Title

Modeling and forecasting of cooling and electricity load demand

Authors
Vaghefi, A
Jafari, MA
Bisse, Emmanuel
et al.

Publication Date
2014-12-01

DOI
10.1016/j.apenergy.2014.09.004

Applied Energy 136 (2014) 186–196

Contents lists available at ScienceDirect

Applied Energy

journal homepage: www.elsevier.com/locate/apenergy

Modeling and forecasting of cooling and electricity load demand

A. Vaghefi [a,*], M.A. Jafari [a], Emmanuel Bisse [b], Y. Lu [b], J. Brouwer [c]

[a] Department of Industrial and System Engineering, Rutgers University, Piscataway, NJ 08904, USA
[b] Siemens Corporate Research, Princeton, NJ 08540, USA
[c] National Fuel Cell Research Center, University of California, Irvine, Irvine, CA 92697, USA

HIGHLIGHTS

- We propose a model for forecasting cooling and electricity load demand.
- The model takes the advantage of both time series and regression methods.
- The model is able to accurately forecast the load demands of the CCHP system.

ARTICLE INFO

Article history:
Received 25 January 2014
Received in revised form 14 August 2014
Accepted 2 September 2014
Available online 30 September 2014

Keywords:
Load forecasting
Time series-regression model
Cochrane–Orcutt estimation
Thermal energy storage
Heat recovery steam generator

ABSTRACT

The objective of this paper is to extend a statistical approach to effectively provide look-ahead forecasts for cooling and electricity demand load. Our proposed model is a generalized form of a Cochrane–Orcutt estimation technique that combines a multiple linear regression model and a seasonal autoregressive moving average model. The proposed model is adaptive so that it updates forecast values every time that new information on cooling and electricity load is received. Therefore, the model can simultaneously take advantage of two statistical methods, time series, and linear regression in an adaptive way. The effectiveness of the proposed forecast model is shown through a use case. The example utilizes the proposed approach for economic dispatching of a combined cooling, heating and power (CCHP) plant at the University of California, Irvine. The results reveal the effectiveness of the proposed forecast model.

1. Introduction

U.S. households and commercial buildings consume approximately 40% of total energy use and account for 72% of total U.S. electricity consumption [1]. Commercial building energy demand, in particular, doubled between 1980 and 2000 and is predicted to increase 50% over the next 15 years [2]. As a result, energy demand management has emerged as a key policy for both public and private organizations. CCHP systems can significantly contribute to reducing buildings energy use, curtail pollutant and carbon emission, and help to decrease risks of blackouts and brownouts in the utility grid [3,4]. CCHP technology integrates processes of production and simultaneous use of cooling, heating, and power at a single site. However, since most commercial and industrial electrical loads are highly dynamic and typically not synchronized with local heating and cooling demands, advanced control strategies will be imperative to economic dispatch of CCHP resources.

A wide range of optimal control strategies has been proposed to improve the CCHP operation based on different objectives including power flow, capacity, operation, energy-use and environmental considerations [5–12]. A common element in almost all optimal control strategies is to have an accurate estimation of cooling, heating, and electricity load demands. Some researchers assume that load demands are known and available over a specific period [8,9]. However, cooling and electricity demands are typically stochastic and unknown mainly because of the complex interaction s between plant facilities and equipment, e.g. chillers and turbines yields. Liu et al. [8] point out that in practical applications, the exact future load profile does not exist; and forecasting methods should be taken into consideration by researchers. Therefore, a forecasting mechanism should be applied by researchers to find the future values of load demands.

A number of researchers employ building simulation platform to generate building load demand based on its physical characteristics and other dynamic input variables such as occupancy, weather, and time information. The cooling and electricity load demands are outputs of running the simulation and are then fed into the optimization model [10–12]. However, the quality of

* Corresponding author.
E-mail addresses: vaghefi@rutgers.edu (A. Vaghefi), jafari@rutgers.edu (M.A. Jafari), Emmanuel.Bisse@siemens.com (E. Bisse), yanlu@siemens.com (Y. Lu), jbrouwer@uci.edu (J. Brouwer).

http://dx.doi.org/10.1016/j.apenergy.2014.09.004
0306-2619/

Nomenclature

CCHP	combined cooling, heating and power
LSE	least square estimate
ARX	autoregressive with exogenous variable
ARMAX	autoregressive moving average with exogenous variable
ARMA	autoregressive moving average
AI	artificial intelligence
ANN	artificial neural network
R^2	coefficient of determination
R^2_{adj}	adjusted coefficient of determination
GT	gas turbine
ST	steam turbine
TES	thermal energy storage
HRSG	heat recovery steam generator
COP	coefficient of performance
W^k_{CHC}	cooling power generated by the kth chiller (kW)
W^k_{CHW}	power consumed by the kth chiller to cooling power (kW)
COP^k	coefficient of performance for the kth chiller
W_{CHW}	total power consumed by chillers to generate total cooling power of the campus (kW)
$Q_{cooling}$	cooling demand from the campus (kW)
T_{CHRw}	returned water temperature to chillers (K)
T_{CHSw}	supply water temperature from chillers (K)
$\dot{m}_{chw}$	chilled water mass flow rate (kg/s)
c_w	specific heat capacity of water (kJ/kg K)
$W_{electricity}$	electricity demand from the campus (kW)
W_{grid}	power purchased from grid (kW)
W_{GT}	power produced by gas turbine (kW)
W_{ST}	power produced by steam turbine (kW)

Symbols

W	electricity power
Q	cooling power
$\dot{m}$	mass flow
T	temperature
ε	random error term
y	dependent/output variable
x	independent/input variable
β	coefficients of input
Φ	autoregressive operator
Θ	moving average operator
B	backward operator

Subscripts

T	time (h)
J	index for input variables
CHC	cooling generated by chiller
CHW	power consumed by chiller
W	water
Chw	chilled water
$CHRw$	water returned to chiller
$CHSw$	water supplied by chiller
GT	gas turbine
ST	steam turbine
$grid$	power grid

Superscripts

k	chiller number

results highly depends on quality of the simulation models and their inputs. In addition, for any CCHP optimization, a detailed building simulation model needs to be accordingly built and run repeatedly. Another way to deal with this problem is to consider uncertainty in CCHP optimization model. Hu and Cho [12] for instance, propose an optimization model with some probabilistic constraints to guarantee that the model is reliable to satisfy the stochastic load demand. They assume load demands are independent and follow normal distributions in which 95% of the area is within the range of ±20% of the average load demands. Another approach to this problem is to develop a forecasting model and embed it into the optimization model. This is the main motivation of this work. In this paper, Cochrane–Orcutt estimation technique is used as an effective linear model to provide look-ahead forecasts for cooling and electricity demand load. It simultaneously fits a regression model and a time series to the data while maintaining least square estimate (LSE) conditions. In addition, the forecast values are modified when a new data is received from the real system. The proposed model is currently working as a part of an integrated optimal dispatch for CCHP plant at the University of California, Irvine and providing accurate forecasts for the entire campus cooling and electricity load demand.

2. Background study

In most real cases, cooling and electricity load demands are highly dynamic oscillating within a wide range of values during course of a day. This is mainly because several physically explicit or latent factors can instantaneously influence cooling and electricity demand patterns. These factors can be any one of the following types: (i) Static factors that are usually set at the design stage and only change due to aging wear and tear. Building characteristics, CCHP components, chiller types and generator nominal capacities are examples of such factors; (ii) Environmental variables extrinsic to the building, such as climate and weather data; (iii) operational variables, e.g. cooling/heating set point values, lighting, time schedule to operate various equipment and system components within plant or building; and (iv) uncontrollable dynamical variables, such as number of occupants at any time, noise due to structural variations etc. It is ideal to know all these factors and their impacts on energy dynamics in order to optimally forecast and control cooling and electricity demands for single building or a cluster of buildings. However, a complete forecast model is not practically attainable due to unknown significant dynamical variables, lack of tools to measure their effects, or that some of these variables are uncontrollable. Therefore, a wide range of different methods has been proposed to model and forecast load dynamics. In overall, these methods can be categorized into three general approaches.

In the first approach, a linear or nonlinear statistical model is used to explain the variability of response (load or energy dynamics) over time. The most popular example of such statistical models is Box and Jenkins time series paradigm where load demands are estimated based upon a linear combination of their past values [13,14]. There are a large family of different models in this category that can deal with many special cases including seasonality, non-stationary, and non-homogeneity of variances (see e.g. [15,16]). The major drawback of such models is that the future values are typically forecasted based upon the past and present values of cooling and electricity load demands without considering any *exogenous* factors in the model. Another example of statistical approach is using regression models (metamodel) where the variability within response is modeled via a number of exogenous factors [17–21]. The major problem of such models is that they often ignore the complex interactions between exogenous factors, which may result in less accurate forecast values. To overcome this problem, a number of studies use a hybrid approach, which employs the main components of both above-mentioned approaches [22]. ARX and ARMAX are two examples of this approach. Although these models perform effectively in many cases, they have many parameters to be estimated since all input and output variables with their past and current values should appear in the forecast model.

The second approach employs artificial intelligence to find the k-step ahead forecasts for load demand. A broad range of numerical methods can be included in this category. Kalogirou [23] and Mellit and Kalogirou [24] provide a comprehensive review of AI techniques in some areas of energy. Although their techniques are not directly related to load forecasting, however, they can easily be used with minor changes. Artificial neural network (ANN) is among most frequent AI techniques and has been widely used in load or energy forecasting. ANN's have particularly evolved based upon different settings of neuron arrangement, neuron connections, training techniques, and internal layers and become a powerful competitor for statistical methods [25–29]. They can be designed to include both past observation of cooling and electricity demands and associated exogenous factors. The main disadvantage of AI approach is that they are often black box and do not show any explicit relationship between response an input variables. For example, the hidden layers of ANN's are difficult to explain and cannot be appeared in an explicit forecasting equation [30].

In addition, by developing computational methods, a third approach has recently been developed which is a combination of any abovementioned techniques. The main purpose of this hybrid approach is to improve the accuracy of the forecast values by combining different numerical–analytical methods. Some hybrid methods also partially include the physical aspects of the real system in their computation and come up with a mixed physical–numerical method, which is often referred to as gray models [30]. A few applications of hybrid models in the area of energy can be found in [31,32,29].

The proposed model can be classified in the statistical groups. It first fits a linear regression to find the correlation between the cooling and electricity load demands and exogenous factors. Any variability that cannot be explained by regression models can be aggregated in residual terms. Then, a seasonal time series model is applied to the residuals to express the remaining variability. Since, the regression parameters should be estimated using least square error method, the process of parameters estimation is applied iteratively and simultaneously. Further details will be explained in the next section.

3. Time series regression model

The common assumption of uncorrelated random error terms (ε's) made in basic regression models is not appropriate to forecast building energy consumption. Historical data shows that error terms are frequently correlated (often positively) over time [33]. In particular, this typically happens when there are some uncontrollable, unknown, or non-measurable input variables. A special case for the regression model with auto-correlated data can be shown as follows:

$$y_t = \sum_{j=0}^{k} \beta_j x_{tj} + \varepsilon_t, \varepsilon_t = \xi(\varepsilon_{t-1}, \ldots, \varepsilon_{t-q}) + \alpha_t, \tag{1}$$

where $\xi(.)$ is a function of previous error terms ε's, y_t is the power consumed at time t and x_{tj} is the jth input variable affecting the building energy consumption at time t and a_t is a white noise. The error terms are typically modeled using Box and Jenkins model as a first order auto-regressive model. A preliminary study of our historical data on cooling and electricity load demands indicates a seasonal pattern with lag of 24 h. Therefore, the error terms in (1) is generalized to include seasonal patterns. To do this, assume that p, q, P and Q are the order of non-seasonal and seasonal autoregressive and moving range parts respectively, and s is the seasonal order. Then a general ARMA model for error terms can be written as follows:

$$\phi_p(B)\Phi_P^s(B)\varepsilon_t = \theta_q(B)\Theta_Q^s(B)\alpha_t, \tag{2}$$

where ϕ_p and Φ_P^s are autoregressive operators, θ_q and Θ_Q^s are moving average operators and B is backward operator. s is set equal to 24 showing the significance of autocorrelation between loads of same time in two consecutive days.

Let

$$\phi_p(B)\Phi_P^s(B) = 1 - \Psi(B)$$

then

$$\varepsilon_t = \Psi(B)\varepsilon_t + \theta_q(B)\Theta_Q^s(B)\alpha_t,$$

Furthermore, Eq. (2) can be written as follows:

$$y_t = \sum_{j=0}^{k} \beta_j x_{tj} + \varepsilon_t \sum_{i=0}^{p} \sum_{j=0}^{P} (-1)^{i+j-1} \varphi_i \phi_j B^{i+s\times j} + \alpha_t \sum_{i=0}^{q} \sum_{j=0}^{Q} (-1)^{i+j} \theta_i \Theta_j B^{i+s\times j}, \tag{3}$$

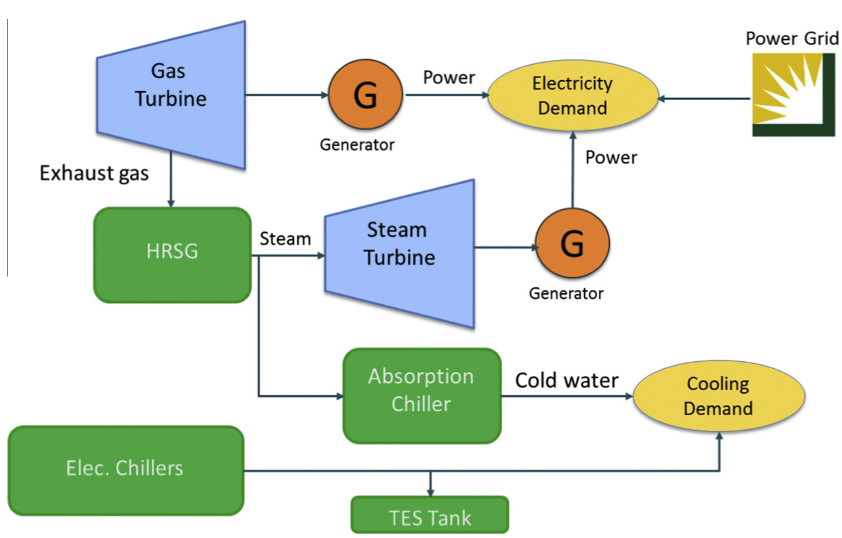

Fig. 1. Schematic framework of CCHP plant at University of California, Irvine.

Note that $\varphi_0 = \phi_0 = 0$ and $\theta_0 = \Theta_0 = 1$. For example, for the ARMA $(1,0) \times (1,0)_{n=24}$, Eq. (3) is written as follows:

$$
\begin{aligned}
y_t &= \sum_{j=0}^{k} \beta_j x_{tj} + \varepsilon_t(\varphi_1 B + \phi_1 B^{24} - \varphi_1 \cdot \phi_1 B^{25}) \\
&= \sum_j \beta_j x_{tj} + \varphi_1 \varepsilon_{t-1} + \phi_1 \varepsilon_{t-24} - \varphi_1 \cdot \phi_1 \varepsilon_{t-25},
\end{aligned} \tag{4}
$$

The main significance of Eq. (3) is that it includes seasonal error and tends to capture statistical similarities between two periods, which are n hours apart. The major problem of multiple linear regression with auto-correlated error terms is the estimation of coefficients. With auto-correlated error terms, the ordinary least square (OLS) procedures can be misleading and does not guarantee estimation with the minimum variance [33]. To overcome this problem, Cochrane and Orcutt [34] proposed a transformation when error terms follow a first order autoregressive process. According to Cochrane–Orcutt model, one should transform the response values in such a way that

$Y'_t = \phi_p(B)\Phi^s_P(B)Y_t, \mathbf{x}'_t = \phi_p(B)\Phi^s_P(B)\mathbf{x}_t$ and $\beta'_0 = \phi_p(B)\Phi^s_P(B)\beta_0$.

Therefore, Eq. (3) can be replaced by

$$
Y'_t = \beta'_0 + \mathbf{x}'_t \boldsymbol{\beta}'_t + a_t, \tag{5}
$$

Eq. (5) is an ordinal multiple linear regressions with independent error terms and can be calculated via OLS estimation method. As a result, the fitted linear function $\widehat{Y}'_t = \hat{\beta}'_0 + \mathbf{x}'_t \hat{\boldsymbol{\beta}}'_t$ can eliminate the autocorrelation structure of the error terms. The following algorithm summarizes our approach:

3.1. Algorithm

{

Step 1: Divide the original dataset into two subsets: training dataset and testing dataset, which are used for model estimation and model verification respectively and denoted by Ω_1 and Ω_2. Set $i = 0$.

Step 2: Fit a multiple regression model to training subset and estimate vector of $\hat{\boldsymbol{\beta}}^i$ in $\mathbf{y}'^i_1 = \mathbf{X}_1 \hat{\boldsymbol{\beta}}^i$, where $\mathbf{y}_1, \mathbf{X}_1 \in \Omega_1$ are response (cooling or electricity load demand) and independent variables (exogenous variables). Then calculate initial residual values by $\hat{\boldsymbol{\varepsilon}}^i = \mathbf{y}_1 - \mathbf{y}'^i_1 = \mathbf{y}_1 - \mathbf{X}_1 \hat{\boldsymbol{\beta}}^i$.

Step 3: If $\hat{\boldsymbol{\varepsilon}}^i$'s are correlated fit an ARMA model, $\hat{\phi}^i_p(B)\hat{\Phi}^{si}_P(B)\varepsilon^i = \hat{\theta}^i_q(B)\hat{\Theta}^{si}_Q(B)\alpha_t$, and find estimation values for $\hat{\phi}^i_p, \hat{\Phi}^{si}_P(B)$ and $\hat{\Theta}^{si}_Q(B)$ using least square error technique or other estimators.

Step 4: Apply following transformations $\mathbf{y}'^i_1 = \phi^i_p(B)\Phi^{si}_P(B)\mathbf{y}_1$ and $\mathbf{X}'^i_1 = \phi^i_p(B)\Phi^{si}_P(B)\mathbf{X}_1$ on $\mathbf{y}_1, \mathbf{X}_1 \in \Omega_1$. Then fit a new multiple regression model to transformed subset and estimate vector of $\hat{\boldsymbol{\beta}}'^i$ in $\mathbf{y}'^i_1 = \mathbf{X}'^i_1 \hat{\boldsymbol{\beta}}'^i$.

Step 5: Check $|\hat{\boldsymbol{\beta}}'^i - \hat{\boldsymbol{\beta}}'^{i-1}| < \delta$. If the criterion is met then set $\hat{\boldsymbol{\beta}}' = \hat{\boldsymbol{\beta}}'^i$ and go to Step 6. Otherwise, calculate the residual values by $\hat{\boldsymbol{\varepsilon}}^i = \mathbf{y}_1 - \mathbf{y}'^i_1 = \mathbf{y}_1 - \mathbf{X}'^i_1 \hat{\boldsymbol{\beta}}'^i$ and go to Step 3.

Step 6: Apply anti-transformations for $\beta'_0 = \phi_p(B)\Phi^s_P(B)\beta_0$ and $\hat{\boldsymbol{\beta}} = \hat{\boldsymbol{\beta}}'$ and use them in Eq. (1).

}

It is quite common to use the estimated parameters as well as subset Ω_2 to check the adequacy of the given model. In this study, coefficient of determination R^2 and adjusted coefficient of determination R^2_{adj} are employed as measures for model adequacy checking. These measures can be calculated as follows:

$$
R^2 = \frac{\hat{\boldsymbol{\beta}}'^T \mathbf{X}'^T_2 (I - H)\mathbf{X}'_2 \hat{\boldsymbol{\beta}}'}{\mathbf{y}'^T_2 (I - (1/n)J)\mathbf{y}'_2}, \tag{6}
$$

and

$$
R^2_{adj} = \frac{\hat{\boldsymbol{\beta}}'^T \mathbf{X}'^T_2 (I - H)\mathbf{X}'_2 \hat{\boldsymbol{\beta}}'/k - 1}{\mathbf{y}'^T_2 (I - (1/n)J)\mathbf{y}'_2/n_2 - k}, \tag{7}
$$

where k is number of exogenous variables, n_2 is sample size for testing dataset, I is identity matrix and H can be calculated by $H = \mathbf{X}_2(\mathbf{X}^T_2\mathbf{X}_2)^{-1}\mathbf{X}^T_2$ as well. R^2 and R^2_{adj} are both between 0 and 1 and explain the percentage of variation that is explained by model. A closer value to 1 depicts a better model.

4. Case study and experimentation

In this section, the forecast model is employed as a part of optimal dispatching of a CCHP plant at the University of California, Irvine. Cooling and electricity forecast values are fed into an optimal control strategy, which searches for optimal set points for 24 h ahead. The forecast model then is used to compute optimal control values to minimize energy consumption during course of a day in a building.

The UC Irvine Central Plant consists of eight electric chillers, providing cold water, a 13.5 MW gas turbine (GT), a 5.7 MW steam turbine (ST), thermal energy storage (TES) tank, and a heat recovery steam generator (HRSG). It provides heating and cooling loads for the entire campus as well as the majority of the campus electric loads. The chillers are able to supply as much as 14,500 tons (51 MW) and the steam driven chiller can provide an additional 2000 tons (7 MW). The TES tank capacity is 60,000 ton-hour (211

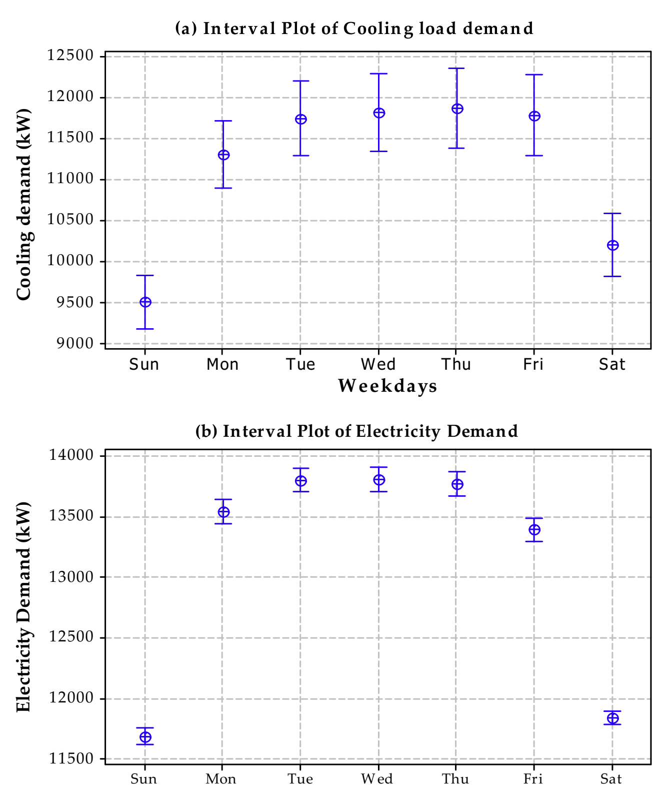

Fig. 2. 95% Confidence interval plots categorized by weekday for (a) cooling load demand and (b) electricity load demand (Day 1 is Sunday, Day 7 is Saturday).

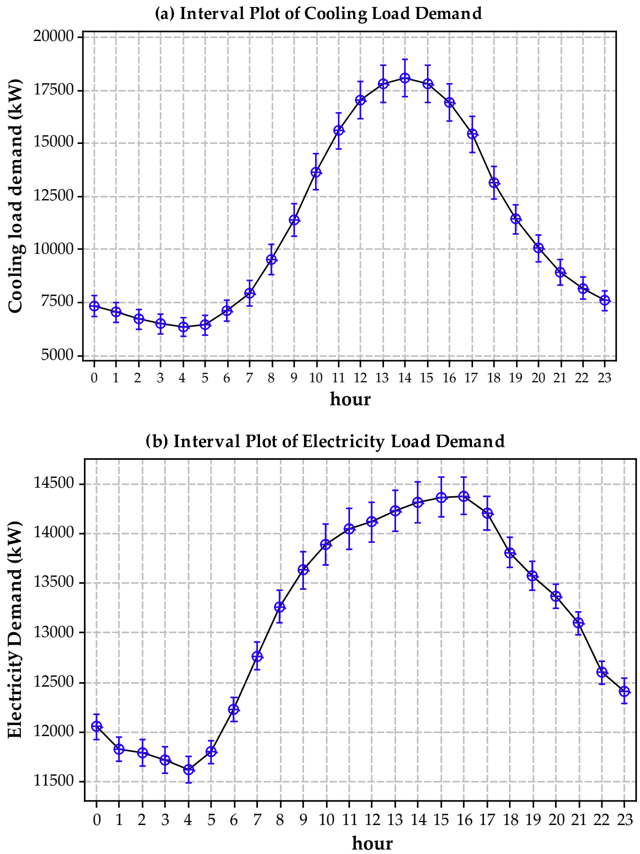

Fig. 3. 95% Confidence interval plots categorized by hours for (a) cooling load demand (kW) and (b) electricity load demand (kW).

megawatt-hour) which is able to shift, on average, 65% of the cooling load during the day to the night when electricity prices are lower and temperature is cooler.

Fig. 1 provides a schematic of the plant, where GT is the primary source of electric power providing electricity for the campus and for the chillers. As a byproduct, the gas turbine generates the exhaust gas, which can be source of extra thermal energy. Such energy is then used to produce steam using HRSG unit. HRSG can supply 23,500 kg/hour and 54,000 kg/hour without and with duct fire, respectively. The generated steam drives the steam turbine (ST). The steam can also be used to produce hot water for the campus needs. A portion of the produced steam is also transferred to use in a steam chiller unit. GT and ST supply about 85% of the total electrical needs on the campus with the balance being served by utility import (14%) and an 893 kW-fixed panel solar photovoltaic (1%).

As mentioned, the electricity produced by two generators are either sent directly to the campus to satisfy electricity demand or supplied as the energy input to the electrical chiller (see [9] for more details), which is mainly responsible to provide cold water. Cold water can be either directly supplied to the campus to meet campus cooling needs or stored in the TES tank for later use. Hence, the chillers and the TES together are the main sources for the campus cooling demands. Any additional electricity demand is provided from the grid.

Such a CCHP system is able to produce thermal energy along with electricity over time. The thermal energy storage (TES) is a flexible component of the plant, which allows the campus to reshape the cooling demand particularly in peak hours. There are many examples of CCHP supervisory control systems in literature [5,6,10]. A key element for such optimal control is to have accurate information about the power (electricity and cooling) demand over the course of a day, which is the central focus of this study.

Suppose that W^k_{CHC} is the cooling load generated by the kth chiller (kW), and that W^k_{CHW} is the power consumed by the kth chiller (kW) to generate W^k_{CHC} units of cooling load. Then W^k_{CHC} is proportional with W^k_{CHW} as follows:

$$W^k_{CHW} = w^k_{CHC}/COP^k \tag{8}$$

where COP^k is the coefficient of performance for the kth chiller which is the ratio between efficient energy acquired by and supplied to the chiller; this is typically determined by the chiller manufacturer. In this study, COP^k is fixed and given by the chillers' manufacturer. However, in reality, it is a function of the real operating temperature and reliability of the absorption chiller. This information is not often available. Therefore, any variation due to change in COP^k is appeared in error term of (1) and should be modeled via time series part of the proposed model.

W^k_{CHW} presents the actual power (electricity) consumed by the kth chiller to produce W^k_{CHC}. The total power consumed by all chillers is given by:

$$W_{CHW} = \sum_{k=1}^{8} W^k_{CHW}. \tag{9}$$

Note that W^k_{CHC}values do not reflect the cooling power supplied to the campus. A portion of cooling load produced by the chillers is sent to the TES tank and stored for peak hours. Thus, W^k_{CHC} values cannot be a good measure for determining the total cooling demand of campus at any time. Instead, the amount of cooling supplied to the campus can be expressed as follows:

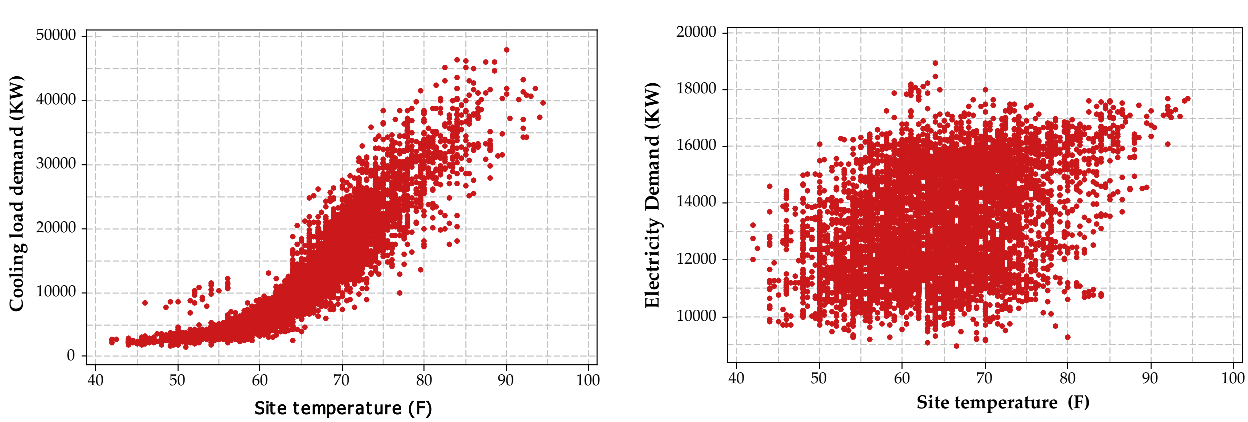

Fig. 4. Scatter plots of cooling and electricity load demand versus site temperature.

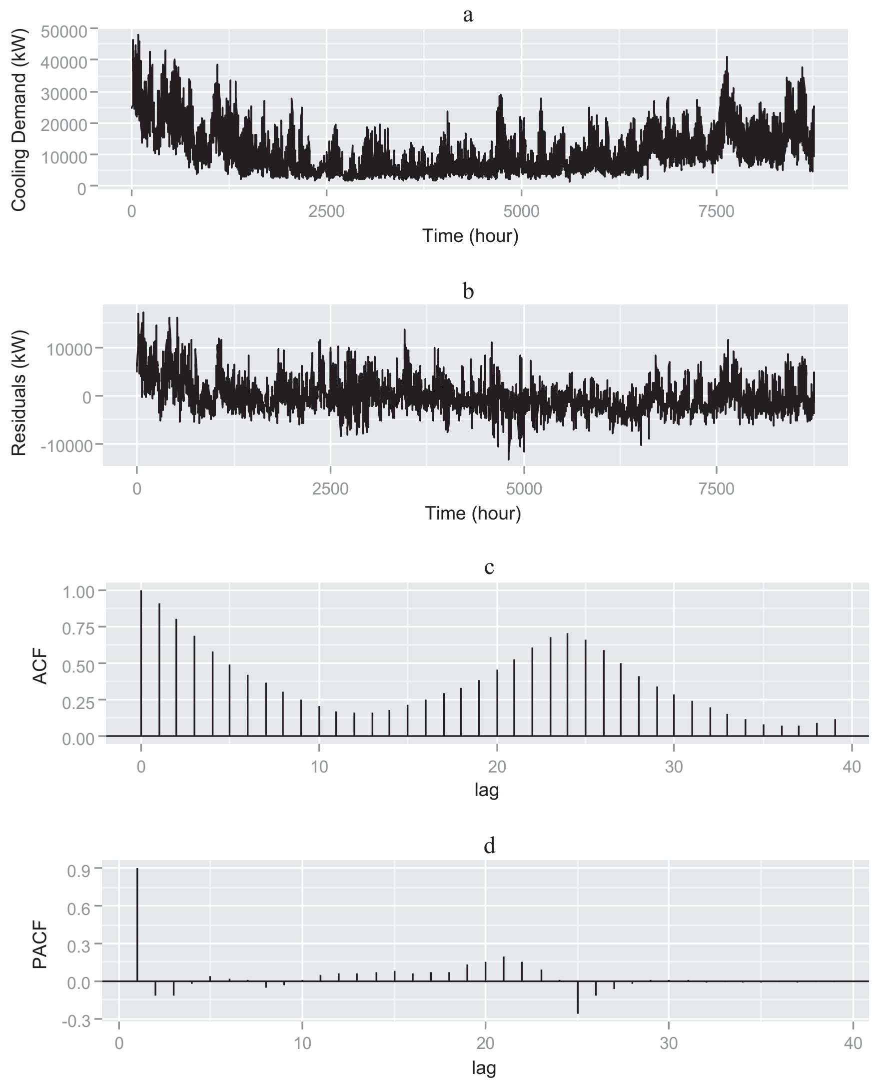

Fig. 5. (a) Time series plot for the cooling load demand, (b) the residuals for a preliminary linear model, (c) autocorrelation plot and (d) partial autocorrelation plot for residual values.

$$Q_{cooling} = \dot{m}_{chw} \times c_w \times (T_{CHRw} - T_{CHSw}), \tag{10}$$

where $Q_{cooling}$ is the total amount of cooling (kW) provided by the chillers and supplied to the campus to meet cooling demands, T_{CHRw} is the temperature of returned water to chillers (K), T_{CHSw} is the supply water temperature from chillers (K), $\dot{m}_{chw}$ is the chilled water mass flow rate (kg/s) and c_w is the specific heat capacity of water (kJ/kg K) [9]. All above parameters are known and available in the plant. This allows us to accurately estimate the actual cooling load demands.

Similar to the cooling load, the direct values for the electricity load demand are not available. However, this can be calculated from the hourly power consumption by the chillers, the total power generated by gas and steam turbines, and the power provided by grid. The electricity load at time t is therefore:

$$W_{electricity} = W_{GT} + W_{ST} + W_{grid} - W_{CHW}, \tag{11}$$

where W_{GT} and W_{ST} are the power produced by gas and steam turbines, respectively, and W_{grid} is the power purchased from grid at any time. W_{CHW} is the total power consumed by all chillers, which is calculated in (9), and $W_{electricity}$ is the electricity load demand at time t. In this study, due to lack of data, we ignore the power consumption by pumps and chiller compressors, which account for a relatively negligible portion of the power consumption throughout the campus. The proposed forecast model is used to forecast both $Q_{cooling}$ and $W_{electricity}$ using a set of weather and time variables as well as historical cooling and electricity data.

5. Results for the CCHP plant data

In this section, the performance of the proposed method is discussed using the CCHP plant data collected from the UCI campus. In this example, one year (September 2009–2010) and 4 months data (September 2009–December 2009) are used for building the forecast models for the cooling and electricity load demands, respectively. Both datasets are provided by the UCI campus plant based on actual values of the cooling and electricity consumption. Each dataset is divided into two subsets. The first set is used for model building and estimation purposes (training dataset). The rest of the data is used for validation purposes (testing dataset). In this work, Matlab is employed for creating and testing the proposed forecast model and plotting and visualization is done by Minitab and R. In this phase, Eqs. (6) and (7) are used to investigate the performance of the forecast models. The testing subset does not share any information with the training dataset.

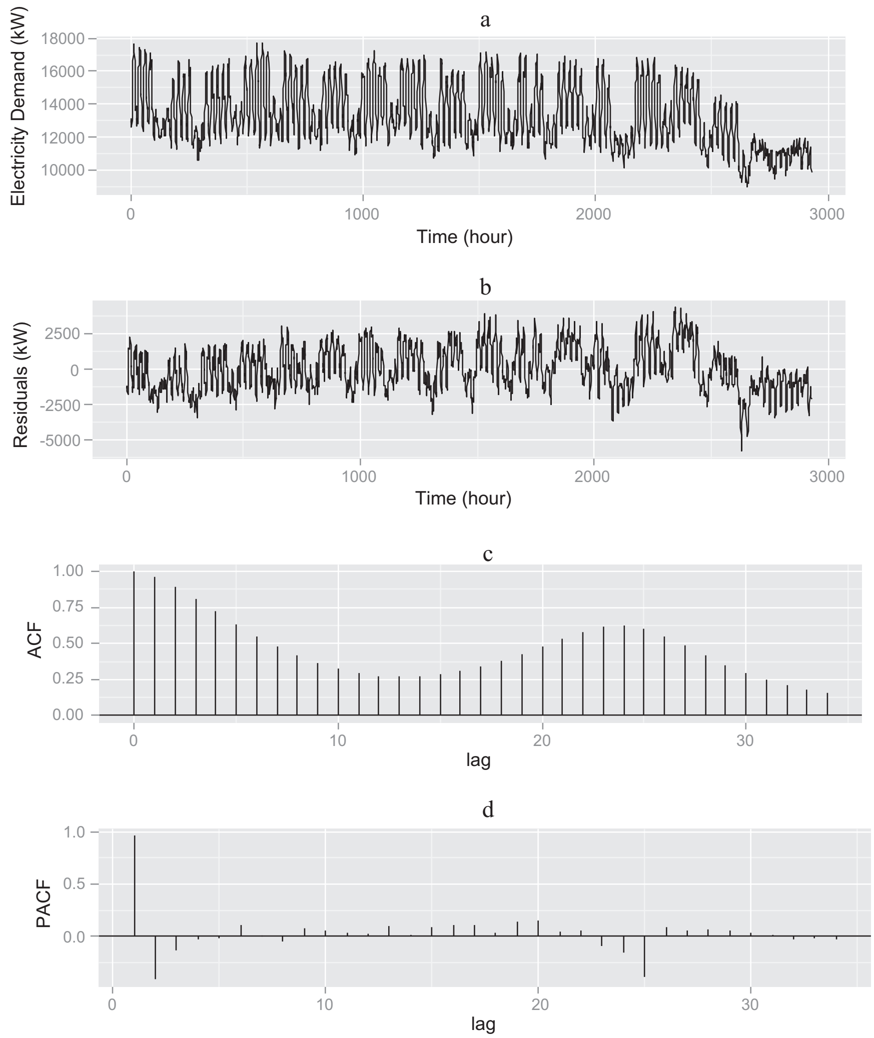

Fig. 6. (a) Time series plot for electricity load demand, (b) the residuals for a preliminary linear model, (c) autocorrelation plot and (d) partial autocorrelation plot.

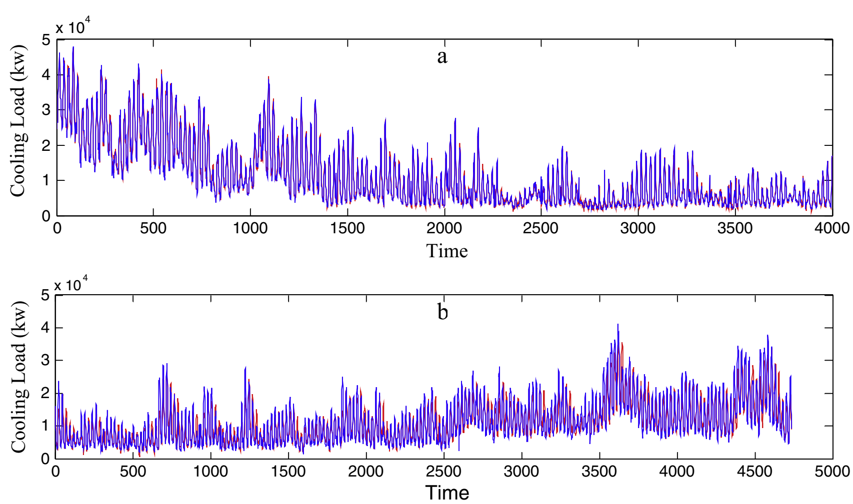

Fig. 7. Comparison of actual and forecasted values for cooling load demand using (a) training dataset (above) and (b) testing dataset (below).

Before building the forecast model, an exploratory data analysis is performed to capture the behavior of data over time. Fig. 2 depicts the 95% confidence interval plots for the cooling and electricity load demands categorized by weekdays. It is observed that both the cooling and the electricity load demands are higher in working days than weekends. This is particularly obvious for the electricity load demand that is less than 12,000 (kW) in weekends and more than 13,000 (kW) for weekdays. This implies that mixing

Table 1
The estimates values for cooling and electricity forecast models.

	Cooling		Electricity			
			Weekdays		Weekends	
	Estimate	Standard error	Estimate	Standard error	Estimate	Standard error
β_0	−13441.85	1851.10	12,783	458.2	9825.557	704.74
β_1	357.28	13.98	21.0180	6.1137	39.168	8.809
φ_1	0.9059	0.016	0.8775	0.0319	1.1882	0.0750
φ_2	0.0513	0.021	0.0018	0.0425	−0.1848	0.1162
φ_3	−0.0830	0.02134	−0.1298	0.0423	−0.0140	0.1168
φ_4	−0.0314	0.02138	0.0606	0.0425	−0.0594	0.1168
φ_5	−0.0299	0.02138	−0.1054	0.0425	−0.2021	0.1169
φ_6	−0.0405	0.02137	0.0073	0.0425	0.1319	0.1178
φ_7	0.0829	0.02138	0.1304	0.0422	0.1240	0.1160
φ_8	−0.0264	0.02142	−0.1095	0.0423	−0.1624	0.1170
φ_9	−0.0289	0.02142	−0.0625	0.0425	0.1140	0.1190
φ_{10}	−0.0094	0.02140	0.0484	0.0425	−0.1592	0.1195
φ_{11}	0.0191	0.02136	0.0155	0.0426	0.1565	0.1210
φ_{12}	0.0017	0.02135	−0.0777	0.0427	−0.0218	0.1219
φ_{13}	−0.0138	0.02135	0.0466	0.0427	0.1390	0.1204
φ_{14}	−0.0112	0.02134	−0.0047	0.0427	−0.1450	0.1190
φ_{15}	0.0275	0.02134	−0.0287	0.0427	−0.1449	0.1189
φ_{16}	−0.0021	0.02134	0.0276	0.0426	0.0813	0.1194
φ_{17}	0.0012	0.02134	0.0957	0.0425	0.2598	0.1193
φ_{18}	−0.0074	0.02129	−0.1241	0.0424	−0.2818	0.1206
φ_{19}	0.0084	0.02128	0.0501	0.0426	−0.0009	0.1222
φ_{20}	0.0191	0.02128	0.0550	0.0425	0.0622	0.1233
φ_{21}	0.0392	0.02128	−0.0531	0.0425	0.0664	0.1228
φ_{22}	0.0728	0.02124	0.0966	0.0423	0.1109	0.1241
φ_{23}	0.0751	0.02125	0.0439	0.0424	−0.0776	0.1239
φ_{24}	−0.0322	0.01576	0.0797	0.0319	−0.0285	0.0814
R^2	0.884		0.708		0.430	
$R^2{}_{adj}$	0.883		0.700		0.405	

all data and building a global forecast model without considering the factor of "day" may result in a less powerful model. Thus, in this work, two different models are constructed for weekdays and weekends.

Fig. 3 presents the 95% confidence interval plots for the cooling and the electricity load demands categorized by 24 h of the day. For example, 17 in x-axis means the 95% confidence interval for the cooling and electricity load demands at time 17:00, which is constructed by all data collected at this particular time slot. This figure can easily represent peak time for the cooling and electricity load demands.

For cooling, the load demand increases constantly from 6:00 and reaches its maximum value at time14:00 then decreases until end of the day. The peak hours for the cooling load demand are between 11:00 and 17:00. This also implies that the cooling demand load is highly correlated with the ambient temperature. Similarly, the peak hours for electricity load demand are between 9:00 and 19:00 as well.

Fig. 4 shows scatter plots of the cooling and electricity load demands versus the ambient temperature. The cooling load values show higher correlation with ambient temperature than the electricity load demand. The estimated correlations between cooling and electricity load demands with ambient temperature are 0.905 and 0.374, respectively. This means that to find an accurate model for the electricity load demand, it is required to add more significant exogenous factors than ambient temperature.

For example, the average number of people in the campus at time t would be a potential exogenous factor for modeling the campus electricity load demand. As number of people in the campus increases, it is logical to presume that the electricity load demand increases. However, in this example, since the number of people in the campus at time t is not available, it is not possible to analyze its effect. As a result, those parts of variation that are related to such missing exogenous factor(s) should be explained and modeled by time series part of the proposed method.

Fig. 5a and b present the hourly cooling load of the campus and the residual values given by fitting a linear model of cooling versus ambient temperature. The residuals are highly autocorrelated over time in different lags (Fig. 5c). Furthermore, Fig. 5d is the partial autocorrelation function (PCAF) for residual values and can identify the extent of lags in an autocorrelation model. In this figure, PACF illustrates a strong autocorrelation structure in the first lag and the 24th lag, which accounts for seasonality in the data. Therefore, a seasonal ARMA$(2,0,0) \times (1,0,0)_{24}$ seems an appropriate candidate for the electricity load dataset.

Similarly, Fig. 6a and b are the electricity load demand and its corresponding residual values when applying a linear model to the data. Again, ACF and PACF in Fig. 6c and d reveal a correlated structure for the electricity load dataset. Particularly, PACF illustrates a positive autocorrelation for the first lag and a remarkable negative correlation for the 24th lag. This means that a seasonal ARMA$(1,0,0) \times (1,0,0)_{24}$ model would be enough for the electricity load demand.

Fig. 7a and b depict the result of forecast modeling for the cooling load demand using training and testing datasets. In Fig. 7a, the forecast values are very close to the corresponding actual values. This is because the training dataset is used for parameter estimation of the forecast model. Therefore, the model includes the information of actual data.

Fig. 7b represents the performance of the model with testing dataset, which does not share any information with the estimated parameters. It is observed that the model adequately fits with the actual data. In addition, Table 1 provides the estimate values of the model parameter, their standard errors as well as coefficient of determinations for both cooling and electricity load demands. For the cooling demand, coefficient of determination R^2 and adjusted

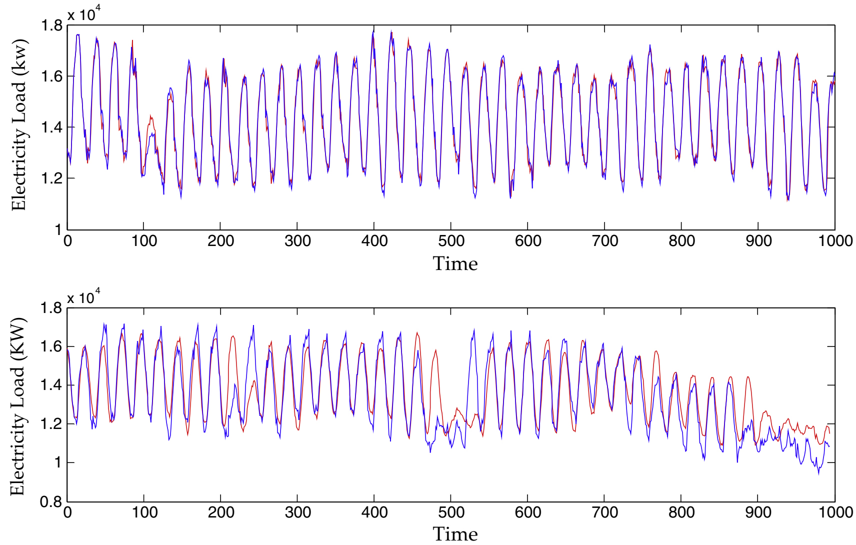

Fig. 8. Comparison of actual and forecasted values for electricity load demand in weekdays using training dataset (above) and testing dataset (below).

coefficient of determination R^2_{adj} are 88.4% and 88.3%, respectively implying that the proposed model can explain more than 88% of the total variability within data.

Figs. 8 and 9 present the actual and forecast values of electricity load demand using both training and testing datasets for weekdays and weekends, respectively. As shown in Fig. 2b, the electricity demand patterns are significantly different in weekends and weekdays, probably because of fewer numbers of people in the campus in weekends. Therefore, to improve the performance of the proposed method, two separate models should be built for weekdays and weekends.

In addition, It is observed from Figs. 8 and 9 that the performance of the proposed model for the electricity load demand is still less than the same model proposed for the cooling load demand. This is mainly due to lack of other exogenous factors in electricity demand model. As shown in Fig. 4, the correlation between electricity load demand and the ambient temperature is moderate. It means that the ambient temperature can only explain a relatively small portion of variation in electricity demand. This can be confirmed by observing Table 1. In this table, R^2 and R^2_{adj} for electricity load demand in weekdays are namely 70.8% and 70% and for electricity load demand in weekends are namely 43% and 40%. Therefore, the electricity load model should be enhanced by adding more exogenous factors e.g. occupancy into the forecast model in order to capture larger amount of variability over time.

Another potential reason for lower performance of the electricity demand forecast model is shown in Fig. 10. In this figure, the values of electricity load demand are plotted over time and are grouped by months. It is shown that the load demand in the last month follows different pattern than the other months. This is because the last month is December and the campus is probably less populated at the last days of December. Since, the model uses the first two months for training and estimation and the rest of data (including December data) for the testing purposes, it cannot fit the last part of December.

A solution for this problem is to add the occupancy as another exogenous variable into the model and re-estimate the model parameters accordingly. This way, the model can differentiate between those days that more people are in campus from the days that less people are in campus including weekends. Another idea is

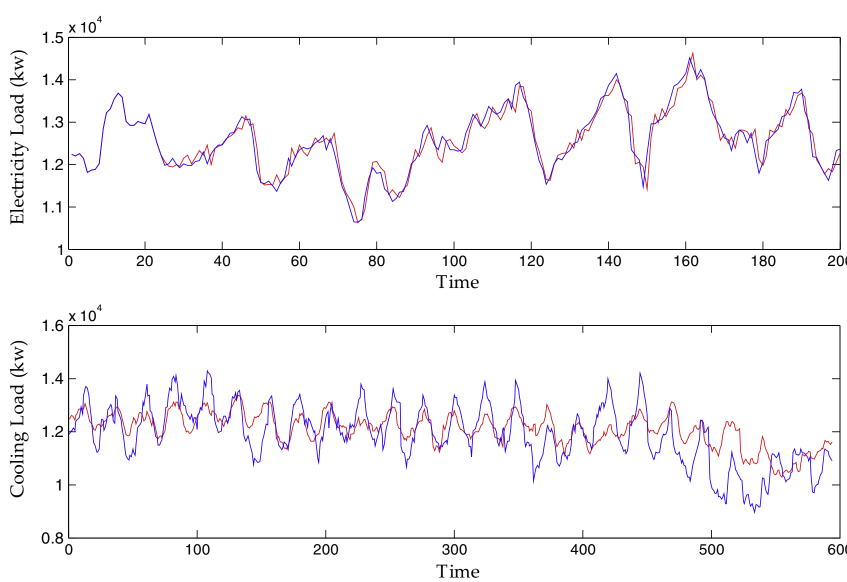

Fig. 9. Comparison of actual and forecasted values for electricity load demand in weekends using training dataset (above) and testing dataset (below).

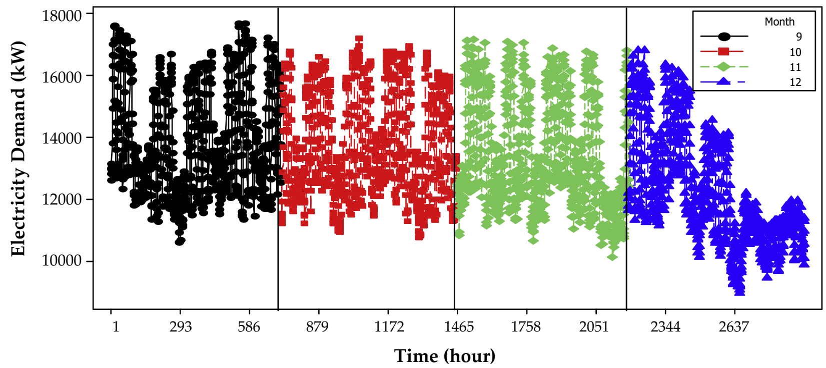

Fig. 10. Time series plot for electricity load demand grouped by month.

to build a new model solely for December. In doing so, the model switch to a new model that is designed and built based on December data as soon as December begins.

6. Conclusions

In this paper, a statistical method for forecasting cooling and electricity load demands was proposed in both campus and building level. The performance of the proposed model was evaluated using the CCHP plant data collected from the UCI campus. The results revealed that the proposed model was able to provide high quality forecasts for both cooling and electricity load demands. Coefficient of determination R^2 and adjusted coefficient of determination $R^2{}_{adj}$ for forecasting cooling demand were 88.4% and 88.3%, implying that the proposed model could explain more than 88% of the total variability within testing data. These indices were 70.8% and 70% for electricity load demand in weekdays and 43% and 40% in weekends respectively. The weekend cooling demand forecasts could significantly improve by using more stable weekend data.

The proposed model is now running in the campus and is forecasting both cooling and electricity load demands as a part of an integrated CCHP optimization platform. Further information from exogenous factors such as occupancy can improve the performance of the proposed model.

Acknowledgment

This work was supported by U.S. Department of Energy and National Energy Technology Laboratory via grants.gov in response to Clean Distributed Generation: Fuel/Feedstock Flexibility and Combined Heat and Power (Solicitation No. DE-PS26-08NT0004312-03).

References

[1] U.S. Department of Energy. Energy efficiency and renewable energy. Energy efficiency trends in residential and commercial buildings; October 2008.
[2] Energy Information Administration. Annual energy review 2004. Washington, DC: U.S. Department of Energy, Annual Report 061–003-01141-1; 2005.
[3] Wu DW, Wang RZ. Combined cooling, heating and power: a review. Prog Energy Combust Sci 2006;32(5–6):459–95. ISSN 0360-1285.
[4] San Martín JI, Zamora I, San Martín JJ, Aperribay V, Eguía P. Trigeneration systems with fuel cells. In: International conference on renewable energy and power quality (ICREPQ); 2008.
[5] Fang Fang, Wang QH, Shi Yang. A novel optimal operational strategy for the CCHP system based on two operating modes. IEEE Trans Power Syst 2012;27(2):1032–41.
[6] Liu Ye, Luo Yi. Optimal design of distributed CCHP system. In: International conference on transportation, mechanical, and electrical engineering (TMEE), 2011, 16–18 December; 2011. p. 1961–65.
[7] Cho Heejin, Mago Pedro J, Luck Rogelio, Chamra Louay M. Evaluation of CCHP systems performance based on operational cost, primary energy consumption, and carbon dioxide emission by utilizing an optimal operation scheme. Appl Energy 2009;86(12):2540–9.
[8] Liu Mingxi, Shi Yang, Fang Fang. Optimal power flow and PGU capacity of CCHP systems using a matrix modeling approach. Appl Energy 2013;102:794–802. ISSN 0306-2619.
[9] Chandan V, Do A, Jin B, Jabbari F, Brouwer J, Akrotirianakis A, et al. Modeling and optimization of a combined cooling, heating and power plant system, accepted for publication in proceedings of the American Control Conference; 2012.
[10] Mago PJ, Chamra LM. Analysis and optimization of CCHP systems based on energy, economical, and environmental considerations. Energy Build 2009;41(10):1099–106.
[11] Wang Jiang-Jiang, Jing You-Yin, Zhang Chun-Fa. Optimization of capacity and operation for CCHP system by genetic algorithm. Appl Energy 2010;87(4):1325–35.
[12] Hu Mengqi, Cho Heejin. A probability constrained multi-objective optimization model for CCHP system operation decision support. Appl Energy 2014;116(1):230–42.
[13] Huang Shyh-Jier, Shih Kuang-Rong. Short-term load forecasting via ARMA model identification including non-Gaussian process considerations. IEEE Trans Power Syst 2003;18(2):673–9.
[14] Muche Thomas. Optimal operation and forecasting policy for pump storage plants in day-ahead markets. Appl Energy 2014;113(January):1089–99.
[15] Taylor JW, McSharry PE, Buizza R. Wind power density forecasting using ensemble predictions and time series models. IEEE Trans Energy Convers 2009;24(3). p. 775, 782.
[16] García-Martos Carolina, Rodríguez Julio, Sánchez María Jesús. Modelling and forecasting fossil fuels, CO_2 and electricity prices and their volatilities. Appl Energy 2013;101(January):363–75.
[17] Goia Aldo, May Caterina, Fusai Gianluca. Functional clustering and linear regression for peak load forecasting. Int J Forecast 2010;26(4):700–11. ISSN 0169–2070.
[18] Serinaldi Francesco. Distributional modeling and short-term forecasting of electricity prices by generalized additive models for location, scale and shape. Energy Econom 2011;33(6):1216–26.
[19] Jain Rishee K, Smith Kevin M, Culligan Patricia J, Taylor John E. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. Appl Energy 2014;123(15 June):168–78. ISSN 0306-2619.
[20] Apadula Francesco, Bassini Alessandra, Elli Alberto, Scapin Simone. Relationships between meteorological variables and monthly electricity demand. Appl Energy 2012;98(October):346–56. ISSN 0306-2619.
[21] Mavromatidis Lazaros Elias, Bykalyuk Anna, Lequay Hervé. Development of polynomial regression models for composite dynamic envelopes' thermal performance forecasting. Appl Energy 2013;104(April):379–91.
[22] Inman Rich H, Pedro Hugo TC, Coimbra Carlos FM. Solar forecasting methods for renewable energy integration. Prog Energy Combust Sci 2013;39(6):535–76. ISSN 0360-1285.
[23] Kalogirou Soteris A. Artificial intelligence for the modeling and control of combustion processes: a review. Prog Energy Combust Sci 2003;29(6):515–66. ISSN 0360-1285.
[24] Mellit Adel, Kalogirou Soteris A. Artificial intelligence techniques for photovoltaic applications: a review. Prog Energy Combust Sci 2008;34(5):574–632.
[25] Kankal Murat, Akpınar Adem, Kömürcü Murat İhsan, Özşahin Talat Şükrü. Modeling and forecasting of Turkey's energy consumption using socio-economic and demographic variables. Appl Energy 2011;88(5):1927–39.
[26] Kulkarni Santosh, Simon Sishaj P, Sundareswaran K. A spiking neural network (SNN) forecast engine for short-term electrical load forecasting. Appl Soft Comput 2013;13(8).
[27] Kandil Nahi, Wamkeue René, Saad Maarouf, Georges Semaan. An efficient approach for short term load forecasting using artificial neural networks. Int J Elect Power Energy Syst 2006;28(8):525–30.
[28] Deihimi Ali, Orang Omid, Showkati Hemen. Short-term electric load and temperature forecasting using wavelet echo state networks with neural reconstruction. Energy 2013;57(1):382–401.

[29] Liu Hui, Tian Hong-qi, Pan Di-fu, Li Yan-fei. Forecasting models for wind speed using wavelet, wavelet packet, time series and Artificial Neural Networks. Appl Energy 2013;107(July):191–208.

[30] Lee Yi-Shian, Tong Lee-Ing. Forecasting nonlinear time series of energy consumption using hybrid dynamic model. Appl Energy 2012;94(June):251–6.

[31] Li Kangji, Su Hongye. Forecasting building energy consumption with hybrid genetic algorithm-hierarchical adaptive network-based fuzzy inference system. Energy Build 2010;42(11):2070–6.

[32] Lin Cheng-Ting, Chou Li-Der, Chen Yi-Ming, Tseng Li-Ming. A hybrid economic indices based short-term load forecasting system. Int J Elect Power Energy Syst 2014;54(January):293–305. ISSN 0142-0615.

[33] Kutner M, Nachtsheim C, Neter J, Li W. Applied linear statistical models. Irwin, Chicago, IL, US: McGraw-Hill; 2004.

[34] Cochrane D, Orcutt GH. Application of least squares regression to relationships containing auto-correlated error. J Am Statistical Assoc 1949;44(245):32–61.