradiation, hour of the day(0~23) were used as inputs and trial-error procedure is conducted using a parametrical optimization process to determine the optimal number of hidden neurons, hidden layers and delays. For one day-ahead forecasting model, CvRMSE and MBE of ANN model were 28.49, -2.22 and those of NARX neural network model were 18.87, -0.53. For one week-ahead forecasting model, CvRMSE and MBE of ANN model were 29.70, -11.91 and those of NARX neural network model were 27.68, -3.85. One day-ahead model showed better prediction than one week-ahead model and in both cases, the NARX neural network models performed better and also satisfy the acceptable error range proposed by ASHRAE guideline 14.

The enhanced heating energy load forecast model was developed taking advantage of both neural network which can implicitly detect complex nonlinear relationships and NARX network which have additional information contained in the previous target values. The proposed energy consumption forecasting model can be used for effective building energy management system aiming of energy conservation and reduced environmental impact.

ACKNOWLEDGMENT

This research was supported by the KUSTAR-KAIST Institute, KAIST, Korea.

REFERENCES

[1] International Energy Agency, Transition to Sustainable Buildings: Strategies and Opportunities to 2050, Paris, 2013.
[2] C. Deb, F. Zhang, J. Yang, S. E. Lee, K. W. Shah, "A review on time series forecasting techniques for building energy consumption," ***Renewable and Sustainable Energy Reviews***, vol. 74, pp. 902-924, June 2017.
[3] M. A. M. Daut, et al., "Building electrical energy consumption forecasting analysis using conventional and artificial intelligence methods: A review," Renewable and Sustainable Energy Reviews, vol. 70, pp. 1108–1118, April 2017. doi.org/10.1016/j.rser.2016.12.015
[4] Z. Wang, R. S. Srinivasan, "A review of artificial intelligence based building energy use prediction: Contrasting the capabilities of single and ensemble prediction models," *Renewable and Sustainable Energy Reviews*, vol. 75, pp. 796–808, Aug. 2017.
[5] W. S. McCulloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *The Bulletin of Mathematical Biophysics*, vol. 5, no. 4, pp. 115-133, Dec. 1943.
[6] M. Q. Raza, A. Khosravi, "A review on artificial intelligence based load demand forecasting techniques for smart grid and buildings," *Renewable and Sustainable Energy Reviews*, vol. 50, pp. 1352–1372, Oct. 2015.
[7] Z. Boussaada, O. Curea, A. Remaci, H. Camblong, N. Mrabet Bellaaj, "A nonlinear autoregressive exogenous (NARX) neural network model for the prediction of the daily direct solar radiation," *Energies*, vol. 11, no.3, 620, Mar. 2018.
[8] I. Leontaritis and S. Billings, "Input–output parametric models for non-linear systems—part I: deterministic non-linear systems," *International Journal of Control*, vol. 41, pp. 303-328, 1985.
[9] E. Pisoni, M. Farina, C. Carnevale, and L. Piroddi, "Forecasting peak air pollution levels using NARX models," *Engineering Applications of Artificial Intelligence*, vol. 22, pp. 593-602, June 2009.
[10] M. Sheppy, A. Beach, S. Pless, "Metering best practices applied in the national renewable energy laboratory's research support facility. A Primer to the 2011 Measured and Modeled Energy Consumption Datasets," US, Apr. 2013.
[11] ASHRAE, *ASHRAE Guideline 14-2002 Measurement of Energy and Demand Savings*, Atlanta, 2002.

Jeong-A Ryu received the bachelor's degree in electrical engineering from Korea University, Seoul, Korea in 2012. She is currently attending a master's course in civil and environmental engineering at Korean Advanced Institute of Science and Technology (KAIST), Daejeon, Korea. Her research interests include building energy management systems and IOT applications in building systems.

Seongju Chang is an expert in the field of smart & green building systems. He is currently an invited professor of the department of Civil & Environmental Engineering at KAIST, Korea. Prof. Seongju Chang got his bachelor's and master's degrees at Seoul National University in Korea majoring in architecture and received Ph.D. degree in Building Performance & Diagnostics at Carnegie Mellon University, Pittsburgh, USA in 1999. His current research interest includes various smart and sustainable architectural/urban systems including smart building components, energy efficient and resource circulative sustainable buildings and cities. He is also listed on the Marquis Who's Who as one of the potential contributors for resolving World's problems.