Energy and Buildings 105 (2015) 302–313

Contents lists available at ScienceDirect

Energy and Buildings

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

Estimation models of heating energy consumption in schools for local authorities planning

Alfonso Capozzoli [a,*], Daniele Grassi [a], Francesco Causone [b]

[a] TEBE Research Group, Department of Energy, Politecnico di Torino, Corso Duca degli Abruzzi 24, 10129 Torino, Italy
[b] Department of Energy, Politecnico di Milano, Via Lambruschini 4, 20156 Milano, Italy

ARTICLE INFO

Article history:
Received 3 March 2015
Received in revised form 2 July 2015
Accepted 11 July 2015
Available online 21 July 2015

Keywords:
School buildings
Multiple Linear Regression
Heating energy estimation
Decision tree
Public authorities planning

ABSTRACT

Large building stocks should be well managed, in terms of ordinary activities and formulating strategic plans, to achieve energy savings through increased efficiency. It is becoming extremely important to have the capability to quickly and reliably estimate buildings' energy consumption, especially for public authorities and institutions that own and manage large building stocks. This paper analyses the heating energy consumption of eighty school buildings located in the north of Italy. Two estimation models are developed and compared to assess energy consumption: a Multiple Linear Regression (MLR) model and a Classification and Regression Tree (CART). The CART includes interpretable decision rules that enable non-expert users to quickly extract useful information to benefit their decision making. The output of MLR model is an equation that accounts for all of the major variables affecting heating energy consumption. Both models were compared in terms of Mean Absolute Error (MAE), Root Mean Square error (RMSE), and Mean Absolute Percentage error (MAPE). The analysis determined that the heating energy consumption of the considered school buildings was mostly influenced by the gross heated volume, heat transfer surfaces, boiler size, and thermal transmittance of windows.

1. Introduction

1.1. Energy consumption analysis in school buildings

Buildings are responsible for about 40% of the total energy consumption in developed countries [1]. In countries like Italy, about 60% of the existing building stock is more than 40-years-old [2]. A rapid and substantial energy retrofit program is therefore required for these existing buildings. There are about two million households in Italy that live in buildings requiring either demolition and rebuilding or refurbishment. Directive 2010/31/EU (EPBD recast) requires buildings, or parts of buildings, to meet a minimum energy performance or be subject to a retrofit or refurbishment. These requirements may be met by renovating a building's envelope and systems, but an effective management can also significantly impact a building's energy consumption. The EU Directive [3–5] requires public buildings to play an exemplary role in terms of energy savings. Public utilities and buildings that are typically owned and managed by municipalities include: street lighting, schools, administrative buildings, public transport, and sport centres, such as swimming pools and gymnasiums [6]. Local governments would clearly benefit from having access to energy consumption data. Further, being able to understand the savings potential of these assets would help to prioritise energy and environmental projects and better illuminate their financial aspects [7]. According to the US Department of Energy, school buildings constitute a major part of the public building stock. Around 25% of the energy expenses in schools could be saved through better building designs and more energy-efficient technologies, combined with improvements in operation and maintenance [8].

De Santoli et al. [9] evaluated the energy performance of public schools in Rome. They defined intervention strategies to reduce energy consumption and identified action priorities by means of a simple payback time analysis (PBT). Dimoudi et al. [10] conducted an energy simulation to study the energy savings potential of school buildings in Greece. Kim et al. [11] analysed the energy consumption of some elementary schools in South Korea by utilising monitoring data from January 2006 to December 2010. They determined that electrical energy was consumed the most, followed by gas and oil. During the monitoring period, electrical energy continued to increase its relevance on the energy breakdown because of cooling/heating system replacements. These and other studies were carried out in recent years to estimate the energy

* Corresponding author. Tel.: +39 011 0904413; fax: +39 011 0904499.
E-mail addresses: alfonso.capozzoli@polito.it (A. Capozzoli), grassi.daniele88@gmail.com (D. Grassi), francesco.causone@polimi.it (F. Causone).

http://dx.doi.org/10.1016/j.enbuild.2015.07.024
0378-7788/

Nomenclature

β	estimated coefficient of Multiple Linear Regression model
CART	classification and regression tree
D–W	Durbin–Watson test
E	error associated to the tree
EUI	energy use intensity
EUI_{st}	standard energy use intensity
$EUI_{st,s}$	standard and specific energy use intensity
F	Fisher–Snedecor test
HDD_{conv}	conventional heating degree days
HDD_{real}	real heating degree days
MAE	mean absolute error
MAPE	mean absolute percentage error
MLR	multiple linear regression
POW	boiler size (heat input)
R^2	coefficient of determination
$R^2{}_{adj}$	adjusted coefficient of determination
RMSE	root mean square error
SUR	heat transfer surface
$U_{windows}$	thermal transmittance of windows
VIF	variance inflation factors
VOL	heated gross volume

consumption of school buildings. The literature shows that there are two main approaches for estimating a building's energy consumption: the direct approach or the inverse approach. The first approach calculates the energy demand by running an energy simulation under a steady state or dynamic conditions. The second approach uses historical data to produce data driven models that estimate the energy consumption.

A large part of the current literature focuses on the inverse approach. Analysts and decision makers have access to several applications of new or recast versions of existing models. Corrado et al. [12] defined a simplified method for predicting future consumption based on climatic and real use data on a stock of 120 school buildings. Corgnati et al. [13] then validated this method, using another stock of 118 schools, as did Ariaudio et al. [14]. Amber et al. [15] gathered daily values of a school building's electrical consumption on the Southwark campus of the London South Bank University from 2007 to 2013 and then developed a multiple regression model to estimate future daily electrical consumption. Beusker et al. [16] evaluated the energy consumption of schools and sports facilities in Germany using different linear and nonlinear regression models. Thewes et al. [17] presented a regression model with categorical variables to predict the electrical and heating energy consumption of school buildings in Luxembourg.

Innovative techniques, including machine learning, data mining, and knowledge discovery in databases, have also been successfully applied to building energy consumption data in recent years [18]. In particular, a classification tree which consists of a multi-stage decision-making process that is useful to categorise observations in a finite number of classes, can be a powerful estimation tool. This method has not yet been applied in other studies to estimate the energy consumption of school buildings.

In this paper, the heating energy consumption of a school building stock located in the north of Italy is analysed using a Multiple Linear Regression (MLR) model and a Classification and Regression Tree (CART). Both MLR model and CART are data driven models that have been successfully applied to estimate a building's energy demand. Nevertheless, the outcome of MLR is an equation, while the output of CART are decision rules that allow users to quickly extract relevant information [19]. This characteristic substantially changes the practical applicability of the two models.

1.2. Implementation of multiple regression analysis and classification tree for buildings' energy use estimations

In recent years, numerous researchers successfully employed multiple regression models as a tool for energy consumption estimations. Al-Garni et al. [20] correlated electrical energy consumption with relevant climatic variables (air temperature, relative humidity, solar radiation), and variable occupant populations through statistical methods (regression model) to forecast the overall electrical energy consumption in Eastern Saudi Arabia. Aranda et al. [21] developed three regression models to predict the Spanish banking sector's annual energy consumption. The first model can be used to estimate the energy consumption of the whole banking sector, while the second estimates the energy consumption for branches under conditions of a low severity winter climate and the third under conditions of a high severity winter climate. The variance reported for the three models is 58%, and 68%, respectively. Korolija et al. [22] developed regression models to predict the annual heating, cooling, and electrical auxiliary energy consumption of five different types of HVAC systems (variable air volume—VAV, constant air volume—CAV, fan–coil system with dedicated air – FC, and two chilled ceiling systems with dedicated air, radiator heating, and either embedded pipes – EMB – or exposed aluminium panels—ALU) for office buildings in the UK. Freire et al. [23] used independent variables like energy consumption, ventilation and air conditioning power, outdoor temperature, relativity humidity, and total solar radiation to develop a regression equation to predict the indoor air temperature and relative humidity for two buildings with low and high thermal mass. The literature demonstrates therefore that regression models offer a robust methodology for estimating a building's energy consumption (e.g., heating, cooling, lighting, etc.).

Decision trees belong to the *machine learning algorithms* family. This method is recognised as an emerging analysis tool and is currently receiving plenty of attention from applied research. Yu et al. [18] used the decision tree to classify and predict building energy consumption. This method was applied to Japanese residential buildings for predicting and classifying building Energy Use Intensity (EUI) levels based on training data. This tool was then evaluated on a sample test.

Zhao et al. [24] used a C4.5 decision tree algorithm, locally weighted naïve Bayes and support vector machine, to classify occupant behaviour and to create schedule models for building energy simulation. The results show that the C4.5 algorithm correctly classified 90% of individual behaviour and this allowed getting closer to the real group schedule. Mikučionienė et al. [25] used a decision tree to increase the sustainability and improve the criteria for evaluating energy efficiency measures in a public building renovation in Lithuania. By analysing and weighting each variable (related to insulation of external walls, roof insulation, heating substation renovation, reconstruction of the entire heating system, and installation of a ventilation system with exhaust air heat recovery), the researchers created a decision tree to evaluate the influence of each variable on energy consumption. The results show that this algorithm reduces the amount of data that must be understood by transforming it into a more compact form while still preserving the basic substance. The researchers determine whether the data are characterised by well-separated object classes and finally, this algorithm determines the precise relationship between attributes and their class.

In this paper, two different estimation models are developed using a database consisting of 80 school buildings located in the province of Turin. The estimation models include climatic, envelope