Energy and Buildings 49 (2012) 380–387

Contents lists available at SciVerse ScienceDirect

Energy and Buildings

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

Multiple regression models to predict the annual energy consumption in the Spanish banking sector

Alfonso Aranda*, Germán Ferreira, M.D. Mainar-Toledo, Sabina Scarpellini, Eva Llera Sastresa

CIRCE – Centre of Research for Energy Resources and Consumption – University of Zaragoza, Mariano Esquillor Gómez, 15 – 50018 Zaragoza, Spain

ARTICLE INFO

Article history:
Received 6 December 2011
Received in revised form 7 February 2012
Accepted 24 February 2012

Keywords:
Banking sector
Multiple regression model
Energy consumption prediction

ABSTRACT

This paper presents a regression analysis of energy consumption in the banking sector. In our case study, the target area is the Spanish banking sector, for which we divide the available data into a prediction and a validation subset. Power models were developed using test data from 55 banks. From the analysis, three models were obtained; where the first proposed model can be used to predict the energy consumption of the whole banking sector, while the rest of the models estimate the energy consumption for branches with low winter climate severity (Model 2) and high winter climate severity (Model 3). Models 2 and 3 differ from the first model in that they need independent variables measured in situ. As a result, the uncertainty of the response variable in the function of the independent variables is reduced by 56.8% for the first model and by 65.2% and 68.5% for the second and third proposed models, respectively. The validation of the first model, which is the model with the lowest determination coefficient, shows that this model is appropriate for predicting the energy consumption of bank branches with good energy consumption performance and detecting inefficiencies in bank branches with poor energy consumption performance.

1. Introduction

Energy consumption and demand trends show that for the 27 members of the European Union (EU 27) greater energy efficiency will be required by 2020. The most recent projections, using 2009 as the reference, show that primary energy consumption will decrease to 70,170,768 TJ (1676 Mtoe) by 2020 (6,950,088 TJ [166 Mtoe] less than the reference year 2005), which is 8,457,336 TJ (202 Mtoe) more than the 2020 objective of achieving a 20% reduction in energy consumption [1]. The total annual energy consumption in the EU 27 increased continuously until 2007, with a total of 48,470,583.6 TJ (1157.7 Mtoe), and the service sector was responsible for 39.5% of the energy consumption, with 11,915,632.8 TJ (284.6 Mtoe) for households, 1,332,480.257 TJ (27.8 Mtoe) for agriculture and 6,079,233.6 TJ (145.2 Mtoe) for services, etc. After the economic downturn that started in 2008, there was a reduction in economic activity and, consequently, energy consumption, but there was also a slowdown on the progression towards energy efficiency [2]. The economic recovery period will see an increase in equipment refurbishing that will allow for progress in energy efficiency to take place.

In 2007, the service sector accounted for nearly 10% of the total final energy consumption in Spain – 408,799.152 GJ (9764 ktoe) – [3], with the largest consumers being in the office sector, and an increasing trend is expected in the coming years. Greenhouse gases (GHG) emitted by the commercial and institutional sector were equivalent to approximately 8.2 million tonnes of CO_2 [4]. Consequently, energy savings in this sector offer the best means of reducing the energy demand. Several European countries, led by the European Commission, are currently interested in improvements in energy efficiency in buildings and in reducing carbon dioxide equivalent emissions (CO_2 eq). As result of Directive 2002/91/EC in the last few years, intense development is taking place in Spain with the intention of reducing carbon dioxide emissions. Since the enactment in November 2007 of the Royal Decree 47/2007, of 19th January, approving the Basic Procedure to certify energy efficiency in new-construction buildings and in certain retrofitting buildings (in which more than 1000 m^2 or more than the 25% of the building envelope is refurbished), buildings that undergo this Royal Decree must be qualified in terms of energy efficiency at the project level for the work to be completed. Directive 2010/31/EU regarding the energy performance of buildings (recast) from 19 May 2010 repeals the current directive that requires the certification of buildings that are for rent or sale. Consequently, a new Royal Decree is expected in the near future. Additionally, recent initiatives focusing on both the improvement of energy efficiency as well as the provision of renewable energy sources in the building sector by Energy Service Companies (ESCOs) are expected to have a positive effect on these issues in the short term [5].

* Corresponding author.
E-mail address: alaranda@unizar.es (A. Aranda).

0378-7788/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.enbuild.2012.02.040

Nomenclature

Latin alphabet

AGE	age
ANOVA	analysis of variance
ATMs	automated teller machines
CO_2 eq	carbon dioxide emissions equivalent
CI	confidence interval
Cp	Mallows' statistic
CTE	building technical code, Código Técnico de la Edificación
E	mathematical expectation
EIF	energy inefficiency factor
ENERCON	annual energy consumption
ESCOs	energy services companies
F-test	*F*-test statistic
GHG	greenhouse gases
GLASUR	glazed surface in the façade
H_0	null hypothesis
HEIGHT	office height
HVAC	heating, ventilation and air conditioning
NCASH	number of automated teller machines
NEMP	number of employees
p	number of independent variables plus one, which corresponds to the intercept
p-value	*p*-value statistic
r	Pearson's correlation coefficient
R^2	determination coefficient
R^2 (adj)	adjusted determination coefficient
s	number of coefficients
S	residuals standard deviation
SE	standard error
SURF	office surface area
SUSEV	summer climatic severity
t-test	Student's *t*-test statistic
Var	variance
VIF	variance inflation factor
WCSEV	winter climatic severity
x_i, x_j	independent variable, predictor, predictive variable
X	independent variable matrix
Y	dependent variable, response variable

Greek symbols

α	confidence level
β	estimates of the regression coefficients
ε, ε_a, ε_b	random errors or perturbations
σ^2	population variance

In spite of significant trends and interest of the research community in the energy performance improvement of the service sector, there is a lack of information regarding the specific service sectors that have dedicated research efforts to study their energy consumption in more detail [6–13]. G.N. Spyropoulos and C.A. Balaras reported that among the different office subcategories, banks and other financial offices are the most energy intensive in the US, but similar data for European buildings has not been published for the banking sector [14].

Modelling techniques that predict the energy performance of buildings have been used, including multiple regression methods, artificial neural networks, decision trees or Fourier series models [15–19]. In this paper a regression model has been selected to find a compromise between the simplicity of the evaluation method and the accuracy in the result without requiring a considerable amount of input data and simulation energy [17]. Multiple regression is used frequently in research; the present work aims to identify explicative variables to develop a model in which the chosen variables influence the response and the variables that do not contribute relevant information are rejected.

The objectives of this study are to develop a regression model that determines how efficient or inefficient a bank branch is in terms of energy consumption, depending on its construction characteristics, climatic area and energy performance, by predicting its annual energy consumption. Furthermore the energy requirements for heating and cooling demand of bank branches are supplied only by electricity. The mathematical model permits researchers to predict the energy consumption without widespread analysis, and the model is validated and used to detect energy-inefficient bank branches in Spain and to propose energy saving measures that could reduce energy consumption. The results provide relevant information on the energy performance of the Spanish banking sector and contribute new data for the energy performance of the service sector.

2. Methodology

An inference analysis was developed to obtain three multiple regression models for the prediction of the annual energy consumption in the Spanish banking sector. The aim was to first obtain a model that will serve as a pre-diagnostic tool for energy performance in the bank branches analysed, and the model was based on easy-to-obtain variables, precluding the necessity of a walk-through audit. Additionally, alternative models were calculated to obtain a more accurate energy performance based on independent variables that represent features that need an in situ measurement or more detailed information than is collected in a walkthrough energy audit finding a balance between accuracy and feasibility in obtaining the predictors.

Relevant independent variables that define the energy consumption were selected to develop models that were obtained by means of regression models, which were validated and discussed for future predictions and reproducibility [20]. Regression coefficients were estimated using the least squares method. This method estimates the regression coefficients by minimising the sum of the squares of the deviations to the proposed regression model [21]. A regression equation is as shown:

$$\hat{Y} = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \quad (1)$$

where $\hat{Y}$ is the fitted value and $\beta_0, \beta_1, \ldots,$ and β_p are the estimations of the regression parameters.

The real value for *Y* is:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p + \varepsilon \quad (2)$$

where ε is the random error [22].

$\beta_0, \beta_1, \ldots, \beta_p$ describe the expected change in the predicted variable *Y* in response to a unitary change in x_i when the rest of predictors remain constant [23].

It is not recommended to predict the response variable for a set of values for predictors that are out of the range of data used for the regression equation obtained, which would lead to an extrapolation error [24]. The graphical and regression analysis were performed using Minitab [25] and SPSS [26].

2.1. Sampling

Fifty-five bank branches were selected from the 12 total climatic areas across Spain. Using the Minitab software, the sample size was obtained according to the 1 sample *Z*-test method outlined by Douglas C. Montgomery [22] for a standard deviation of 7224 for the population analysed, for a confidence level of 90% and an acceptable