Energy and Buildings 111 (2016) 184–194

Contents lists available at ScienceDirect

Energy and Buildings

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

Artificial neural network model for forecasting sub-hourly electricity usage in commercial buildings

Young Tae Chae [a,*], Raya Horesh [b], Youngdeok Hwang [b], Young M. Lee [b]

[a] Department of Architectural Engineering, Cheongju University, 298 Daesung-ro, Cheongju-si, Republic of Korea
[b] IBM, T.J. Watson Research Center, 1101 Kitchwan Road, Yorktown Heights, United States

ARTICLE INFO

Article history:
Received 1 October 2014
Received in revised form 13 October 2015
Accepted 14 November 2015
Available online 19 November 2015

Keywords:
Building electricity consumption
Short-term load forecasting
Artificial Neural Network
Bayesian regularization

ABSTRACT

Short-term load forecasting of building electricity usage is of great importance for anomaly detection on electricity usage pattern and management of building energy consumption in an environment where electricity pricing is dynamically determined based on the peak energy consumption. In this paper, we present a data-driven forecasting model for day-ahead electricity usage of buildings in 15-minute resolution.

By using variable importance analysis, we have selected key variables: day type indicator, time-of-day, HVAC set temperature schedule, outdoor air dry-bulb temperature, and outdoor humidity as the most important predictors for electricity consumption. This study proposes a short-term building energy usage forecasting model based on an Artificial Neural Network (ANN) model with Bayesian regularization algorithm and investigates how the network design parameters such as time delay, number of hidden neurons, and training data effect on the model capability and generality.

The results demonstrate that the proposed model with adaptive training methods is capable to predict the electricity consumption with 15-minute time intervals and the daily peak electricity usage reasonably well in a test case of a commercial building complex.

1. Introduction

Forecasting electricity load has become one of the more important topics recently, not only for the electricity power suppliers, but also for the consumers of the electricity, especially for commercial and industrial buildings. Building energy administrators, such as building owners or engineers, need to have an accurate forecast of energy demand in the near future or one day ahead to be able to better manage energy usage. Although the forecast time horizon can range from minutes to years, the short-term load forecast (STLF), especially for a period shorter than a day, has been more of an interest in the perspective of buildings because the utility prices may change by seasonality, time-of-use in on/off peak period, and contract demand [1,2].

A large variety of forecasting models and approaches, such as regression models, time series models, and machine-learning-based models have been developed and used. Regression models, in which the load forecasting is formulated usually as a linear function of input variables, have been effective to predict building energy consumption with number of experiments [3,4]. These models are attractive because the model components can have direct physical interpretation, such as the total load, base load, heating load, cooling load, and weather components, etc.

Time series models assume that the current and future energy usage is a function of the past observed energy usage. Examples include the autoregressive integrated moving average (ARIMA) [5], multiplicative autoregressive models [6], autoregressive moving average with exogenous input model (ARMAX) [7], Kalman filtering [8], and Fourier series model [9] among many. Although these models can predict short-term load pattern, they may be unstable in the case of nonlinear load or non-stationary conditions, due to its restrictive assumptions.

Other machine learning techniques have received significant attention in the context of STLF problems since the late 1980's including Support Vector Machine (SVM) model [10,11], neuro-fuzzy system [12], and Artificial Neural Networks (ANNs) for both electricity supply and demand side [13–15]. Several studies have shown that ANNs have produced better results compared with other approaches [16,17].

Although STLF models have been used in practice with different degree of success, these were mostly designed and evaluated with the hourly load forecasting. However, if the utility price depends on time of use, season, and the characteristics of building energy usage pattern in sub-hourly time resolutions [18], building

* Corresponding author. Tel.: +82-43-229-8479; fax: +82-43-229-8479.
E-mail addresses: ychae@cju.ac.kr, y.t.chae@gmail.com (Y.T. Chae).

http://dx.doi.org/10.1016/j.enbuild.2015.11.045
0378-7788/

Nomenclature

MSE Mean squared error, $\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(S_i - M_i)^2$ where,M: measured energy consumption in the period (n), S:simulated energy consumption in the period (n).

MBE Mean bias error, $\mathrm{MBE} = \frac{\sum_{i=1}^{n}(S_i - M_i)}{\sum_{i=1}^{n} M_i}$ where, M:measured energy consumption in the period (n), S:simulated energy consumption in the period (n).

APE Absolute percentage error, $\mathrm{APE} = \frac{|S-M|}{M} \times 100$ where, M:measured energy consumption in a time step, S:simulated energy consumption in a time step.

CV(RMSE) Coefficient of variation of the root-mean-squared error, $\mathrm{CV(RMSE)} = \frac{\mathrm{RMSE}}{A}$, $\mathrm{RMSE} = \sqrt{\mathrm{MSE}}$, and $A = \frac{\sum_{i=1}^{n} M_i}{n}$, where RMSE: root-mean-square-error, A:mean of measured data

engineers need a precise estimation of the peak electricity load in a day and the load shape to better control the utility costs. In addition, the energy forecast can provide useful information in making an electricity purchase plan when the building has an on-site electricity generation system with renewable energy sources under a dynamic electricity pricing grid [19]. For this problem, the hourly load forecast model may be insufficient to estimate the peak demand and to plan on-site energy generation.

Previous studies on STLF for the sub-hourly electricity consumption of buildings are limited. Escriva-Escriva et al. [14] proposed STLF model using ANNs to forecast the building energy consumption using 96 time steps for a day using an independent algorithm that first searches for four or more days which have similar outdoor air temperature patterns and types of day in the previous year for model training. Then the algorithm trains 96 ANNs for each time interval. This model requires an entire whole year's data set and the performance may not be stable when the energy consumption pattern has large daily or annual variability. Therefore, it is useful to explore a model that can perform well under more general setting, in particular, not requiring a large amount of input data for forecasting electricity usage of buildings.

For this need, we develop a short-term load forecasting model using data mining and machine learning technique while assuming limited availability of data. In particular, we investigate ANNs model to predict the energy consumption of a commercial building complex. This paper is structured as follows. The building complex for the case study and data processing approaches are illustrated in Section 2. The section also presents a feature extraction identifies key predictors that impacts the electricity consumption of the case, as well as a model selection procedure that selects the most suitable machine learning algorithm for this problem. Section 3 describes the sensitivity analysis study on the configuration variables of the model and implementation results, which are compared with the actual measurement of energy use. Lastly, in Section 4 summary of the results and future research directions are presented.

2. Methodological approach

2.1. Description of a case study: a building complex

All data set for this study was obtained from a building management system (BMS) of a commercial office building complex, and the data are periodically pulled into a relational database (IBM DB2$^{\mathrm{TM}}$). The site consists of three office buildings in urban area, each of which has different number of floors; five in building 1 (BLDG1), four in building 2 (BLDG2), and two in building 3 (BLDG3). A total floor area of 15,224 m^2 spreads over typical office area, small laboratories, cafeteria, parking garage, and small gymnasium (Fig. 1). Although the buildings are separated, they are all managed by one utility billing system.

Two absorption chiller systems provide chilled water in summer and hot water in winter for a constant air volume (CAV) system for each floor of BLDG1 and fan coil units (FCU) of perimeter zones of each floor of BLDG2. All three buildings have electric heat pump (EHP) systems with multi-indoor units. EHP systems are supplementary system incorporated with CAV systems for BLDG1, but operate as the main air-conditioning system for BLDG2 and BLDG3. BMS system monitors operational conditions of both primary/secondary system and EHPs. The system also controls all secondary system operation, whilst EHPs are locally controlled by occupants' indoor thermal demand. For the electricity usage monitoring, one main electric meter and several sub-meters are installed as illustrated in Fig. 2. The main meter measures electricity usage, both the instantaneous power in kW with a minute interval and aggregated electricity usage at every 15 minutes in kWh.

The short-term monitoring of electricity usage is very important to the building engineers since 15-minute peak energy consumption in the current year impacts the annual utility expenditure of the following year. According to the actual electricity price of the building complex in Table 1 from the electricity grid supplier [18], a monthly electricity bill is composed of a demand and meter charge for commercial buildings. The monthly demand charge, except in residential buildings, is set by a peak usage in three summer months, July, August, and September of the previous year. The supplier sets a building's peak usage as the maximum electricity usage for 15-minute intervals during the three months. If the peak electricity usage exceeded that of the last year at a certain time interval, the monthly demand charge is reset with the new peak usage for the next 12 months.

2.2. Data collection and processing

The relational database management system (RDBMS) has been used to collect and store environmental variables, BMS data, and electricity usage as illustrated in Fig. 3. Software engineers, building engineers, and the predictive model developers worked together to create data schema for RDBMS. Although the data set has over 1000 data points with 20 different measurement types, some variables may not eventually be used for electricity usage forecast. For example, the supply and return air temperature of an air handling unit are important in estimating the space thermal load and the electricity usage by the primary/secondary system, but they are not available at the time when forecast is made. Thus after taking the data availability into account, the predictors are divided into three categories: environmental data, time indicator, and operational condition, as illustrated in Table 2.

For the environmental variables, the database retrieves data from local weather forecasting service. The data include outdoor air temperature, relative humidity, wind speed and direction, sky condition, and precipitation type in every 3-hour interval for the next 72 hours. The forecast data is updated eight times a day. RDBMS system converts string variables like sky condition and wind