Building and Environment 86 (2015) 177–190

Contents lists available at ScienceDirect

Building and Environment

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

Comparisons of inverse modeling approaches for predicting building energy performance

Yuna Zhang [a,*], Zheng O'Neill [b], Bing Dong [c], Godfried Augenbroe [a]

[a] College of Architecture, Georgia Institute of Technology, Atlanta, GA, USA
[b] Department of Mechanical Engineering, University of Alabama, Tuscaloosa, AL, USA
[c] Department of Mechanical Engineering, The University of Texas at San Antonio, San Antonio, TX, USA

A R T I C L E I N F O

Article history:
Received 29 October 2014
Received in revised form
24 December 2014
Accepted 26 December 2014
Available online 14 January 2015

Keywords:
Inverse modeling
Energy performance prediction
Gaussian regression
Comparison

A B S T R A C T

In building retrofit projects, retrofit savings can be estimated by comparing building energy use before and after installing Energy Conservation Measures (ECMs). A complicating factor is that there is no direct measurement of the reduced energy use that is solely attributable to the retrofit. Indeed, simple comparisons by subtracting the post-retrofit energy use from the pre-retrofit would ignore the impact of other factors, such as weather and occupancy with constantly changing patterns, on the total building energy use. Data-driven models (i.e., derived by inverse modeling approaches) that are trained with monitored pre-retrofit building data can be used as the baseline models in a retrofit project. However, to be effective, the baseline energy models must be capable of singling out the impact of ECMs and ignoring the influence of other factors. A commonly used method to achieve this goal is to develop a statistical model that correlates energy use with weather and other independent variables.

This paper first reviews four mainstream baseline data-driven energy models used to characterize building energy performance: change-point regression model, Gaussian process regression model, Gaussian Mixture Regression Model, and Artificial Neural Network model, These models are then applied to an office building to predict the Heating, Ventilation, and Air-Conditioning (HVAC) hot water energy consumption. Several model accuracy measures such as R^2, *RMSE*, *CV-RMSE*, and sensitivity to sample frequency, and reliability, are evaluated and compared.

1. Introduction

Nowadays, fossil fuel based energy resources are decreasing and energy demand is increasing. This requires an immediate attention for improved energy efficiencies and better use of alternative energy resources. Environment related concerns, such as ozone layer depletion and global warming (e.g., increased carbon dioxide level in the air), necessitate a greater priority for such attention. According to the building energy data book 2011 [1], commercial and residential buildings accounted for 41% of the primary energy consumption in the United States in 2010. Enhancing building efficiency represents one of the easiest, most immediate and most cost effective ways to reduce the nation's energy consumption. One effective method to reduce energy consumption in the building sector is to retrofit existing buildings. Decisions for energy efficiency retrofits are typically made based on predictions of how much energy and money a retrofit will save and the expected payback period of certain Energy Conservation Measures (ECMs). In performance contracts provided by the typical Energy Service Company (ESCO), the service fee to the ESCO is related to predicted and actual savings associated with ECMs.

Fig. 1 illustrates the general methodology to define energy savings from the ASHRAE guideline 14, Measurement of Energy and Demand Savings [2]. The energy savings brought by ECMs is equal to the baseline energy use under post-retrofit conditions minus the measured post-retrofit building energy use.

The retrofit energy savings cannot be measured directly. However, it can be calculated by comparing energy consumption measurements in pre-retrofit and post-retrofit periods. Simple comparison of pre-retrofit and post-retrofit energy consumption cannot differentiate the impact from ECMs and weather and occupancy schedules variations. The ASHRAE Guideline 14 [2] presents three approaches for measuring savings: whole building

* Corresponding author. 723 Cherry St. NW Atlanta, GA 30332, USA. Tel.: +1 (404) 713 3022.
E-mail address: yuna.zhang@gatech.edu (Y. Zhang).

http://dx.doi.org/10.1016/j.buildenv.2014.12.023
0360-1323/

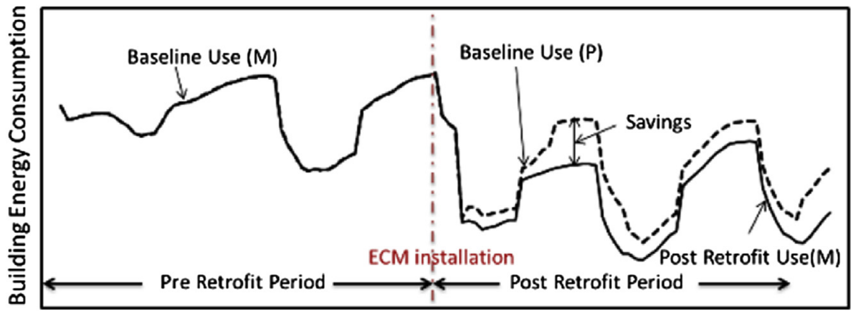

Fig. 1. Illustration of energy savings for retrofit.

approach, retrofit isolation approach, whole building calibrated simulation approach.

2. Background

In the last decade, there is an increasing recognition that data-driven models can serve as building baseline energy models [3–10]. Such baseline models can be used for determining retrofit savings, energy system fault diagnostics, and acquiring physical insight into the operating patterns. Energy conservation retrofits are typically initiated based on predictions of how much energy and money a retrofit could save. Therefore, defining a baseline prior to a retrofit is essential.

The 2013 ASHRAE Handbook-Fundamentals [11] classified modeling approaches into two basic categories: forward (classical) modeling and inverse (data-driven) modeling. The forward modeling approach generally takes the physical parameters that describe the building as input, which can include building location, local weather, geometry, envelope construction materials, operational schedule, and HVAC system type, etc. The forward modeling approach is typically used in the design phase to facilitate building designers to make early design decisions. Inverse models take the monitored building energy consumption data (and possible other monitored behavior data) as inputs and are expressed in terms of one or more driving variables and a set of empirical parameters. Typically, a model form is a priori assumed and measured data are used to find the parameters that provide the best fit for the chosen model form and data set. Inverse models have been widely used in building retrofit projects [9], performance monitoring and system fault diagnostics [12], control strategy development [13], and online control applications [14].

Generally speaking, two most commonly used approaches to define the baseline suggested by the ASHRAE Guideline 14 [2] are a calibrated simulation data-driven regression analysis and calibrated computer simulation tools such as EnergyPlus [15], eQuest [16], etc. Compared to the first approach, the second approach has two major limitations: (1) it requires detailed building information which is not easy to get and (2) creating and calibrating the predictive building energy model is time-consuming and labor intensive. The data-driven model statistically derives a relationship between a set of inputs (e.g., the ambient conditions) and outputs (e.g., energy consumption). In practice, inverse models trained by actual building energy consumption data can provide reliable estimation and have been widely adopted for measurement and verification, and ongoing commissioning of building performance [17].

Currently, the most popular approaches for inverse modeling of building energy performance are based on regression techniques. This is typically done in an ad-hoc manner relying on the assumption that the nonlinear energy behavior arising from complex multivariable relationships between ambient conditions, occupancy levels, and building operating conditions can be captured adequately by the regression.

Literature shows a blossoming development of data-driven baseline modeling methods. ASHRAE [18] introduces constant-base degree-days models. Degree days are calculated as the sum of the differences between daily average temperatures and the base temperature. Heating degree days and cooling degree days are used extensively in calculations related to building energy consumption. Heating degree days are a measure of how much (in degrees), and for how long (in days), the outside air temperature is below a certain level. They are commonly used in calculations of the heating energy consumption. On the other hand, cooling degree days are a measure of how much (in degrees), and for how long (in days), the outside air temperature is above a certain level. They are commonly used in calculations of cooling energy consumption. In a constant-base degree-days (CBDD) model, the degree day is used as an independent variable while building total electricity consumption as a dependent variable. Fels [19] and Rabl et al. [20] utilized variable-base degree-day (VBDD) method to estimate retrofitting energy use. Letherman [21] presented a method to predict the heating and cooling energy demand with degree hours method.

Thamilseran and Haberl [22] introduced a bin method model as a baseline model. In the bin method, the bin model predicts average hourly corresponding pre-retrofit electricity use during any hour of day in the post-retrofit period. Then, the predictions are compared with measured hourly post-retrofit energy use for every hour of day to get the total energy savings from a certain energy conservation retrofit. The bin method model is based on Box–whisker–mean plots theory [23]. It is convenient to use Box–whisker–mean plots to display the main features of a set of data and these plots facilitate the comparison of multiple data sets. Fig. 2 shows a typical bin method model application.

Thomas [24] introduced four linear regression methods to forecast residential building energy demand. Kwok [25] conducted a study using artificial neural networks to predict building cooling load. Kissock [26], Krarti et al. [27] utilized neural networks to estimate energy and demand savings from retrofitting of commercial buildings. Dhar et al. [28] developed a temperature based Generalized Fourier series model to estimate hourly heating and cooling energy use in commercial buildings. Dong et al. [6] and Solomon et al. [29] use support vector regression method to predict building energy consumption. Heo and Zavala [9] and Zhang et al. [30] implemented a Gaussian Process regression baseline model to access the total building energy consumption in the post-retrofit phase, which is a less time-consuming and easy to accomplish process. Srivastav et al. [31] presents a Gaussian Mixture Regression Model to predict the building energy use with parameterized and locally adaptive uncertainty quantification for simulation data only. Granderson and Price [10] reviewed and compared five whole