CHAOS TIME SERIES MODEL FOR NONLINEAR MULTI-STEP AHEAD PREDICTION

Young Jin Kim[1], Cheol Soo Park[2], Kwang Woo Kim[3]
[1]Division of Architecture, Architectural Engineering and Civil Engineering, Sunmoon University, South Korea
[2]School of Civil & Architectural Engineering, SungKyunKwan University, South Korea
[3]Department of Architecture and Architectural Eng., College of Eng., Seoul National University, South Korea

ABSTRACT

This paper addresses a chaos time series model which can provide multi-step ahead prediction. The chaos time series model can reconstruct a deterministic and low-dimensional phase space by transforming irregular behaviors of nonlinear time-varying systems into a strange attractor (e.g. Rossler attractor and Lorenz attractor), and then it can predict multi-step ahead future states. For this study, electricity consumption measured in a building was used, and it was modeled into phase spaces with strange attractors. In the paper, it was discussed that the chaos time series model could be effectively used for multi-step ahead prediction.

INTRODUCTION

Recently, Building Energy Management System (BEMS) has been used to operate various building systems (dynamic blinds, lights, boiler, chiller, fan, pump, etc.) (IBPSA, 1987-2013). Model Predictive Control (MPC), as a part of BEMS, is based on the sensor data, a mathematical model of the systems, and optimization techniques. In general, an optimal control logic in the MPC is an iterative process to search optimal control variables which minimize a cost function over a time horizon for a given system. The time horizon usually consists of multiple sampling times. However, the system model in the MPC often fails in terms of accuracy and reliability since most models are not ready to produce multi-step ahead predictions.

The system model can be divided into a forward model and an inverse model (ASHRAE, 2013). The forward model can predict the system's behavior using the Building Performance Simulation (BPS) tools (e.g. EnergyPlus, ESP-r, and TRNSYS) which are based on the first-principles (heat transfer, thermodynamics). However, it is not easy to develop the first principles-based simulation model due to limited time and budget. In addition to the development (or mathematical formulation) of the model, calibration and validation processes are required to deal with unknown parameters and uncertainties (epistemic uncertainty, aleatory uncertainty, etc.).

In contrast, the inverse model has been accepted as a surrogate to the first-principle model due to recent development of machine learning techniques. The inverse model is constructed based on observed data and machine learning algorithms. It demands less computation time and modeling effort compared to the conventional BPS tools if measured data are available and reliable. In particular, Gaussian process model, one of the inverse models, is based on the correlation between the observed data and is able to account for the uncertainty of the model.

In general, the examples of the inverse model are as follows: Auto Regressive Integrated Moving Average (ARIMA), Artificial Neural Network (ANN), Support Vector Regression (SVR), Gaussian Process Model (GPM), etc. The aforementioned inverse models are able to produce relatively accurate outputs, but their robustness significantly depends on the quantities and qualities of the observed data. When a system is influenced by complex and dynamic interactions with other HVAC systems and when dominant time-varying data are not measured (or not available during the development of the model), the inverse model usually fails to produce multi-step ahead prediction. For example, occupants' behavior is a dominant unknown input influencing the output prediction and usually hard to be reflected in the inverse models. Hence, the inverse models usually produce probabilistic multi-step ahead outputs with high variances.

For reliable multi-step ahead prediction of the inverse model, this study presents a chaos time series model using coupling methods between the inverse models (SVR and GPM) and a chaos theory. The chaos time series model can reconstruct a deterministic and low-dimensional phase space by transforming irregular behaviors of nonlinear time-varying systems into a strange attractor (e.g. Rossler attractor and Lorenz attractor), and then it can predict multi-step ahead future states (Jiang & Li, 2005; Karatasou & Santamouris, 2010; Lu et al., 2010a; Lu et al., 2010b; Wang et al., 2011).

This study aims to compare two inverse models (SVR, GPM) to the chaos time series inverse models. The SVR and GPM have been widely used for prediction of building systems (Cam et al., 2013). In particular, the SVR and GPM models have excellent prediction capabilities, even with a few training dataset (Rasmussen, 2004; Cam et al., 2013; Kim &

Park, 2014). For this study, the electric energy consumption measured in an existing office building was used. The comparison was made focused on the multi-step ahead prediction. The predictions of the models were assessed in terms of the Mean Bias Error (MBE) and the Coefficient of Variance of the Root Mean Square Error (CVRMSE) (ASHRAE guideline 14, 2002).

CHAOS TIME SERIES MODEL

The physical phenomena of real building systems are complex and transient where a variation of the initial state of a certain system may influence overall systems. In other words, the variation of the initial state of a system may cause a critical state or tipping-point of an interconnected system, and then they have a significant effect on its performance. For example, occupants' behaviors affect control of openings (windows or doors) and systems (blinds, lights, HVAC systems, etc.). Occupants move between rooms or respond to system information or environment such as indoor temperature. Such actions influence various building system's functions as well as the total energy consumption. In other words, building systems are influenced by such occupants' behaviors and they can result in an unexpected state or tipping-point. The aforementioned phenomena can be modeled as a chaos state. However, the current inverse models (SVR, GPM) are not enough to reflect the aforementioned unexpected phenomena and may provide predictions with high variances.

The chaos theory reconstructs the complex irregular behaviors of nonlinear systems into a phase space with deterministic rules, such as a strange attractor (Jiang & Li, 2005; Karatasou & Santamouris, 2010; Lu et al., 2010a; Lu et al., 2010b; Wang et al., 2011). By coupling between the reconstructed phase space and the inverse model, the chaos time series model can be developed. According to Taken's theory (Taken, 1981), one-dimensional time series data $x(t)$ can be transformed into a state vector $z(t)$ by calculating an embedding dimension d and time delay τ as shown in Equation (1). The embedding dimension and time delay are calculated respectively using a False Nearest Neighbor (FNN) algorithm and an Average Mutual Information (AMI). The reconstructed d-dimensional phase space can be used to predict the next time horizon throughout the inverse model $F(z_t)$ as shown in Equation (2).

$$\overrightarrow{z_t} = \left[x_t, x_{t-\tau}, \ldots, x_{t+(d-1)\tau} \right] \tag{1}$$

$$\overrightarrow{z_{t+1}} = F(z_t) \tag{2}$$

The FNN algorithm proposed by Kennel et al. (1992) is used for estimating the minimal sufficient embedding dimension. When state vectors are reconstructed while a random embedding dimension is increased, the state vectors are divided into their nearest neighbors D_i^{nn} (Equation (3)) and false nearest neighbors D_i^{fnn} (Equation (4)) according to Euclidean distance of each state vector. The closer to the embedding dimension FNN ratio (R_i , Equation (5)) gets, the less FNN ratio becomes. Based on the aforementioned characteristics, the embedding dimension can be obtained using the FNN ratio.

$$D_i^{nn} = \left\| \overrightarrow{z_i} - \overrightarrow{z_j} \right\| \tag{3}$$

$$D_i^{fnn} = \left\| \overrightarrow{z_{i+1}} - \overrightarrow{z_{j+1}} \right\| \tag{4}$$

$$R_i = \frac{D_i^{nn}}{D_i^{fnn}} \tag{5}$$

The AMI developed by Fraser & Swinney (1986) is used to calculate the time delay of nonlinear data. It can quantify the dependency between the pairs of random variables through the entropy of their joint probability density function as shown in Equation (6). The AMI determines the delay time from the first minimum time.

$$I = -\sum_{ij} p_{ij}(\tau) \ln \frac{p_{ij}(\tau)}{p_i p_j} \tag{6}$$

In this study, the one-dimensional time series data are reconstructed as a d-dimensional phase space using the aforementioned FNN algorithm and AMI. The phase space is used to predict outputs of the next time horizon through a coupling with the inverse model.

TARGET BUILDING

An existing high-rise office building was selected. The building has 10 stories above ground and 3 underground levels. In this study, the BEMS data were used including total electric energy consumption at a sampling time of 1 hour during 20 weekdays in July (Figure 1). The data were divided as follows: (1) 15 days for development of the model, (2) 5 days for validation of the model.

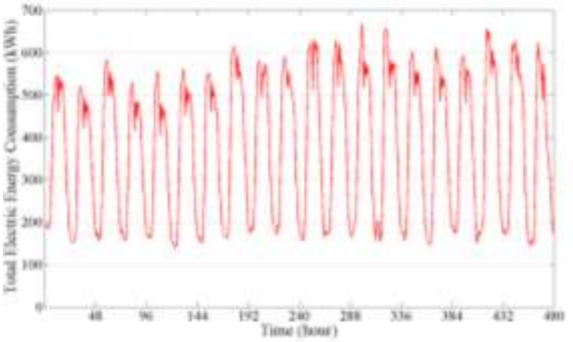


Figure 1 BEMS data

INVERSE MODEL

Support Vector Regression

Support Vector Regression (SVR) (Vapnik, 1995; Vapnik et al., 1997; He, 2008; Li et al., 2009; Niu et al., 2010; Cam et al., 2013) is used for this study. The SVR can be formulated to a linear regression $f(x)$ with a weight factor ω, a high-dimensional feature space $\varphi(x_i)$ and a threshold value b as shown in Equations (7-9). ω is expressed as a non-negative Lagrange multipliers λ, training input samples x_i and target values y_i. $\varphi(x_i)$ denotes a set of nonlinear transformations for mapping the training input samples x_i with target values y_i. The linear regression was reformulated using the Lagrange multipliers and kernel matrix $k(x_i, x_j)$ as shown in Equations (10-11).

$$f(x) = \omega \times \varphi(x_i) + b \tag{7}$$

$$\omega = \sum_{i=1}^{n} \lambda_i y_i \varphi(x_i) \tag{8}$$

$$b = -\sum_{i=1}^{n} \lambda_i y_i \varphi(x_i) \varphi(x_i) \tag{9}$$

$$f(x) = \sum_{i=1}^{n} (\lambda_i - \lambda_i^*) k(x_i, x_j) + b \tag{10}$$

$$k(x_i, x_j) = \exp\left(-\|x_i - x_j\|^2 / (2\sigma^2)\right) \tag{11}$$

To estimate the weight factor ω and threshold value b, the linear regression model is required. The linear regression model results in a quardratic programming optimization problem. In this study, the optimization problem is constrainted by a set of inequality equations with canonical hyperplans as shown in Equations (12-13). C denotes a penalty parameter which is a constant to determine the trade-off between training error and model flatness (Li et al., 2009). ξ_i denotes a slack variable used to deal with noise errors in the data. With the Lagrange multipliers λ, the optimization problem can be expressed as shown in Equation (14). To maximize the Lagrange multipliers in the cost function (Equation (14)), Equations (15-16) are obtained. It should be noted that the predictability of the SVR was influenced by estimates of three unknown parameters (C , σ , and ε). To estimate such parameters, various optimization techniques (e.g. gradient based optimization, Genetic Algorithm [GA], Particle Swarm Optimization[PSO]) can be used. In the study, the gradient based optimization was used. It has an advantage of finding an optimal solution quickly in the feasible region.

$$\arg\min_{(\omega,\xi,b)} \left\{ \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n} \xi_i \right\} \tag{12}$$

$$\begin{aligned} s.t.\quad & y_i - (\omega\varphi(x_i)) + b) \le \varepsilon + \xi_i \quad \text{for } \xi_i \ge 0 \\ & (\omega\varphi(x_i)) + b) - y_i \le \varepsilon + \xi_i \quad \text{for } \xi_i \ge 0 \end{aligned} \tag{13}$$

$$\arg\min_{(\omega,\xi,b)} \max_{(\lambda_i,\lambda_i^*)} \left\{ \begin{aligned} & \frac{1}{2}\|\omega\|^2 + C\sum_{i=1}^{n} \xi_i \\ & -\sum_{i=1}^{n} \lambda_i \left[(\omega x_i - b) - 1 + \xi_i \right] - \sum_{i=}^{n} \lambda_i^* \xi_i \end{aligned} \right\} \tag{14}$$

$$\begin{aligned} \tilde{L}(\lambda_i, \lambda_i^*) = & \sum_{i=1}^{n} y_i \left(\lambda_j - \lambda_j^*\right) - \varepsilon \sum_{i=1}^{n} \left(\lambda_i - \lambda_i^*\right) \\ & -\frac{1}{2}\sum_{i,j=1}^{n} \left(\lambda_i - \lambda_i^*\right)\left(\lambda_j - \lambda_j^*\right) k(\mathrm{x}_i, x_j) \end{aligned} \tag{15}$$

$$s.t. \quad \lambda_i, \lambda_i^* \in [0, C] \tag{16}$$

Gaussian Process Model

The GPM has been widely used for sensitivity analysis (Oakley & O'Hagan, 2004; Marrel et al., 2009), optimal design (Eisenhower et al., 2012; Kim et al, 2013), real-time optimal control (Kim et al., 2014), and retrofit analysis (Heo, 2011). Based on the BEMS data, the GPM was constructed using Gaussian Process (GP) and Bayesian inference (Kennedy & O'Hagan, 2001; Oakley & O'Hagan, 2004; Rasmussen & Williams, 2006; Marrel et al., 2009; Heo et al., 2012; Kim et al., 2013).

The BEMS data, observed during 15 days as mentioned earlier, were used. The data were classified into inputs $x = [x_1, ..., x_{d-k}]^T$ and outputs $y = [x_{1+k}, ..., y_{d+k}]^T$. With the inputs and outputs, the correlations between inputs and outputs were determined by the Gaussian Process. The Gaussian

Process is composed of a mean function $m(x_i)$ and a kernel matrix $k(x_i, x_j)$ which have a covariance function as shown in Equation (17). In general, the Gaussian Process is represented as the kernel matrix having zero mean function (Rasmussen & Williams, 2006; Kim et al., 2013), and the kernel matrix was calculated using the square exponential covariance function as shown in Equation (18). The GPM was formulated as a linear regression model with Gaussian noise as shown in Equations (19-20).

$$f(x_i) \sim gp(0, k(x_i, x_j^{'})) \quad (17)$$

$$k(x_i, x_j^{'}) = \sigma_{se}^2 \exp(-0.5\sum_{k=1}^{d} \left| x_{i,k} - x_{j,k} \right|^2 / l_k^2) \quad (18)$$

$$y_i = f(x_i) + \varepsilon_i \quad (19)$$

$$\varepsilon_i \sim N(0, v_n) \quad (20)$$

The GPM has three hyperparameters such as a scaling parameter σ_{se}^2 , a length-scale $l_{1:d}$ in the squared exponential covariance function and a variance v_n in the Gaussian noise (Rasmussen & Williams, 2006; Kim et al., 2013). Bayesian inference was used to estimate posterior distributions of the hyperparameters using a conditional probability as shown in Equation (21). In this study, the Maximum A Posteriori (MAP) estimate was chosen. The MAP estimate can provide posterior distributions of the hyperparameters using a gradient-based optimization as shown in Equation (22).

$$p(\sigma_{se}^2, l_{1:d}, v_n \left| y \right.) \propto p(y \left| \sigma_{se}^2, l_{1:d}, v_n \right.) p(\sigma_{se}^2, l_{1:d}, v_n) \quad (21)$$

$$\arg \min_{(\sigma_{se}^2, l_{1:n}, v_n)} \left\{ -\log p(D \left| \sigma_{se}^2, l_{1:n}, v_n \right.) - \log p(\sigma_{se}^2, l_{1:n}, v_n) \right\} \quad (22)$$

CHAOS TIME SERIES MODEL VS. INVERSE MODEL

To reconstruct the *d*-dimensional phase space with state vectors in the chaos time series model, the embedding dimension and time delay were calculated using the FNN algorithm and AMI, respectively. Figures 2-3 show the results of the embedding dimension and time delay. The minimum embedding dimension and time delay were set to 4 and 6, respectively. Figure 4 shows the three dimensional phase space (i.e. strange attractor) having a delay time of 6 hours, resulting from the measured one-dimensional time series data.

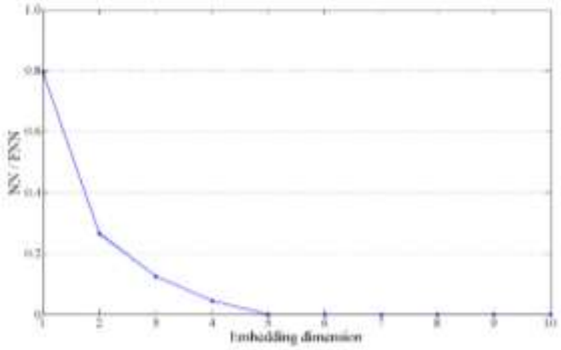


Figure 2 Embedding dimension using the FNN ratio

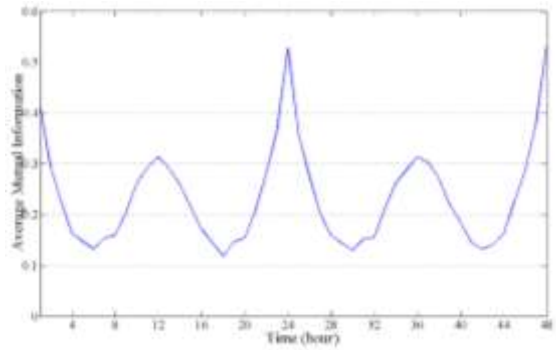


Figure 3 Time delay using the AMI

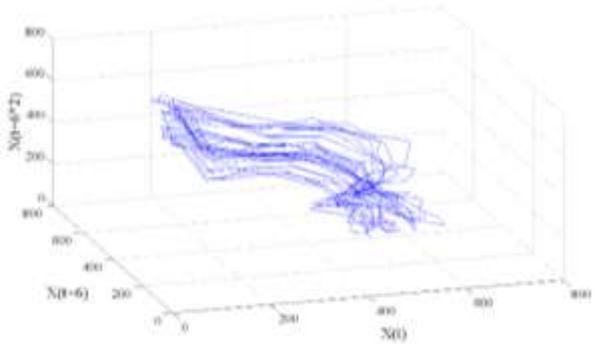
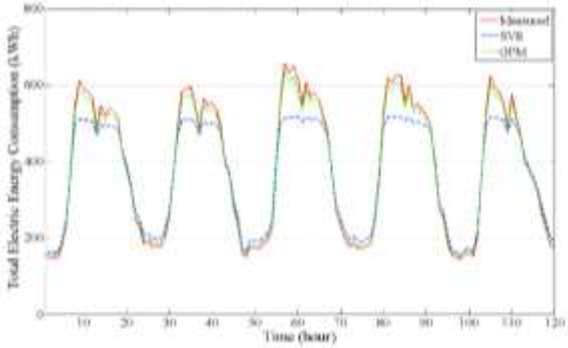
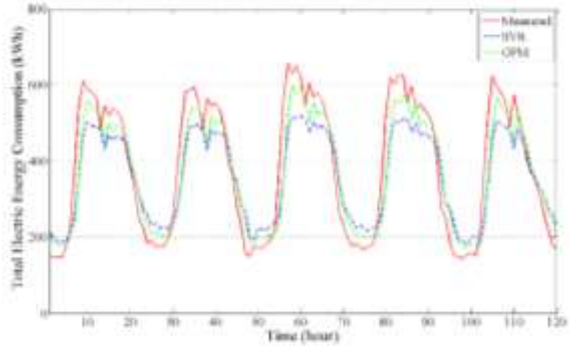


Figure 4 Three dimensional phase space obtained from the BEMS data

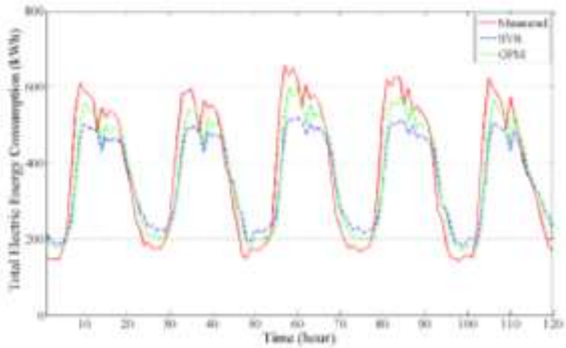
Figure 5 shows the predicted outputs of the inverse models (SVR and GPM) with varying *k*-steps (1, 2, 3, 4, 5, and 6). The prediction accuracy of the inverse models was significantly decreased when the number of steps (*k*) increased. It can be inferred that the inverse models were unsuitable for multi-step ahead predictions (e.g. for multi-step ahead MPC).



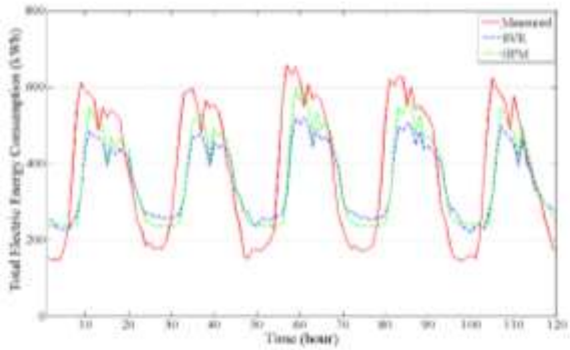
(a) 1-step ahead prediction



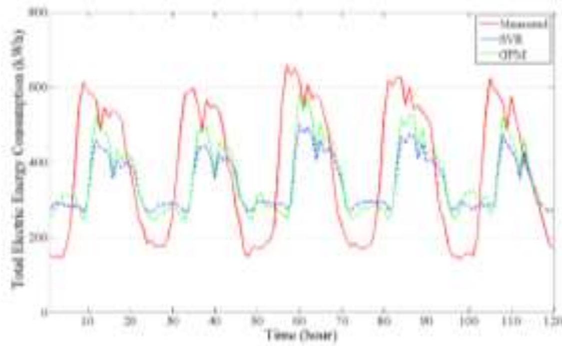
(b) 2-step ahead prediction



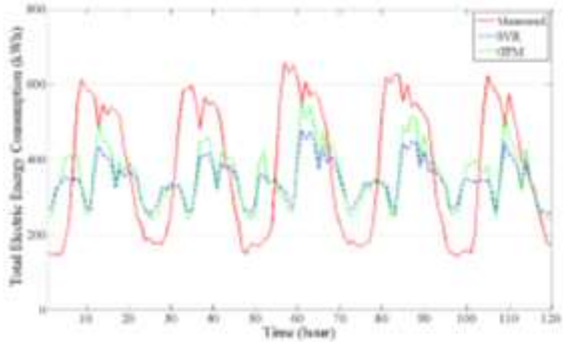
(c) 3-step ahead prediction



(d) 4-step ahead prediction



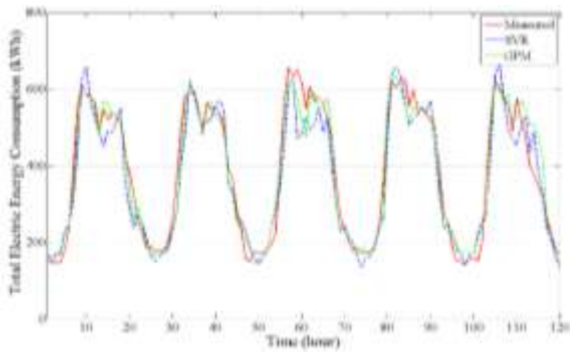
(e) 5-step ahead prediction



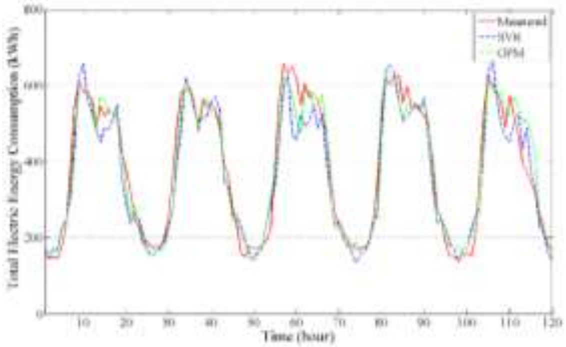
(f) 6-step ahead prediction

Figure 5 Comparison between measured data and inverse models (SVR and GPM)

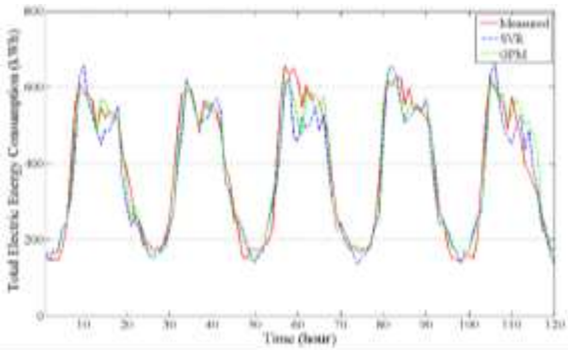
Figure 6 shows the predicted outputs of the chaos SVR and chaos GPM using state orbits of the reconstructed *d*-dimensional phase space. The prediction accuracy is good compared to the inverse models (Figure 5). The chaos time series model can provide more accurate long-term outputs than the inverse models (SVR, GPM). As noted above, the chaos time series model is advantageous to provide multi-step ahead predictions.



(a) 1-step ahead prediction



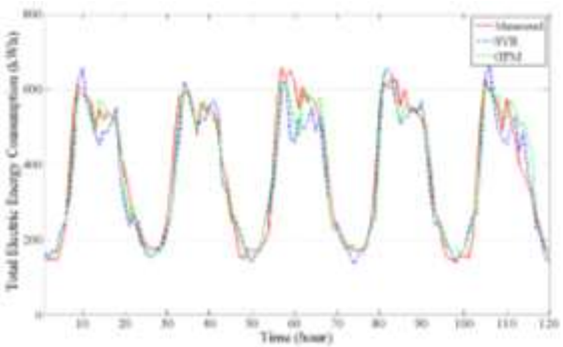
(b) 2-step ahead prediction



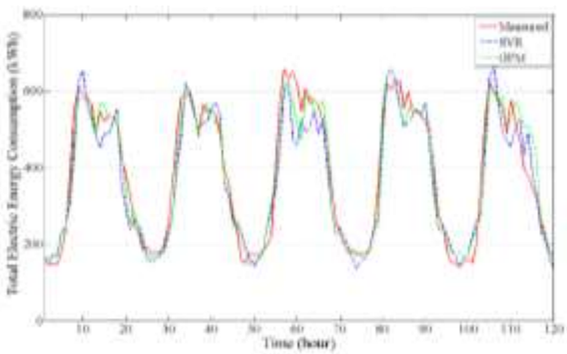
(e) 5-step ahead prediction

Table 1
Comparison between inverse models and chaos time series models using MBE and CVRMSE

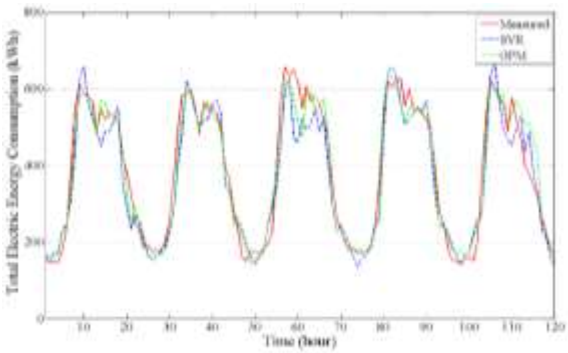
k-step	**Inverse models**				**Chaos time series inverse models**			
	SVR		**GPM**		**Chaos SVR**		**Chaos GPM**	
	MBE	CVRMSE	MBE	CVRMSE	MBE	CVRMSE	MBE	CVRMSE
1 step	4.85	12.26	1.36	3.25	4.13	13.75	0.58	10.95
2 step	6.43	21.39	3.67	17.83	4.19	13.95	0.63	11.12
3 step	8.32	31.33	5.65	29.48	4.20	13.97	0.65	11.16
4 step	10.15	38.93	7.12	37.96	4.18	13.90	0.63	11.09
5 step	10.87	43.08	7.56	42.72	4.18	13.90	0.67	11.12
6 step	10.36	43.87	6.63	43.90	4.21	13.97	0.73	11.16



(c) 3-step ahead prediction



(f) 6-step ahead prediction



(d) 4-step ahead prediction

Figure 6 Comparison between measured data and chaos time series model

Table 1 summarizes comparison of the chaos inverse models to the inverse models in terms of the MBE and CVRMSE. The predictions of the chaos time series models (from 1-step to 6-step ahead predictions) are satisfactory when assessed by ASHRAE guideline 14 (2002) (at the sampling time of 1 hour, the recommended values of MBE and CVRMSE are less than 10% and 30%, respectively). In other words, the chaos time series SVR and GPM models based on state orbits of the reconstructed *d*-dimensional phase space can provide reliable multi-step ahead predictions.

CONCLUSIONS

This study presents the multi-step ahead predictability of the chaos time series inverse models compared to the inverse models (SVR and GPM). The prediction accuracy of the SVR and GPM models decreases as the number of steps (k) increases. However, the chaos SVR and GPM models provide multi-step ahead predictions very close to the measured data.

The chaos time series inverse models are enhanced due to reconstructing the measured one-dimensional time series data into a phase space, which reduces prediction risks. In terms of the multi-step ahead prediction, the chaos time series models (SVR or GPM + chaos behavior) can accurately forecast future states of the real system which has complex and irregular behaviors. The chaos inverse models perform better than the inverse models (SVR, GPM) and can be used for real-time MPC and other control applications. Future work may include the following:

- Application of the chaos time series inverse model to a variety of building systems: the proposed chaos time series model will be applied to various building systems (lights, blinds, HVAC system, plants, etc.).
- Real-time optimal control: To implement the real-time MPC, the validated chaos time series model will be coupled with optimization techniques (e.g. genetic algorithm, particle swarm optimization) and used to find optimal control variables.

NOMENCLATURE

x_i = inputs in training dataset

y_i = outputs in training dataset

t = time

z_i = state vector in time t

d = embedding dimension

τ = time delay

D_i^{nn} = euclidean distance for nearest neighbor (nn) of the i^{th} data in the time series

D_i^{fnn} = euclidean distance for false nearest neighbor (fnn) of the i^{th} data in the time series

R_i = FNN ratio of the i^{th} data in the time series

p_i = probability density function to find a time series value in the i^{th} interval

p_j = probability density function to find a time series value in the j^{th} interval

$p_{ij}(\tau)$ = joint probability density function to find a time series value in the i^{th} and j^{th} interval

$f(x)$ = linear regression model

ω = weight factor

b = threshold value

λ = Larrange multiplier

$\varphi(x)$ = high-dimensional feature space

$k(x_i, x_j)$ = kernel matrix

ξ = slack variable

C , σ , ε = free parameter

gp = Gaussian Process

ε = Gaussian noise

σ_{se}^2 = scaling parameter

$l_{1:d}$ = length-scale parameters 1-d

v_n = variance in Gaussian noise

ACKNOWLEDGEMENT

This research was supported by a grant(11 AUDP G02) from Architecture & Urban Development Research Program funded by Ministry of Land, Infrastructure and Transport of Korean government.

REFERENCES

ASHRAE. 2002. Guideline14-Measurement of energy and demand savings, American Society of Heating, Refrigerating, and Air-Conditioning Engineers, Atlanta, GA.

ASHRAE. 2013. ASHRAE Handbook Fundamentals. Atlanta: American Society of Heating, Refrigerating and Air-Conditioning Engineers, Inc.

Cam, M.L., Zmeureanu, R., Daoud, A. 2013. Comparison of inverse models used for the forecast of the electric demand of chillers, Proceedings of the 13th IBPSA Conference, August 26-28, Chambery, France, pp.2044-2051.

Eisenhower, B., O'Neil, Z., Narayanan, S., Fonoberov, V.A., Mezic, I. 2012. A methodology for meta-model based optimization in building energy models, Energy and Buildings, Vol.47, pp.292-301.

Fraser, A.M., Swinney, H.L. 1986. Independent coordinates for strange attractors from mutual information. Physical Review, Vol.33, No.2, pp.1134-1140.

He, W. 2008. Forecasting electricity load with optimized local learning models, International Journal of Electrical Power & Energy Systems, Vol.30, No.10, pp.603-608.

Heo, Y.S. 2011. Bayesian calibration of building energy models for energy retrofit decision-making under uncertainty, Ph.D. thesis, Georgia Institute of Technology, Atlanta, GA. USA.

Heo, Y.S., Zavala, V.M. 2012. Gaussian process modeling for measurement and verification of building energy savings, Energy and Buildings, Vol.53, pp.7-18.

IBPSA. 1987-2013. Proceedings of the IBPSA (International Building Performance Simulation Association) conference ('87. '91, '93, '95, '97, '99, '01, '03, '05, '07, '09, '11, '13).

Jiang, C., Li, T. 2005. Forecasting method study on chaotic load series with high Embedding dimension, Energy Conversion and Management, Vol.46, pp.667-676.

Karatasou, S., Santamouris, M. 2010. Detection of low-dimensional chaos in buildings energy consumption time series, Communications in Nonlinear Science and Numerical Simulation, Vol.15, pp.1603-1612.

Kennedy, M.C., O'Hagan, A. 2002. Bayesian calibration of computer models, Journal of the Royal Statistical Society, Series B (Statistical Methodology), Vol.63, pp.425-464.

Kennel, M., Brown, R., Abarbanel, H. 1992. Determining embedding dimension for phase-space reconstruction using a geometrical construction, Physical Review, Vol.45, pp.3403-3411.

Kim, Y.J., Ahn, K.U., Park, C.S., Kim, I.H. 2013. Gaussian emulator for stochastic optimal design of a double glazing system, Proceedings of the 13th IBPSA Conference, August 25-28, Chambery, France, pp.2217-2224.

Kim, Y.J. and Park, C.S. 2014. Nonlinear Predictive Control of Chiller System using Gaussian Process Model, The 2nd Asia conference of International Building Performance Simulation Association, November 28-29, Nagoya, Japan, pp.594-601.

Li, Q., Meng, Q., Cai, J., Yoshino, H., Mochida, A. 2009. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networs, Energy Conversion and Management, Vol.50, pp.90-96.

Lu, X., Clements-Croome, D., Viljanen, M. 2010a Integration of chaos theory and mathematical models in building simulation Part Ⅰ: Literature review, Automation in Construction, Vol.19, pp.447-451.

Lu, X., Clements-Croome, D., Viljanen, M. 2010b Integration of chaos theory and mathematical models in building simulation Part Ⅱ: Conceptual frameworks, Automation in Construction, Vol.19, pp.452-457.

Marrel, A., Iooss, B., Laurent, B., Roustant, O. 2009. Calculations of Sobol indices for the Gaussian process metamodel, Reliability Engineering and System Safety, Vol.94, pp. 742-751.

Niu, D., Wang, Y., Wu, D.D. 2010. Power load forecating using support vector machine and ant colony optimization, Expert Systems with Applications, Vol.37, pp.2531-2539.

Oakley, J.E., O'Hagan, A. 2004. Probabilistic sensitivity analysis of complex models: a Bayesian approach, Journal of the Royal Statistical Society, Series B, Vol.66, pp.751–769.

Rasmussen, C.E. 2004. Gaussian Processes in Machine Learning, Technical report, Max Planck Institute for Biological Cybernetics, 72076 T¨ubingen, Germany.

Rasmussen, C.E., Williams, C.K.I. 2006. Gaussian Processes for Machine Learning, the MIT Press, ISBN 026218253X.

Takens, F. 1981. Detecting strange attractors in turbulence, In: Rand DA, Yong L-S, editors, Dynamical systems and turbulence, Warwick, Lecture notes in mathematics, N0.898, Berlin: Springer-Verlag, pp.366-381.

Vapnik, V.N. 1995. The nature of statistical learning theory, Springer, Berlin

Vapnik, V.N., Golowich, S.E., Smola, A. 1997. Support vector method for function approximation, regression estimation, and signal processing, Advances in Neural Information Processing Systems 9. MIT Press.

Wang, J., Chi, D., Wu, J. Lu, H. 2011. Chaotic time series method combined with particle swarm optimization and trend adjustment for electricity demand forecasting, Expert Systems with Applications, Vol.38, pp.8419-8429.