ASSESSMENT OF DIFFERENT DATA-DRIVEN ALGORITHMS FOR AHU ENERGY CONSUMPTION PREDICTIONS

Fuxin Niu[1], Zheng O'Neill[1], Wangda Zuo[2], Yanfei Li[1]
[1]Department of Mechanical Engineering,
The University of Alabama, Tuscaloosa, USA
[2]Department of Civil, Architectural and Environmental Engineering,
University of Miami, Coral Gables, Florida, USA

ABSTRACT

In this paper, four different data-driven algorithms including AutoRegressive with eXternal inputs (ARX), State Space (SS), Subspace state space (N4S) and Bayesian Network (BN) are evaluated and compared using a case study of predictions of Air Handler Unit (AHU) thermal energy consumption. Training and testing data are generated from a dynamic Modelica-based AHU model. Four evaluation metrics of Root Mean Squared Error (RMSE), coefficient of determination (R2), Normalized Mean Bias Error (NMBE) and Coefficient of Variation of the Root Mean Square Error (CV-RMSE) are used to compare the model prediction performance of different algorithms. The best algorithm is selected and proposed following the criteria recommended by ASHRAE Guideline 14. Using the proposed data driven algorithm, the relation of AHU energy consumption with mixed air temperature, air flow rate, and supply water temperature are obtained. In the future, such correlations will be employed for an optimization analysis of AHU energy consumption.

INTRODUCTION

Accurate energy performance predictions of Heating, Ventilation and Air-conditioning (HVAC) system play a significant role for building energy system performance optimization (e.g., optimal controls) to reduce energy consumption. Currently, there are mainly two kinds of prediction models: physics based model prediction and data driven based model prediction (Zhao, 2012). Physics based model prediction uses physical principles to present an entire or sub-system inherent thermodynamic performance. It can indicate the heat and mass transfer process in all individual components. Nowadays, the complex building energy performance modeling can be conducted using simulation programs such as DOE-2, EnergyPlus, Modelica, TRNSYS and Simulink (Ma et al. 2011, Candanedo et al. 2011, Henze et al. 2005, Morosan et al. 2010, and Karlsson et al. 2011). Physics based modeling in general is a time-consuming and labor intensive method. In addition, sometimes, it also needs high processing device to compute linear or nonlinear, steady or dynamic mathematic model. Even so, precise prediction using physics based models is not easy to achieve due to assumptions and uncertainty associated with input variables.

In actual HVAC systems, large amounts of raw data are monitored, trended and saved in Building Automation System (BAS). In fact, such data represents the performance of building energy system, which includes inherent information and relation of each subsystem or component of building. Therefore, how to utilize these building data to do energy performance analysis, energy diagnosis, and system operation optimization is needed. Researchers started to use data driven algorithms to explore historical performance data from BAS for energy performance prediction. The data-driven approach overcomes the drawbacks of physical modeling and has been applied for energy consumption prediction and fault diagnosis in buildings. Kusiak et al. (2010) built prediction models based on real test data using multiple-linear perceptron (MLP) algorithm for a chiller, a pump, a fan, and a reheat device. An energy optimization model integrated these four models with two dominant variables of the supply air temperature setpoint and the static pressure setpoint. The optimization results indicated that a 7.66% energy saving can be achieved. An artificial neural network (ANN) based data-driven approach ensembling with five multi-layer perceptron performed the best among tested data-mining algorithms and therefore was selected for the prediction model (Kusiak, et al. 2010). Outdoor air temperature and relative humidity were used as input variables of the prediction model. Practical archetype of building operation data was collected by University of Iowa for training and testing. Most of the existing ANN models for building energy prediction are static, while this research evaluates the performance of adaptive ANN models that are capable of adapting themselves to unexpected pattern changes in the input data, and therefore can be applied to the real-time and on-line building energy prediction (Yang et al. 2005). Two adaptive ANN models are proposed and tested in Yang's study: accumulative training and sliding window training. Yun et al. (2010) predicted building hourly thermal load using an ARX435 model. Results indicated that the ARX model was more accurate than the models of MLR, AR and ANN in that case. Yoshida et al. (2001) used ARX model in the fault detection technique for an off-line analysis. The performance of a VAV sub-system was modelled using an ARX model created from actual building data. It was concluded that the off-line analysis of data by this model was likely to detect most of faults such as problems in damper control. Ma et al. (2011) used ARX models to mimic the behaviour of EnergyPlus models. The model inputs were temperature setpoints for each zone and outputs were actual zone temperature and power measurements. The optimization problem with economic objective and several constraints was transformed into a linear programming and solved in each time step. It was shown by a continuous weekly simulation that for a smart grid with time-based pricing, the proposed method could bring substantial cost savings by automatically triggering pre-cooling effect and shifting the peak demand away from on-peak period. Cai and Braun (2013) introduced the theory of state space sub-space (N4S) identification and applied it to a single wall and a multi-zone building. Performance comparisons of N4S and gray-box models showed that N4S had a high accuracy and was computationally efficient. However, N4S prediction method requires more training data and the data with sufficient operation range to guarantee the model prediction accuracy.

Both linear regression and Gaussian process regression were used to develop an inverse model for a building case (Zhang et al. 2013, Zhang et al. 2015). From the comparisons, it is concluded that the inverse modeling is a realistic and efficient way to estimate the baseline building performance in the post-retrofit phase. The Gaussian Process (GP) approach leads to a highly flexible model, which can easily capture the complex building behaviour. It can provide more realistic results compared with linear regression model. The predictive quality of the GP model is strongly influenced by the range covered by the training and testing data set. O'Neill (2014) presented the development of a data driven probabilistic graphic model to predict building HVAC hot water energy consumption. A directed graphical model, namely, a Bayesian Network (BN) model was created for such a purpose. Each node in the BN represents a random variable such as outside air temperature, energy end usage, etc. The links between the nodes are probabilistic dependencies among these corresponding variables. These dependencies are statistically learnt and/or estimated by using measured data and augmented by domain expert knowledge. The prediction result by BN indicated that it was acceptable prediction method while providing more information such as uncertainty associated with predictions.

Many other data driven algorithms are being explored in the field of building energy performance prediction and operation managment. Each algorithm has its own advantages and disadvantages. Choosing an appropriate method for a specific case is critical to guarantee a successful energy operation management in buildings. Currently, there is a lack of research work on assessment of different data driven algorithms using the same data set. In this paper, four commonly used data driven algorithms of ARX, SS, N4S and BN are evaluated and compared with the criteria recommended by ASHRAE Guideline 14 (ASHRAE 2002).

We first briefly introduce the fundamentals of four data-driven algorithms. Then, a Modelica-based dynamic model will be presented for data generation. This will be followed by results and conclusions.

METHODOLOGIES

Autoregressive with external inputs (ARX)

The most used model structure for ARX model is the simple linear difference equation:

$$ARX : A(q)y(t) = B(q)u(t-nk) + e(t) \qquad (1)$$

Where *n* is the number of time-step delay transferring from input *u(t)* to output *y(t)*, *e(t)* is error. *A(q)* and *B(q)* are polynomials in terms of the time delay operator (q^{-1}) given in Equation 2 and Equation 3.

$$A(q) = 1 + a_1 q^{-1} + a_2 q^{-2} \ldots\ldots + a_{na} q^{-na} \qquad (2)$$

$$B(q) = 1 + b_1 q^{-1} + b_2 q^{-2} \ldots\ldots + b_{nb} q^{-nb-1} \qquad (3)$$

The example of ARX with shift operator polynomials is as follows:

$$\begin{aligned} ARX : y(t) + a_1 y(t-1) + \ldots + a_{na} y(t-na) = \\ b_1 u(t-nk) + \ldots + b_{nb} u(t-nk-nb+1) + e(t) \end{aligned} \qquad (4)$$

This ARX model relates the current output *y(t)* to a finite number of past output *y(t-na)* and input *u(t-nb)*. The structure is entirely defined by three integers *na*, *nb*, and *nk*. *na* and *nb* are the order of input and output, while *nk* is the pure time-delay in the system. For a system under sampled-data control, typically *nk* is equal to 1 if there is no dead-time.

State space (SS)

The state space of a dynamical system is the set of all possible states of the system. The values of all the state variables completely describe the system state. In other words, each point in the state space corresponds to a different state of the system. State-space models are models that use state variables to describe a system by a set of first-order differential or difference equations, rather than by one or more n^{th}-order differential or difference equations. The SS model has been successfully applied in engineering, statistics, computer science and economics to solve a broad range of dynamical systems problems. State variables *x(t)* can be reconstructed from the measured input-output data, but are not themselves measured during an experiment.

$$\dot{x}(t) = \mathrm{A}x(t) + \mathrm{B}u(t) \tag{5}$$

$$y(t) = \mathrm{C}x(t) + \mathrm{D}u(t) \tag{6}$$

The internal state variables are the smallest possible subset of system variables that can represent the entire state of the system at any given time (Nise, 2010). The minimum number of state variables required to represent a given system, is usually equal to the order of the system's defining differential equation.

Subspace state space (N4S)

Subspace state space can be derived from state space form (Cai and Braun, 2013). Assuming state space form with noise terms can be described in the following.

$$\dot{x}(t) = \mathrm{A}x(t) + \mathrm{B}u(t) + \omega(t) \tag{7}$$

$$y(t) = \mathrm{C}x(t) + \mathrm{D}u(t) + v(t) \tag{8}$$

Defining $Y_r(t) = \begin{bmatrix} y(t) \\ y(t+1) \\ \vdots \\ y(t+r-1) \end{bmatrix}$; $U_r(t) = \begin{bmatrix} u(t) \\ u(t+1) \\ \vdots \\ u(t+r-1) \end{bmatrix}$;

$W_r(t) = \begin{bmatrix} w(t) \\ w(t+1) \\ \vdots \\ w(t+r-1) \end{bmatrix}$,

Therefore, the state space equation is transformed:

$$\begin{aligned} y(t+k) &= \mathrm{C}x(t+k) + \mathrm{D}u(t+k) + v(t+k) \\ &= \mathrm{CA}^k x(t) + \left[\mathrm{CA}^{k-1}\, \mathrm{CA}^{k-2} \cdots \mathrm{C}\right] \times \\ &\quad (\mathrm{B}U_k(t) + W_k(t)) + v(t+k) \end{aligned} \tag{9}$$

Stacking and abbreviating the equation, the final model is obtained:

$$Y_r = O_r X + S_r U + V \tag{10}$$

Where $O_r = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{r-1} \end{bmatrix}$ $S_r = \begin{bmatrix} D & 0 & \cdots & 0 & 0 \\ CB & D & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ CA^{r-2}B & CA^{r-3}B & \cdots & CB & D \end{bmatrix}$

Bayesian network (BN)

A BN is a graphical model that encodes probabilistic relationships among variables of interest. When used in conjunction with statistical techniques, the graphical model has several advantages for data analysis. First, because the model encodes dependencies among all variables, it readily handles situations where some data entries are missing. Second, a BN can be used to learn causal relationships, and hence it can be used to gain understanding of a problem and to predict the consequences of intervention. Third, because the BN has both a causal and probabilistic semantics, it is an ideal representation for combining prior knowledge (which often comes in causal form) and data. Fourth, Bayesian statistical methods in conjunction with Bayesian networks offer an efficient and principled approach for avoiding the overfitting of data (Heckerman, 1997). A BN is a probabilistic graphical model that represents a set of random variables and conditional dependence through a direct acyclic graph (DAG). In the graphical model, the node that causes another node is called a parent and the affected node is called its child. The child is conditioned by the parent. Given *A* is a parent and *B* is a child of *A*, the probability of *B* conditioned by *A* is noted *P(B|A)*. Bayes theorem describes probabilistic dependencies between *A* and *B* as follows (Jensen, 2001):

$$\mathrm{P(B|A)} = \frac{\mathrm{P(A|B)P(B)}}{\mathrm{P(A)}} \tag{11}$$

For continuous variable, the conditional probability distribution will obey the normal distribution regulation $N(\mu, \sigma^2)$. The probabiltiy formula is shown in Eq. (12). The BN approach can encode the background knowledge of the system as prior expert knowledge and also discover new relationships within data streams using structure learning algorithms. Thus this approach would allow leveraging of both system domain knowledge and statistical data-mining algorithms.

$$P(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{\frac{-(x-\mu)^2}{2\sigma^2}} \tag{12}$$

It is common to consider BNs consisting of nodes with discrete variables (Laur, 2009). If a distribution is continuous, then marginalisation becomes a challenging problem. The marginal distribution of a subset of a collection of random variables is the probability distribution of the variables contained in the subset. It gives the probabilities of various values of the variables in the subset without reference to the values of the other variables. Although many families of exponential distributions are known to be closed under the multiplication, the multivariate normal distribution is also closed under the marginalisation. To have a multivariate normal distribution $N(\mu, \Sigma)$, first, we need to derive formulas for multiplication and marginalisation of normal distributions. Then, we need to verify that when the conditional distributions are specified through a normal distribution. In this case study, all output and input variables are continuous. Therefore, BN network with continuous variables will be utilized to predict AHU energy consumption.

DATA GENERATION

In this case study, we are using simulations to generate data for training and testing different data-driven algorithms. Dymola with a friendly interface based on Modelica programming is applied to simulate the

AHU system with air-side economizer. The LBNL Modelica Buildings library (Wetter et al. 2014) is a free open-source library with building envelope and HVAC system/component models. The library contains models for air-based HVAC systems, chilled water plants, water-based heating systems, controls, heat transfer among rooms and the outside and multizone airflow. This library also includes natural ventilation and contaminant transport. Fig.1 shows the Dymola simulation schematics of the AHU system, which mainly has an air-side economizer with a damper control and an air-to-water heat exchanger with a valve control. Supply air temperature was controlled at the setpoint of 50 ºC using a PI controller. Three variables of supply water temperature, mixed air temperature and mixed air mass flow rate are used as inputs for a data-driven model to predict water thermal energy consumption.

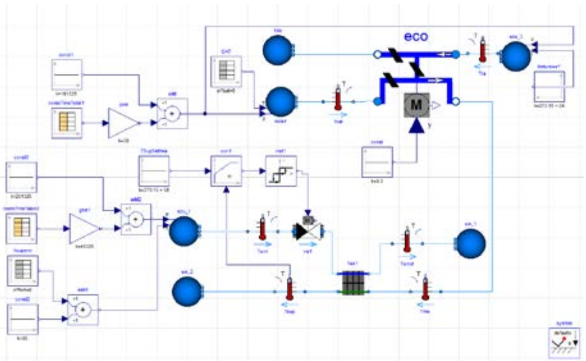

Fig.1 Diagram of Modelica models for AHU system

In order to get sufficient data for covering entire operation condition, excitation signals to outdoor air pressure, return air pressure and supply water pressure are needed. A common excitation signal was generated by Simulink program as presented in Fig. 2. The excitation signal, as shown in Fig.3, was random distribution in a range between 0.4 and 1. Based on the signal, outdoor air pressure and return air pressure are defined using Equation (13). Therefore, the pressure difference of air sources and sinks can provide sufficient excitation data to train and test different prediction algorithms. For the water side, supply water pressure is given by Equation (14). The water pressure difference is 40325ω Pa, which can guarantee the sufficient excitation range. ω is the excitation signal.

$$P_{oa,ra} = 101325 + 30 \cdot \omega \quad \text{(Pa)} \tag{13}$$

$$P_{w\text{-}supply} = 201325 + 40325 \cdot \omega \quad \text{(Pa)} \tag{14}$$

Mixed air temperature and supply water temperature after excitations are presented in Fig.4. The supply water temperature is varied from minimum value of 70 ºC to peak value of 95 ºC. The mixed air temperature is varied from 16 ºC to 21 ºC. Fig. 5 shows the mixed air mass flow rate. Because of the excitation signal to outdoor and return air pressure, the mixed air mass flow rate is ranged from 10 kg/s to 70 kg/s.

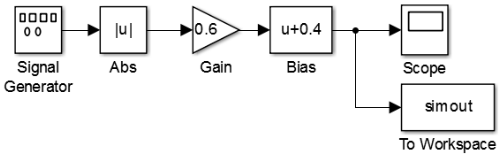

Fig.2 Excitation signal generation Simulink program

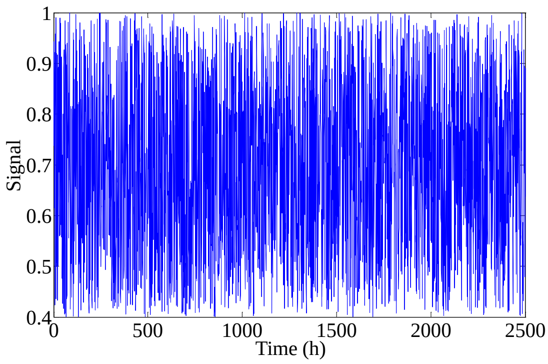

Fig.3 Excitation signal

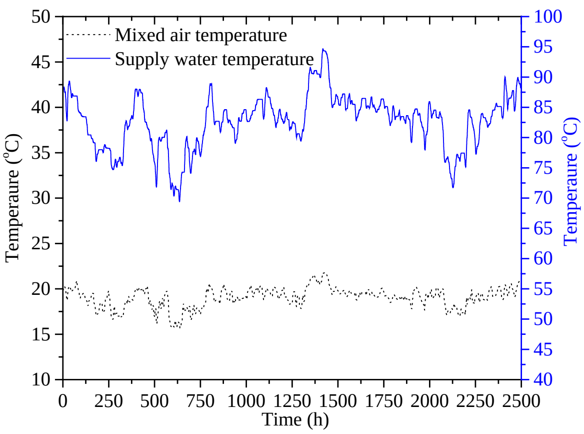

Fig.4 Supply water and mixed air temperatures

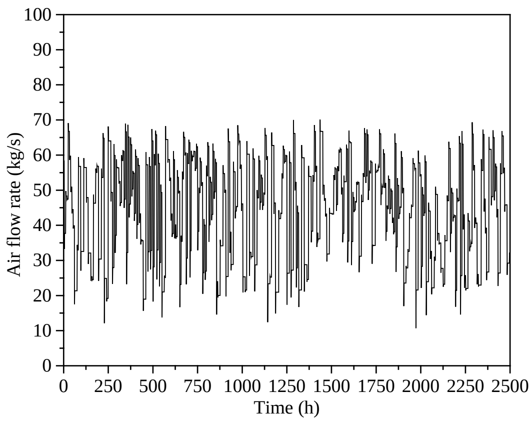

Fig.5 Mixed air mass flow rate

Prediction output variable water energy consumption from the simulation is shown in Fig.6. Enough excitation can satisfy the prediction requirement from 500 kW to 2000 kW.

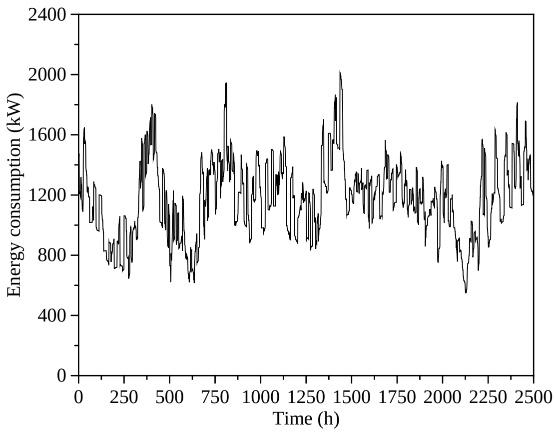

Fig.6 Water energy consumption from Dymola Simulation

For BN algorithm, the structure of direct acyclic graph affects the prediction result. In this case, three input variables of mixed air temperature, mixed air mass flow rate and supply water temperature is considered to be independent each other. Therefore, the BN direct acyclic graph for energy consumption prediction is defined as in Fig.7.

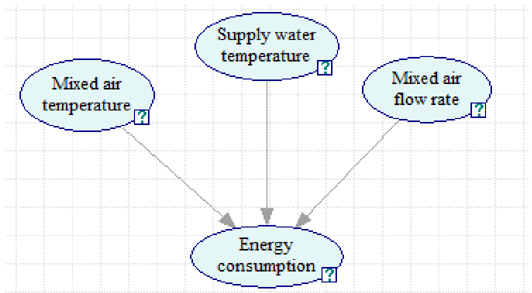

Fig.7. Bayesian network direct acyclic graph for energy consumption prediction

RESULTS AND DISCUSSIONS

Firstly, the training data covering 2,500 hourly data points, as shown in Fig. 4, 5 and 6, was used to train the prediction models. Then, hourly testing data of totally 570 hours was applied to test the accuracy of the trained model for energy consumption predictions. All algorithms were implemented using Matlab.

Of all ARX prediction models, *na* and *nb* are the orders of input and output which are suggested in the range of 1 to 10, while *nk* is the pure time-delay in the system. In this case study, there is no dead time. Therefore, *nk* is equal to 1. Of all possible ARX models, ARX 3-3-1 (i.e., *na* and *nb* are both equal to 3) had the best performance and was utilized to predict energy consumption. According to results shown in Fig.8, the big error happened when the actual energy consumption was at either high or low conditions. The absolute prediction error is 600 kW when the actual energy consumption is at the highest value of 1750 kW with a relative error of 34.3%. The absolute error is 640 kW when the actual energy consumption is at the lowest value of 860 kW with a relative error of 74.4%. ARX model failed to predict water thermal energy consumption in this case study.

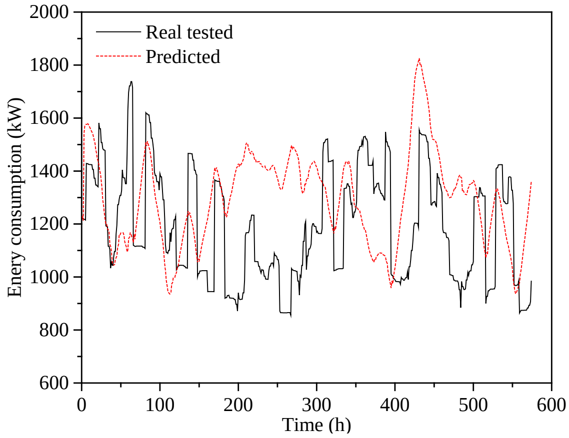

Fig.8 Energy consumption prediction by ARX model

Fig.9 and Fig.10 present the prediction results of SS and N4S models, respectively. The tested and predicted curves have the same tends. In addition, the prediction results from both models are similar. There is very little difference between SS and N4S prediction results. However, a critical issue is time-delayed prediction compared with tested value. Three or four hours delay existed for SS and N4S prediction. The prediction results will be marginally accepted if the prediction objective only focuses on total building energy performance auditing. However, the prediction performances do not satisfy the requirement if the objectives are for real-time building operation control and management. There is a need for more training data spanned over the full range of operation to improve data-driven model prediction accuracy. On the other hand, continuous BN model has the best prediction performance as presented in Fig.11. Comparing with the above three prediction methods, BN model has the least error and no time-delay prediction phenomenon. Fig.12 shows the comparisons of the tested and predicted energy consumption from BN model. The dark dash lines are ±15% error lines. Almost all of the predictions are with ±15% error band.

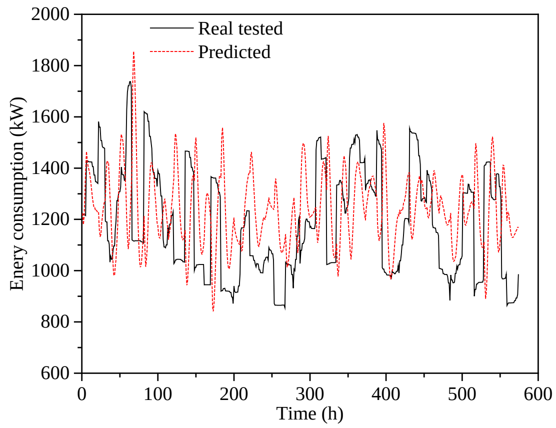

Fig.9 Energy consumption prediction by SS model

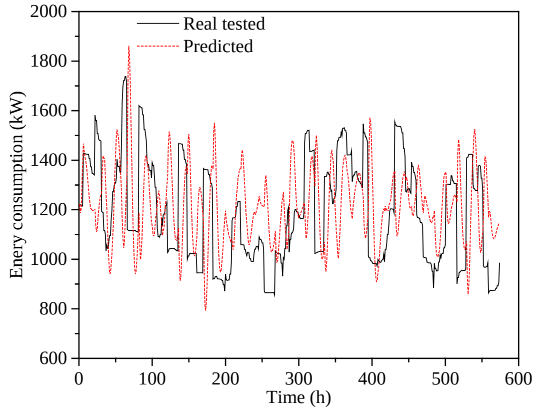

Fig. 10 Energy consumption prediction by N4S model

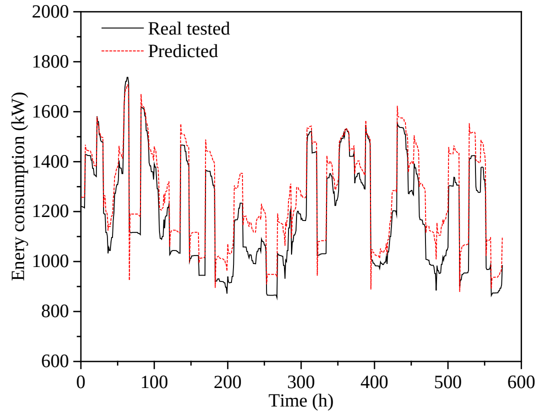

Fig.11 Energy consumption prediction by BN model

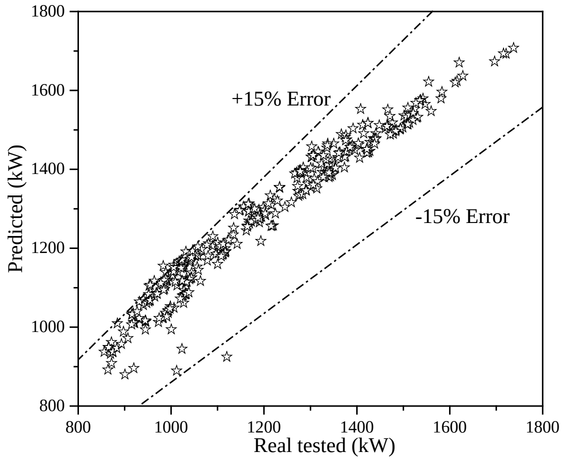

Fig.12 Comparisons of tested and predicted energy consumption prediction by BN model with ±15% error lines

In order to quantitatively evaluate and judge whether each method can meet the recommendations from ASHRAE guideline 14. R^2, RMSE, CV-RMSE NMBE, and NRMSE are used to indicate the statistic feature of the prediction. R^2, which measures the proportion of total variation explained by the fitted regression model is computed from:

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y}_i)^2} \qquad (15)$$

The Root Mean Squared Error (RMSE) is computed from:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-p}} \qquad (16)$$

The coefficient of variation of the root mean square error (CV-RMSE) is computed from:

$$CV - RMSE = 100 \frac{RMSE}{\bar{y}} \qquad (17)$$

The normalized mean bias error (NMBE) is computed from:

$$NMBE = \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)}{(n-p) \times \bar{y}} \times 100 \qquad (18)$$

The normalized root mean square error (NRMSE) is computed from:

$$NRMSE = \frac{RMSE}{y_{\max} - y_{\min}} \qquad (19)$$

ASHRAE Guideline 14 (ASHRAE 2012) 5.3.2.1 requires that the baseline model shall meet the CV-RMSE and NMBE requirement. The required values are dependent of data sampling frequency as listed in Table 1 ASHRAE Guideline 14 only provides requirements for monthly and hourly model. The required value of the daily model is interpolated based on monthly and hourly model.

Table 1
Recommended value for baseline model from ASHRAE Guideline 14

	Monthly	Daily	Hourly
CV-RMSE	15%	22%	30%
NMBE	5%	7%	10%

Table 2 shows all the staticistics indices resulting from four different data-driven models. Compared with ASHRAE's recommoned values listed in Table 1, all algorithms satisfy the ASHRAE recommned criteria for a sampling frequncey of one hour. The CV-RMSE of ARX model is 23.96%. Both SS and N4s have a close CV-RMSE about 21%. BN model has the minimum CV-RMSE of 7.58%. BN is the best prediction algorithm in this case. For RMSE index, ARX, SS and N4S models is more than 250. Nevertheless, the RMSE of BN model is less than 100. Throught the above analysis, BN model is the best prediction model of all four models. Although, all four models meet the ASHRAE recommoned CV-RMSE and NMBE criteria.. ARX model has the biggest error and at the same time the error becomes bigger when the energy consumption is at the low or the high conditions. In addition, SS and N4S models have the time delay prediction problem. It will affect the real time building energy control and analysis.

Table 2
Evaluation performance of four prediction models

	ARX	SS	N4S	BN
R2	-0.832	-0.444	-0.423	0.817
RMSE	284.53	252.60	250.86	90.04
CV-RMSE	23.96%	21.27%	21.12%	7.58%
NMBE	-9.05%	-4.34%	-2.18%	-6.63%
NRMSE	0.323	0.287	0.285	0.102

CONCLUSIONS

Four different data driven algorithms including AutoRegressive with eXternal inputs (ARX), State Space (SS), Subspace state space (N4S) and Bayesian Network (BN) are evaluated and compared using a simulation based case study of predictions of AHU thermal energy consumption.

Modelica simulation program is used to generate training and testing data. In order to obtain sufficient range of input and output variables, excitation signals are given to outdoor air pressure, return air pressure and supply water pressure. Based on the prediction results analysis, the following conclusions are made:

1) Of all four prediction models, ARX model has the worst prediction performance. BN model has the most accurate prediction result.
2) SS and N4S models have the time-delay prediction.
3) Although four models satisfy the ASHRAE Guideline 14's criteria of CV-RMSE and NMBE, other indexes are needed to further evaluate these data driven prediction algorithms for real-time intelligent building operations.

Our future works includes: 1) analyze and evaluate more data-driven algorithms such as artificial neural network (ANN) and support vector machine (SVM); 2) apply these prediction algorithms to the other HVAC and building systems; 3) utilize the selected best algorithm for example, Bayesian network, for model based optimal controls in buildings

ACKNOWLEDGEMENT

This work emerged from the Annex 60 project, an international project conducted under the umbrella of the International Energy Agency (IEA) within the Energy in Buildings and Communities (EBC) Programme. Annex 60 will develop and demonstrate new generation computational tools for building and community energy systems based on Modelica, Functional Mockup Interface and BIM standards.

NOMENCLATURE

a = coefficent
b = coefficent
na = order
nb = order
nk = time delay
q = time delay opeator
u = input
y = output
x = state variable
e = error
P = probability
oa = outdoor air
ra = return air
w = water
μ = mean value
σ = standard deviation
ω = signal
A = state or system matrix
ARX = autoregressive with external inputs
B = input matrix
BN = Bayesian network
C = output matrix
CV-RMSE = coefficient of variation of the root mean
D = feed forward matrix
N4S = subspace state space
NMBE = normalized mean bias error
NRMSE = normalized root mean square error
P = pressure
RMSE = root mean squared error
R2 = coefficient of determination
SS = state space
Y = output matrix
U = input matrix
O = state process matrix
V = noise process matrix
S = input process matrix

REFERENCES

ASHRAE. Guideline 14. (2002) Measurement of energy and demand savings. Atlanta, Georgia: American Society of Heating, Ventilating, and Air Conditioning Engineers.

Cai, J., Kim, D., & Braun, J. E. (2013). Application of Subspace Identification Method in Modeling Multi-zone Buildings.

Candanedo, J. A., & Athienitis, A. K. (2011). Predictive control of radiant floor heating and solar-source heat pump operation in a solar house. HVAC&R Research, 17(3), 235-256.

Heckerman, D. (1997). Bayesian networks for data mining. Data mining and knowledge discovery, 1(1), 79-119

Henze, G. P., Kalz, D. E., Liu, S., & Felsmann, C. (2005). Experimental analysis of model-based predictive optimal control for active and passive building thermal storage inventory. HVAC&R Research, 11(2), 189-213.

Jensen, F. V. (2001). Bayesian networks and decision graphs. Statistics for engineering and information science. Springer, 32, 34.

Karlsson, H., & Hagentoft, C. E. (2011). Application of model based predictive control for water-based floor heating in low energy residential buildings. Building and environment, 46(3), 556-569.

Kusiak, A., Li, M., and Tang, F. (2010). Modeling and optimization of HVAC energy consumption. Applied Energy, 87(10), 3092-3102.

Kusiak, A., Li, M., and Zhang, Z. (2010). A data-driven approach for steam load prediction in buildings. Applied Energy, 87(3), 925-933.

Laur, S. (2009). Bayesian Networks with Continious Distributions.

Ma, J., Qin, J., Salsbury, T., and Xu, P. (2012). Demand reduction in building energy systems based on economic model predictive control. Chemical Engineering Science, 67(1), 92-100.

Ma, J., Qin, S. J., Li, B., & Salsbury, T. (2011). Economic model predictive control for building energy systems (pp. 1-6). IEEE.

Moroşan, P. D., Bourdais, R., Dumur, D., & Buisson, J. (2010). Building temperature regulation using a distributed model predictive control. Energy and Buildings, 42(9), 1445-1452.

Nise, N. S. (2007). CONTROL SYSTEMS ENGINEERING, (With CD). John Wiley & Sons.

Nouidui, T., Phalak, K., Zuo, W. and Wetter, M. (2012) "Validation and application of the room model of the Modelica Buildings library". In: Proceedings of the 9th International Modelica Conference, pp. 727-736.

Nouidui, T., Wetter, M. and Zuo, W. (2012) "Validation of the window and the room models of the Modelica Buildings library". In: Proceedings of the 5th National Conference of International Building Performance Simulation Association -USA (SimBuild2012), pp. 529-536.

O'Neill, Z. (2014). Development of a Probabilistic Graphical Energy Performance Model for an Office Building, ASHRAE conference, Alanta

Wetter, M., Zuo, W., Nouidui, T. S., & Pang, X. (2014). Modelica buildings library. Journal of Building Performance Simulation, 7(4), 253-270.

Yang, J., Rivard, H., & Zmeureanu, R. (2005). On-line building energy prediction using adaptive artificial neural networks. Energy and buildings, 37(12), 1250-1259.

Yoshida, H., & Kumar, S. (2001). Development of ARX model based off-line FDD technique for energy efficient buildings. Renewable energy, 22(1), 53-59.

Yun, K., Luck, R., Mago, P. J., & Cho, H. (2012). Building hourly thermal load prediction using an indexed ARX model. Energy and Buildings, 54, 225-233.

Zhang, Y., O'Neill, Z., Dong, B., & Augenbroe, G. (2015). Comparisons of Inverse Modeling Approaches for Predicting Building Energy Performance. Building and Environment.

Zhang, Y., O'Neill, Z., Wagner, T., & Augenbroe, G. (2013). An inverse model with uncertainty quantification to estimate the energy performance of an office building. In Proceedings of 13th International Building Performance Simulation Association Conference.

Zhao, H. X., and Magoulès, F. (2012). A review on the prediction of building energy consumption. Renewable and Sustainable Energy Reviews, 16(6), 3586-3592.