Energy and Buildings 82 (2014) 437–446

Contents lists available at ScienceDirect

Energy and Buildings

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

Modeling heating and cooling loads by artificial intelligence for energy-efficient building design

Jui-Sheng Chou*, Dac-Khuong Bui

Department of Civil and Construction Engineering, National Taiwan University of Science and Technology, 43, Sec. 4, Keelung Rd., Taipei, 106, Taiwan

ARTICLE INFO

Article history:
Received 15 April 2014
Received in revised form 15 July 2014
Accepted 16 July 2014
Available online 27 July 2014

Keywords:
Cooling load
Heating load
Energy performance
Energy-efficient building
Artificial intelligence
Data mining

ABSTRACT

The energy performance of buildings was estimated using various data mining techniques, including support vector regression (SVR), artificial neural network (ANN), classification and regression tree, chi-squared automatic interaction detector, general linear regression, and ensemble inference model. The prediction models were constructed using 768 experimental datasets from the literature with 8 input parameters and 2 output parameters (cooling load (CL) and heating load (HL)). Comparison results showed that the ensemble approach (SVR +ANN) and SVR were the best models for predicting CL and HL, respectively, with mean absolute percentage errors below 4%. Compared to previous works, the ensemble model and SVR model further obtained at least 39.0% to 65.9% lower root mean square errors, respectively, for CL and HL prediction. This study confirms the efficiency, effectiveness, and accuracy of the proposed approach when predicting CL and HL in building design stage. The analytical results support the feasibility of using the proposed techniques to facilitate early designs of energy conserving buildings.

1. Introduction

Since all three primary economic sectors, *i.e.*, industry, transportation, and building, have high energy use, conservation of energy is a critical task [1]. Buildings consume a substantial share of global energy consumption. Therefore, substantial energy savings can be realized by buildings that are properly designed and operated. Heating load (HL) and cooling load (CL) are measures of energy that must be added or removed from a space by heating ventilation and air conditioning (HVAC) system to provide the desired level of thermal comfort within space. Therefore, early predictions of building CL and HL can help engineers design energy-efficient buildings.

Building simulation tools have been widely used to predict and analyze building energy consumption. Building energy consumption and load simulation software developed in recent years include DOE-2 [2], ESP-r [3], Energy Plus [4], and DeST (Designer's Simulation Toolkit) [5]. Although they accurately predict building load in many projects [6–8], the prediction values differ according to the simulation software used to predict the energy use of occupied buildings [9]. Building simulation programs are complicated and time-consuming due to the diverse disciplines involved and the use of varying parameter settings. Current building simulation tools are also difficult to use for identifying and comparing the impacts of variables that affect the observed quantity of interest [10].

Because the building sector accounts for a large proportion of global energy consumption and has enduring adverse environmental impacts, studies of energy performance of buildings (EPB) have recently increased [11]. Rapid increases in building energy consumption in recent decades have also resulted from rising living standards. Notably, the building sector now consumes more than 30% of the total energy worldwide [12]. In Europe, the building sector accounts for 40% of energy use and 36% of CO_2 emissions [13] whereas, in the United States, the building sector accounts for 38.9% of the total primary energy requirement (PER) [14]. In China, building stocks accounted for about 24.1% of total national energy use in 1996 and for 27.5% in 2011. Building stocks are projected to increase to about 35% by 2020 [15].

Moreover, for a modern city like Hong Kong, 60% of carbon emissions are produced by electricity generation, and buildings account for 89% of total electricity consumption [16]. Hence, increasing the energy efficiency and energy performance of buildings is essential for mitigating the increasing demand for additional energy supply as well as CO_2 emission. While many researchers have developed methods of optimizing the operation of various components in HVAC and refrigerating systems [17,18], accurately predicting

* Corresponding author. Tel.: +886 2 2737 6321; fax: +886 2 2737 6606.
E-mail addresses: jschou@mail.ntust.edu.tw, rayson.chou@gmail.com (J.-S. Chou), M10205810@mail.ntust.edu.tw (D.-K. Bui).

http://dx.doi.org/10.1016/j.enbuild.2014.07.036

building HL and CL is critical for effective energy conservation strategies.

In practice, identifying parameters that substantially affect building energy consumption can help optimize a building design. The influential parameters, *e.g.*, relative compactness [19], climate [20], surface area, wall area, and roof area [21], orientation [22], can be grouped into two main categories: the physical properties of a building and meteorological conditions. These factors make the relationship between EPB and its influential parameters very complicated. Consequently, accurately predicting HL and CL of building is a challenging task.

The artificial intelligence (AI) inference model has recently proven to be a viable alternative approach to predict EPB [23]. The AI is employed to develop models that simulate the human inference processes. Thus, AI can infer new facts from previously acquired information and can adaptively change in response to changes in historical data. Tsanas and Xifara [10] stated that AI not only obtains solutions very quickly, it also assists building designers in analyzing the influence of input parameters. Many studies have explored the use of AI models for predicting various interests in the context of EPB [9,24–30]. However, most works have reported unsatisfactory error rates, and most have considered only a few factors that affect building energy use.

Therefore, the objective of this research is to compare the performance of various AI techniques, including support vector regression (SVR), artificial neural network (ANN), classification and regression tree (CART), chi-squared automatic interaction detector (CHAID) and general linear regression (GLR). The best performing models were then combined into ensemble models. A k-fold cross-validation algorithm was used for validation. A synthesized performance index and hypothesis testing were used to compare performance measures between the proposed models and those in previous works. The contribution to the body of knowledge is the development of an AI technique that can predict building CL and HL with improved accuracy in predicting energy consumption and can facilitate early building design for the energy conservation.

The remainder of this paper is organized as below. The following section introduces the study context by reviewing the related literature, including studies of EPB and predictive techniques. Section 3 then describes research methodology and evaluation methods. Section 4 describes the building information and experimental data obtained in this study. Section 5 presents modeling processes, discusses prediction results, and compares model performance. Concluding remarks and research contributions are given in the final section.

2. Literature review

Recently, researchers have studied the use of AI techniques for predicting energy consumption. Many statistical and artificial intelligence techniques for inverse modeling of building heating and cooling loads have been developed [28,31–34]. Particularly, ANNs are very convenient and easy to use by an ordinary operator after the model has been established and are the most popular AI technique in many applications [24,26,27,35–39]. The ANN models do not require definition of explicit relationships between inputs and outputs as in conventional regression. They can model building heating and cooling loads of complex systems from independent or dependent parameters [40].

Olofsson et al., for example, combined ANN with quasi-physical description to predict the annual supplied space heating demand for numerous small single-family buildings [26]. Additionally, Aydinalp et al. [41] summarized the literature on ANN modeling and used outdoor temperature and electric home appliance usage as inputs in the ANN model for predicting annual residential energy usage. Their predictions of the annual energy consumption of 247 houses in Canada had an overall correlation coefficient of 0.909 with the annual actual energy consumption. In Kwok et al., an ANN model was employed to predict energy use by a Hong Kong office building. The best root-mean-squared-percentage-error (RMSPE) was 11.409% [31]. In Hou et al., an ANN based on data-fusion technique was used to forecast air-conditioning load and achieved a small relative error (below 4%) [27]. Although these studies show that ANN can achieve a moderate fit in predicting HL and CL, model performance is generally unsatisfactory.

In recent years, SVR, a variation of support vector machine (SVM), has been widely used in forecasting and regression [42]. Dong et al. evaluated the use of SVR to predict energy consumption in tropical regions [28]. The SVR was used to make hourly forecasts of building cooling load. Their study showed that SVR has superior prediction accuracy compared to conventional back propagation neural networks. Likewise, Li et al. used SVR to predict the hourly cooling load of a building [29,43]. The authors showed that SVR was better than ANNs. In addition, Jain et al. used SVR to forecast the energy consumption of multi-family residential buildings [32]. A sensor-based forecasting model using SVR was applied to an empirical data set for a multi-family residential building in New York City.

Many other AI techniques have been proposed for improving energy consumption prediction accuracy in the energy field. In Yu et al., CART provided accurate predictions of building energy demand with low errors [9]. Li et al. hybridized genetic algorithm with adaptive network-based fuzzy inference system to enhance accuracy in forecasting building energy consumption [33]. Tsanas and Xifara used random forest (RF) technique to estimate CL and HL in residential buildings [10]. Other proposed machine learning techniques for identifying and predicting system behavior include neural network system [44], general regression neural network [24], regression models [34,45], and hybrid system [33].

The above studies agree that AI model performs satisfactorily in predicting energy performance in building. Nevertheless, most works have used individual forecasting models rather than investigating the power of ensemble models. Moreover, the prediction performance of the aforementioned AI techniques needs further study. Therefore, the objective of this study was to fill this gap by using these models individually and in combination (ensemble models) to predict CL and HL *via* cross-fold validation and multiple performance measures.

3. Methodology

3.1. Support vector regression

Support vector machine was first introduced by Vapnik [46]. This supervised learning method generates an SVM by input-output mapping functions from a labeled training dataset. This function solves both classification and regression problems. Typically, regression is performed by using SVR, variation of SVM, to find a function $f(x)$ that has at most ε deviation from the actually obtained targets y_i for all the training data, and at the same time is as flat as possible. The input is first mapped onto an m-dimensional feature space by using non-linear functions as follows:

$$f(x,\omega)=\langle\omega,x\rangle+b \quad \text{with} \quad \omega\in\chi, b\in\chi \tag{1}$$

where $\langle\bullet,\bullet\rangle$ denotes the dot product in χ.

Thus, the goodness of the $f(x)$ can be estimated based on the loss function $L(x)$ as follows:

$$L(x,\omega)=[y,f(x,\omega)]=\begin{cases}0 & \text{if} \quad |y-f(x,w)|\le\varepsilon\\ |y-f(x,w)| & \text{otherwise}\end{cases} \tag{2}$$