Energy and Buildings 109 (2015) 209–216

Contents lists available at ScienceDirect

Energy and Buildings

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

# Using multiple regression analysis to develop energy consumption indicators for commercial buildings in the U.S.

Shideh Shams Amiri, Mohammad Mottahedi, Somayeh Asadi *

*Department of Architectural Engineering, The Pennsylvania State University, 104 Engineering Unit A, University Park, PA 16802, United States*

ARTICLE INFO

*Article history:*
Received 29 July 2015
Received in revised form
28 September 2015
Accepted 30 September 2015
Available online 3 October 2015

*Keywords:*
Multiple regression analysis
Monte Carlo simulation
DOE-2
Sensitivity analysis
Energy consumption

ABSTRACT

Multiple regression analysis plays an important role in evaluating the energy consumption of buildings. These models are commonly used to assess the energy performance of commercial buildings and to predict any potential for energy consumption reduction. In this study, the building simulation software DOE-2 was used to predict energy consumption. A total of 17 key building design variables were identified related to building envelope, building orientation, and occupant schedule. Since, building energy consumption depends on many operational and design parameters; large numbers of simulations are needed to generate data for the multiple regression models. To tackle this problem, a randomized approach was adopted to reduce the required number of simulations examining the whole design space. Monte Carlo simulation technique was used to generate thirty thousand combinations of design parameters, covering the full range for each climate region. In order to implement the Monte Carlo simulation, an in-house computer program was developed to interface with DOE-2 energy simulation software. Stepwise regression was used to reduce the number of parameters and only include the most effective parameters. *R* statistical analysis program was also used to develop the set of linear regression equations. Parametric study and sensitivity analysis between levels of most effective parameters were performed. The developing models can be used to estimate the energy consumption of office buildings in early stages of design.



## 1. Introduction

Building energy efficiency has come to the forefront of political debates due to high energy prices and climate change concerns. Energy consumption in commercial buildings accounts for a significant proportion of worldwide energy consumption. In the U.S., commercial buildings consume 19% of the total national energy use and emit more than one billion metric tons of carbon dioxide [1]. The demand for energy by the commercial sector is projected to increase by 1.2% per year from 2006 to 2030, driven by trends in population and economic growth [1]. Due to this relatively large share of a nation's total energy consumption, any increase in the energy efficiency of the energy systems for commercial buildings would result in notable energy savings and emissions reductions. Improving energy efficiency in new commercial buildings is one of the easiest and lowest cost options to decrease a building's energy use, owner operating costs, and carbon footprint.

In the traditional design process, the project in the early stages is very unclear, as design detail is still low and uncertain. However, it is at this phase that the designer has to study design alternatives, aiming to satisfy key design requirements, such as urban landscape integration, esthetics, functional qualities and energy performance. Therefore, modeling techniques predicting energy performance of buildings are playing an important role in the designs and analysis of energy-efficient buildings. These models simulate the effect of different design parameters such as building characteristic, HVAC system, occupant's behavior, and weather conditions on building energy consumption [2].

Several researches have been carried out to assess the influencing factors of building energy consumption using multiple regression analysis [3–6]. The statistical models are a special class of simplified models that are obtained by regression techniques from dynamic models. They combine the speed of simple models and the precision of dynamic models [7]. They were used to evaluate the energy demand as a function of the overall heat transfer coefficient [8,9] or of the shape [10]. The principle of the statistical models is to propose a function which relates the energy demand to environmental variables (e.g., temperatures, solar radiation, wind speed, occupation) and design variables (e.g., wall type and thickness) and to identify the coefficients of this function by a regression method [11–14].

* Corresponding author.
*E-mail address:* asadi@engr.psu.edu (S. Asadi).

http://dx.doi.org/10.1016/j.enbuild.2015.09.073
0378-7788/

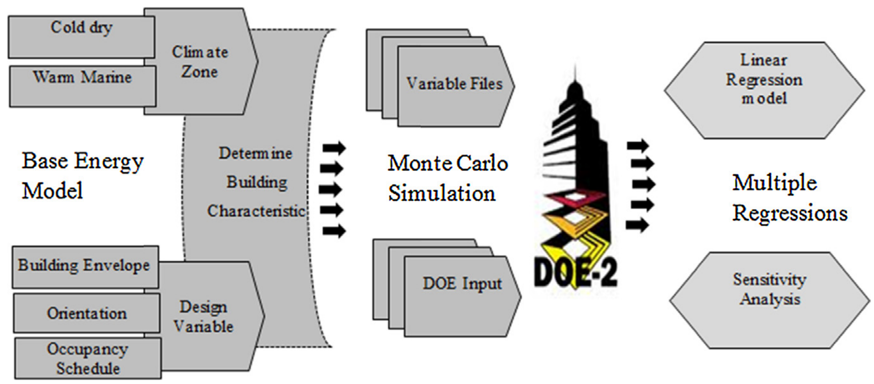


**Fig. 1.** Framework of the study.

By establishing the linear regression equations, Carlo and Lamberts [15] developed regression equations to assess the effects of different building envelope influencing factors such as building volume indicator, roof heat transfer coefficient, solar heat gain coefficient, etc. on the electricity consumption on commercial buildings in Brazil. In another study conducted by Bansal and Bhattacharya [16], the detailed simulation results were used to develop simplified regression equations which can predict a single zone building annual energy demand in addition to maximum heating and cooling loads for the central India weather conditions. The developed regression equations showed a very good fit with coefficient of determination above 0.9. Lam et al. [17] also used DOE-2 energy simulation software to study the effects of 12 key building design variables on energy consumption in five different climate regions in China. It was found that the developed regression equations could also be used to estimate the energy savings when different building schemes were considered. In another study conducted by Lam et al. [18] regression equations were developed to predict the electricity consumption of commercial buildings in Hong Kong. They found that shading coefficient and window to wall ratio are the most significant envelope variables affecting energy consumption.

In addition to analyzing the relationship between building energy consumption and their influencing factors, regression models can also be used for the prediction of energy load and energy consumption. Catalina et al. [19] developed regression models to predict the monthly heating load in residential buildings in France. They considered the building envelop U-value, the window to wall ratio, and building shape factor as design factors in their study. In another study, Li et al. [20] used EnergyPlus simulation software to estimate the energy consumption index of a typical office building in Hong Kong. They developed linear and nonlinear regression models to assess the energy performance of different building envelope designs with daylighting controls.

Most of the previous researches seldom focused on the relationship between energy consumption of office buildings and the variations of building envelop influencing factors in different climate regions in the United States. To surmount these hurdles, present study developed regression equations to predict energy savings in different climate regions in the United States. This study considers seventeen influencing factors including building construction materials and their thickness, building orientation, glass type, etc. Of particular interest in this article is the application of the parametric study in order to understand and determine the most dominant set of input parameters affecting total energy consumption in commercial buildings.

The structure of the article is as follows. Firstly, the methods for conducting parametric study using two-dimensional Monte Carlo analysis are described. Then the development of an energy model of an office building within DOE-2 is presented. Ten Thousand simulations for each climate zone are performed to create a comprehensive dataset covering the full ranges of design parameters. At the end, the results of the energy simulation are used to generate a set of regression equation to predict the energy consumption.

## 2. Methodology

Building energy simulation is conducted using dynamic energy simulation software DOE-2.2. The DOE-2.2 was chosen primarily due the fact that the graphical user interface in eQUEST software enhanced our ability to generate multiple simulation models with different geometries. Also, DOE-2.2 has a comprehensive documentation that was needed for manipulating the simulation input files during Monte Carlo simulation process. Key parameters relevant to building envelope, such as orientation, occupancy schedules, fenestration, etc. were considered in this study. The remaining inputs, such as HVAC system, lighting load, etc. which are needed to run an annual energy simulation were considered as constant parameters. Fig. 1 shows the framework of the study.

### 2.1. Baseline model development

One of the important factors in developing energy models for building is a deep understanding of its physical and operational characteristics. In this study, a typical rectangular office building (Fig. 2) is created as a baseline to compare its energy demand in two different climate regions (i.e. cold, dry and warm, marine) in the United States. ASHRAE 90.1 was used as a reference to design this building. This building is a two-story office with the area of 2322.6 $m^2$. The total number of occupants was 105 and one HVAC zone was defined in this study. The cooling and heating systems were chilled water coils and hot water coils, respectively. Furnace was installed in this building as a heating source. Net floor to ceiling window ratio and floor-to-floor window ratio were 53.3 and 40%, respectively. A full description of the design variables has been presented elsewhere [2].

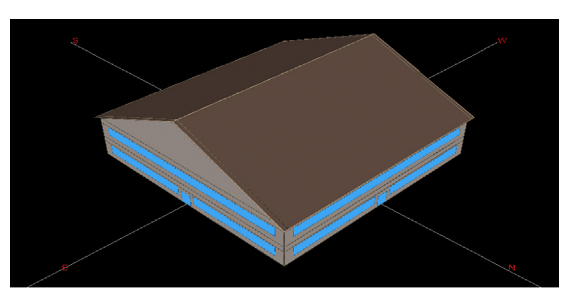

**Fig. 2.** Building shape.

**Table 1**
Regression coefficients.

| Parameters | San Jose | | | | | Billings | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Levels | Cooling | | Heating | | Cooling | | Heating | |
| | | β | P-value | β | P-value | β | P-value | β | P-value |
| Occupant schedule | (Intercept) | 27.5 | **<.001** | 130.73 | **<.001** | 32.45 | **<.001** | 15.99 | **<.001** |
| | b | −2.33 | **<.001** | −18.57 | **<.001** | −1.94 | **<.001** | −4.01 | **<.001** |
| | c | −3.79 | **<.001** | −7.19 | **<.001** | −5.28 | **<.001** | −0.45 | **<.001** |
| | d | −5.4 | **<.001** | −22.66 | **<.001** | −6.27 | **<.001** | −3.82 | **<.001** |
| | e | −2.44 | **<.001** | 2.44 | **<.001** | −3.06 | **<.001** | 1.07 | **<.001** |
| | f | −5.03 | **<.001** | −15.97 | **<.001** | −5.73 | **<.001** | −3.22 | **<.001** |
| | g | −5.52 | **<.001** | −5.86 | **<.001** | −6.85 | **<.001** | 0.4 | **<.001** |
| | h | −7.49 | **<.001** | −21.08 | **<.001** | −8.47 | **<.001** | −3.29 | **<.001** |
| Exterior wall* | ASH Wall-12 | 0.75 | **<.001** | −43.54 | **<.001** | 3.21 | **<.001** | −8.39 | **<.001** |
| | ASH Wall-19 | 1.13 | **<.001** | −47.26 | **<.001** | 3.88 | **<.001** | −8.48 | **<.001** |
| | ASH Wall-1 | 0.72 | **<.001** | −45.24 | **<.001** | 3.16 | **<.001** | −8.71 | **<.001** |
| | ASH Wall-25 | 0.84 | **<.001** | −42.23 | **<.001** | 3.26 | **<.001** | −8.01 | **<.001** |
| | ASH Wall-27 | 1.15 | **<.001** | −46.76 | **<.001** | 3.9 | **<.001** | −8.4 | **<.001** |
| | ASH Wall-29 | 1.09 | **<.001** | −25.15 | **<.001** | 3.16 | **<.001** | −3.93 | **<.001** |
| | ASH Wall-30 | 0.56 | **<.001** | −43.72 | **<.001** | 2.89 | **<.001** | −8.69 | **<.001** |
| | ASH Wall-3 | 0.1 | **<.001** | −24.5 | **<.001** | 1.41 | **<.001** | −5.29 | **<.001** |
| | ASH Wall-6 | 0.21 | **<.001** | −27.77 | **<.001** | 1.7 | **<.001** | −5.78 | **<.001** |
| Exterior wall absorbance | Medium | 0.32 | **<.001** | −0.9 | **<.001** | 0.44 | **<.001** | −0.46 | **<.001** |
| | Dark | 0.79 | **<.001** | −2.33 | **<.001** | 1.07 | **<.001** | −1.06 | **<.001** |
| Interior wall* | ASH Wall-14 | 0.49 | **<.001** | −1.07 | **<.001** | 0.86 | **<.001** | 0.59 | **<.001** |
| | ASH Wall-25 | 0.49 | **<.001** | −1.84 | **<.001** | 0.82 | **<.001** | 1.54 | **<.001** |
| | ASH Wall-28 | 0.45 | **<.001** | −2.14 | **<.001** | 0.84 | **<.001** | 1.9 | **<.001** |
| | ASH Wall-29 | 0.33 | **<.001** | −1.99 | **<.001** | 0.56 | **<.001** | 2.18 | **<.001** |
| | ASH Wall-36 | 0.47 | **<.001** | −1.85 | **<.001** | 0.85 | **<.001** | 1.62 | **<.001** |
| | ASH Wall-3 | 0.38 | **<.001** | −0.66 | **<.001** | 0.61 | **<.001** | 0.52 | **<.001** |
| | ASH Wall-5 | 0.21 | **<.001** | −0.01 | 0.929 | 0.28 | **<.001** | −0.16 | **<.001** |
| Floor construction | Concrete 4 in | −0.17 | **<.001** | 0.26 | **<.001** | −0.32 | **<.001** | −0.57 | **<.001** |
| | Concrete 6 in | −0.3 | **<.001** | 0.01 | 0.926 | −0.54 | **<.001** | −1.01 | **<.001** |
| | Concrete 8 in | −0.38 | **<.001** | −0.07 | 0.295 | −0.67 | **<.001** | −1.19 | **<.001** |
| Ground floor interior finish | Vinyl tile | −0.22 | **<.001** | 1.76 | **<.001** | −0.45 | **<.001** | −0.33 | **<.001** |
| | Ceramic/stone tile | −0.26 | **<.001** | 1.76 | **<.001** | −0.53 | **<.001** | −0.37 | **<.001** |
| Ground floor construction | Concrete 4 in | 0.22 | **<.001** | 0.87 | **<.001** | 0.33 | **<.001** | 0.02 | 0.485 |
| | Concrete 6 in | 0.18 | **<.001** | 0.56 | **<.001** | 0.27 | **<.001** | −0.01 | 0.662 |
| | Concrete 8 in | 0.12 | **<.001** | 0.47 | **<.001** | 0.19 | **<.001** | −0.07 | **0.025** |
| Building orientation | 180° | 0.12 | **<.001** | 0.48 | **<.001** | 0.03 | **0.007** | 0.07 | **0.015** |
| | 270° | −0.01 | 0.068 | −0.01 | 0.943 | 0 | 0.964 | −0.04 | 0.17 |
| | 360° | 0.12 | **<.001** | 0.57 | **<.001** | 0.03 | **0.022** | 0.04 | 0.21 |
| Top floor Batt insulation | R-19 | 0.07 | **<.001** | −1.13 | **<.001** | 0.1 | **<.001** | −0.21 | **<.001** |
| | R-30 | 0.09 | **<.001** | −1.46 | **<.001** | 0.12 | **<.001** | −0.27 | **<.001** |
| | R-45 | 0.12 | **<.001** | −2 | **<.001** | 0.17 | **<.001** | −0.3 | **<.001** |
| Floor interior finish | Carpet & no pad | 0.04 | **<.001** | −1.65 | **<.001** | 0.35 | **<.001** | 0.69 | **<.001** |
| | Vinyl tile | 0.02 | **0.013** | −0.08 | 0.277 | 0.08 | **<.001** | 0.08 | **0.012** |
| | Stone 1 in | −0.07 | **<.001** | −0.03 | 0.677 | −0.12 | **<.001** | −0.21 | **<.001** |
| Ceiling insulation | Wool Batt R-19 | 0.04 | **<.001** | −2.65 | **<.001** | 0.1 | **<.001** | −0.27 | **<.001** |
| | Wool Batt R-30 | 0.08 | **<.001** | −4.64 | **<.001** | 0.2 | **<.001** | −0.47 | **<.001** |
| Top floor ceiling exterior insulation | Polyurethane 1.25 in | 0.04 | **<.001** | −0.79 | **<.001** | 0.07 | **<.001** | −0.08 | **0.022** |
| | Polyurethane 1/2 in | 0.01 | 0.104 | −0.23 | **0.004** | 0.03 | **0.009** | −0.1 | **0.002** |
| | Polyurethane 2 in | 0.06 | **<.001** | −0.98 | **<.001** | 0.08 | **<.001** | −0.2 | **<.001** |
| | Polyurethane 3in | 0.08 | **<.001** | −1.45 | **<.001** | 0.11 | **<.001** | −0.27 | **<.001** |
| Roof absorbance | Medium | 0 | 0.65 | | | 0.02 | **0.037** | −0.06 | **0.015** |
| | Dark | 0.01 | **0.031** | | | 0.04 | **<.001** | −0.06 | **0.031** |
| Top floor ceiling interior finish | Gypsum board 5/8 in | | | 0.31 | **<.001** | | | | |
| | Gypsum 3/4 in | | | 0.34 | **<.001** | | | | |
| Ceiling interior finish | Gypsum board 5/8 in | | | 0.11 | 0.067 | | | | |
| | Gypsum 3/4 in | | | 0.16 | **0.008** | | | | |
| Roof construction | ASH Roof-16 | | | | | | | 0.06 | 0.188 |
| | ASH Roof-20 | | | | | | | −0.02 | 0.644 |
| | ASH Roof-22 | | | | | | | 0.06 | 0.181 |
| | ASH Roof-26 | | | | | | | −0.06 | 0.228 |
| | ASH Roof-28 | | | | | | | −0.04 | 0.432 |
| | ASH Roof-2 | | | | | | | −0.07 | 0.122 |
| | ASH Roof-3 | | | | | | | 0.07 | 0.145 |
| | ASH Roof-35 | | | | | | | 0 | 0.986 |
| | ASH Roof-9 | | | | | | | 0.03 | 0.53 |
| | $R^2$/adj. $R^2$ | | .989/.989 | | | | | | |

The bold values show the parameters that are statistically significant.
[b]08:00:00 AM to 06:00:00 PM (Monday–Thursday) +HVAC[1].
[c]07:00:00 AM to 05:00:00 PM (Monday–Thursday) +HVAC [1].
[d]07:00:00 AM to 04:00:00 PM (Monday–Friday) +HVAC[1].
[e]08:00:00 AM to 05:00:00 PM (Monday–Friday) +HVAC[2].
[f]08:00:00 AM to 06:00:00 PM (Monday–Thursday) +HVAC[2].
[g]07:00:00 AM to 05:00:00 PM (Monday–Thursday) +HVAC[2].
[h]07:00:00 AM to 04:00:00 PM (Monday–Friday) +HVAC[2].
[*]From ASHRAE construction from Tables 13 and 18 in chapter 26 ASHRAE Handbook, Fundamentals, 1989.

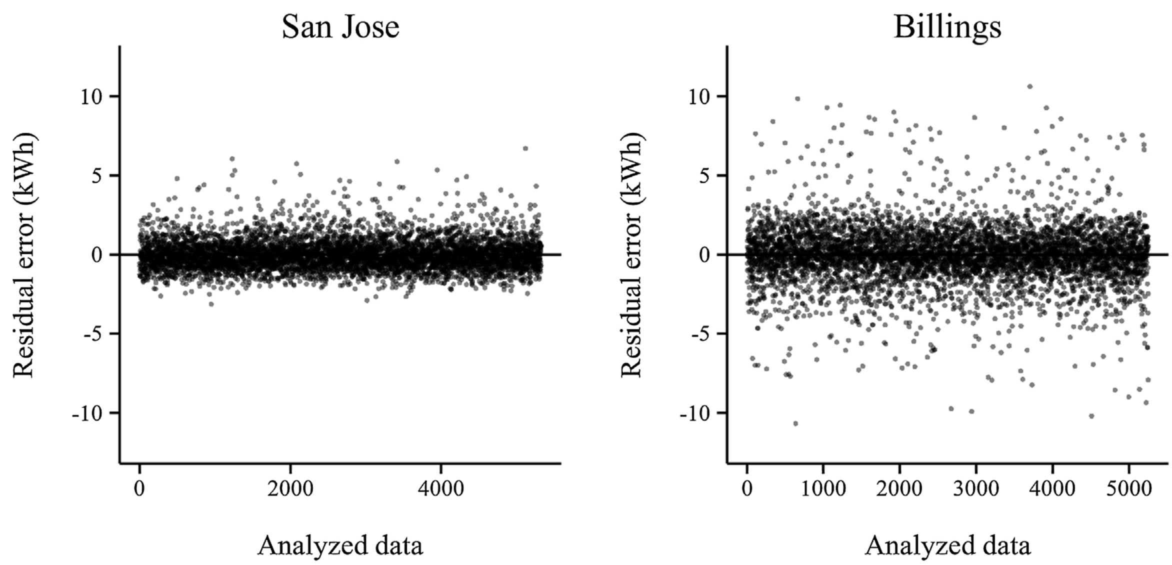


**Fig. 3.** Residual plots.

Typical meteorological year 3 (TMY3) containing 8760 h of weather data suitable for use with the DOE-2 energy program was utilized for both climate regions. Billings (cold, dry) and San Jose (warm, marine) were selected as a typical city in each climate zone.

### 2.2. Design variables

In this study different components of the building including walls, roofs, ceilings, foundation, and floors were considered. Typical occupancy, HVAC, lighting, miscellaneous equipment, and service hot water schedules were provided by 90.1-2004 User's Manual [22]. Eight different occupant schedules were considered in this study as follows:

1) 08:00:00 AM to 05:00:00 PM (Monday–Friday) +HVAC[1]
2) 08:00:00 AM to 06:00:00 PM (Monday–Thursday) +HVAC [1]
3) 07:00:00 AM to 05:00:00 PM (Monday–Thursday) +HVAC [1]
4) 07:00:00 AM to 04:00:00 PM (Monday–Friday) +HVAC[1]
5) 08:00:00 AM to 05:00:00 PM (Monday–Friday) +HVAC[2]
6) 08:00:00 AM to 06:00:00 PM (Monday–Thursday) +HVAC[2]
7) 07:00:00 AM to 05:00:00 PM (Monday–Thursday) +HVAC[2]
8) 07:00:00 AM to 04:00:00 PM (Monday–Friday) +HVAC[2]

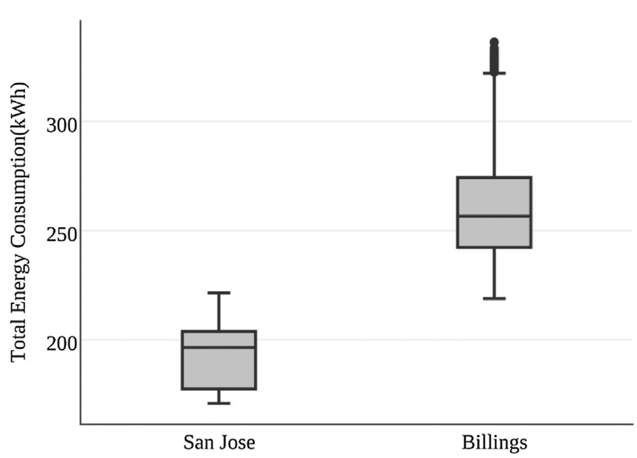


**Fig. 4.** Distribution of total energy consumption for two climates.

where for the case 1–4, the HVAC[1] system turns on 1 h before working hours and turn off 1 h after working hours, and for the case 5–8, the HVAC[2] system is on only during working hours. Lighting, miscellaneous equipment, and service hot water schedules were matched to occupancy schedules with limited usage during unoccupied times. In addition, various building envelope variables including insulation, flooring, glazing, materials, as well as building orientation were considered as input parameters. In total, 17 building design variables were considered in this study.

### 2.3. Implementing Monte Carlo simulation

The Monte Carlo method is a sampling-based technique that performs multiple model runs with random samples generated from the input distributions. The Monte Carlo method provides approximate solutions to a variety of mathematical problems by performing statistical sampling experiments on a computer. To perform Monte Carlo simulation, an in-house python program [21] was developed to create a uniformly distributed random dataset from available levels of each variable. For each set of variables, a new DOE-2 input file was generated and the energy simulation was performed for each scenario. The required data was extracted automatically from the generated report files and stored in a data frame. In total, 30,000 simulations were carried out to cover the wide range of possible parameters.

### 2.4. Regression-based techniques

In this study multiple linear regression models were developed to predict annual energy consumption for a given set of values as independent variables. The multiple linear regression analysis is the extension of the simple linear regression that can relate the variations in multiple predictors to the response. The developed regression model for predicting the heating and cooling has the following form (Eq. (1)):

$$Y = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i3} + \cdots + \beta_p x_{ip} + \varepsilon \quad (1)$$

where $Y$ is the predicted heating, cooling, or total energy, and $\beta$ is the corresponding regression coefficient. The regression coefficient

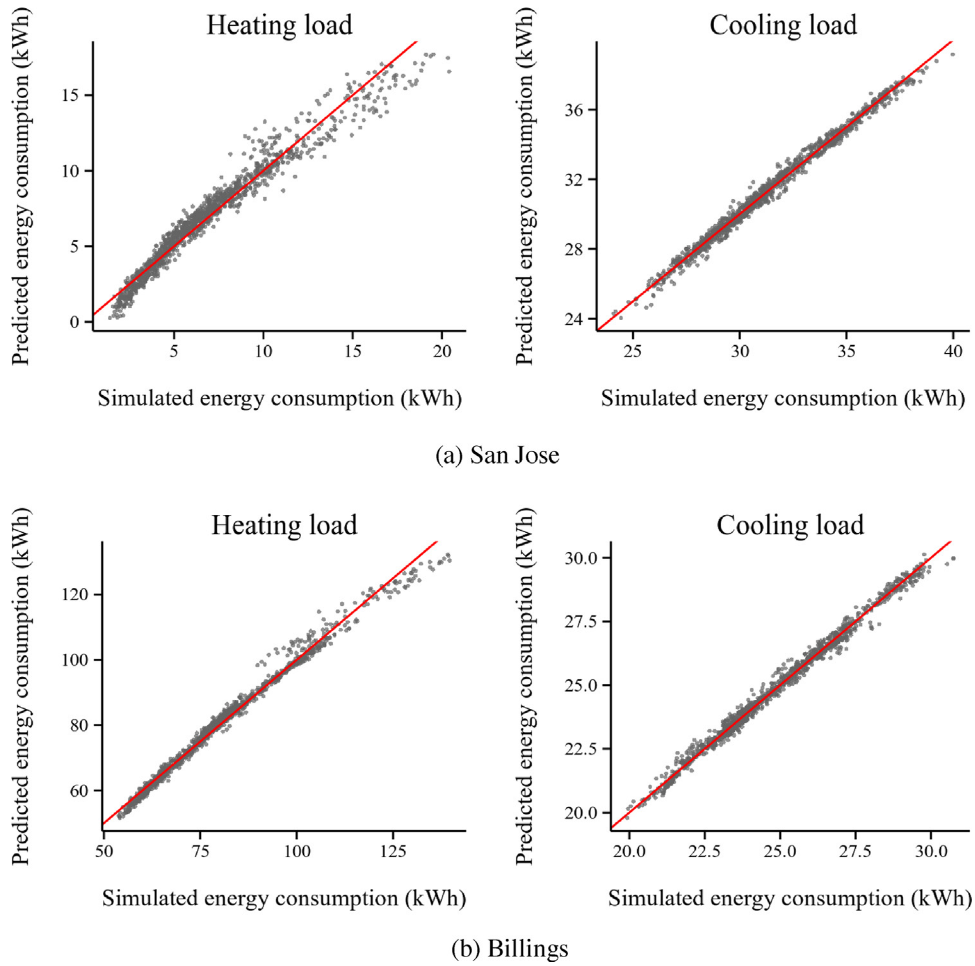


**Fig. 5.** Validation of total energy, heating and cooling load in two climates.

is determined using least square technique to minimize the sum of squared error (RSS), which is defined by:

$$RSS = \sum_{i=1}^{n} \left( y - \beta_0 - \beta_1 x_{i1} - \beta_2 x_{i2} - \beta_3 x_{i3} - \cdots - \beta_p x_{ip} \right) \quad (2)$$

Since all the predictors in this study were qualitative variable, a dummy coding was used for each level to include the qualitative variables in the regression model. The available data was divided into two groups, eighty percent 80% of the runs were selected randomly and used to train the regression model and remaining 20% were used to validate the developed regression model.To find the best regression model that explains the variation in building energy consumption, stepwise regression [23] was used. Stepwise regression is an automatic method that is useful when the number of explanatory variables is large and it is not possible to fit all the possible models. In this study both forward selection and backward elimination were performed to find the best model.

## 3. Results and discussion

### 3.1. Regression models

The goal of regression analysis is to create a model to predict the energy consumption in the building. Also, the regression models make it possible to study how building energy demand fluctuates with changes in individual building parameters. To this end, two regression models were developed to predict the building's heating and cooling energy consumption for each climate. Regression coefficients ($\beta$) are presented in Table 1. The regression coefficients show the changes in energy consumption for each variable while all other predictors remained constant. The $p$-values show the significance of the corresponding regression coefficient. Traditionally, predictors with the $p$-values equal or less than 0.05 are considered significant in regression model. 95% confidence interval was considered in the present study.

From Table 1, it can be seen that even after performing the stepwise regression, still some of the parameters have a $p$-value >0.05. For example, the ASH Wall-5 construction for interior wall has $p$-value greater than 0.05 for heating load regression model while the cooling load regression model has $p$-values less than 0.01.

It is necessary to ensure that regression models satisfy all the assumptions of correlation including normality, independence, linearity and homoscedasticity. Fig. 3 presents sample residual errors

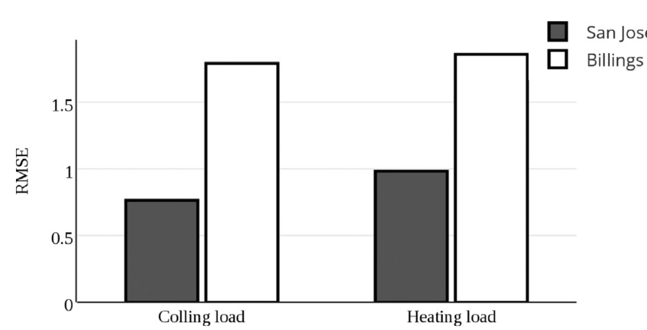


**Fig. 6.** Root mean squared error (RMSE) for, heating and cooling load.

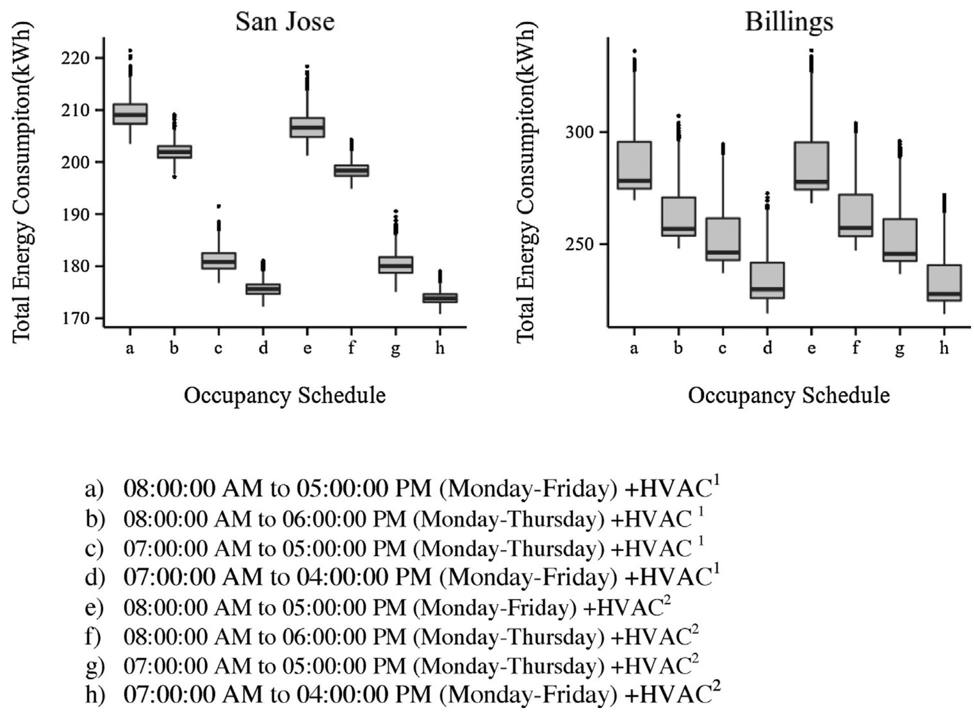


**Fig. 7.** Distribution of total energy consumption vs occupancy.

for a sample regression model in each climate, which confirms that the residual errors are randomly distributed and show no discernible pattern.

Fig. 3 shows that the observed residual errors have higher variation in Billings while they are more concentrated in San Jose. These variations can be explained with the distribution of the total energy in both climates (Fig. 4). In San Jose, which has a warm, marine climate, the median annual energy consumption is equal to 211 kWh with very few outliers and minimum and maximum energy consumption close the median. However, the total energy consumption in Billing has higher variation and many outliers which can be explained by the cold, dry with hot summers and cold, dry winters. The higher variation in the energy consumption resulted in a regression model with higher residual errors (Fig. 3).

### 3.1.1. Regression model validation

Model validation is perhaps the most important step in the model building sequence. The validation process can include examining the goodness of fit of the regression, investigating whether the regression residuals are random, and testing whether the model's predictive performance depreciates substantially when applied to data that were not used in model estimation. In this section, the data that have not been used to build the model was utilized for validation. Fig. 5 illustrates the predicted heating and cooling energy consumptions against their simulated values. As it can be seen there is a good agreement between simulated and predicted energy data. In addition, the stepwise regression shows a good fit. The obtained $R^2$ values are presented in Table 2.

Another measure that quantifies the goodness-of-fit of the regression model is root mean square error (RMSE), which is presented in Fig. 6 for cooling and heating load. The larger RMSE in Billings could be explained by higher variation in temperature during winter and summer. As it can be seen, the RMSE for the cooling loads are approximately equal in both climates, while the RMSE for the heating load is higher in Billings and constitutes great portion of the total load RMSE.

**Table 2**
Coefficient of determination ($R^2$).

| | Cooling load | Heating load |
|---|---|---|
| San Jose | 97.8% | 95.5% |
| Billings | 98.9% | 99% |

### 3.1.2. Interaction between parameters

To better understand the relation between predictors and the response, it is necessary to determine the presence of interaction among predictors. If two variables of interest interact, the relationship between each of the interacting variables and the response depends on the value of the other interacting variable. Multicollinearity can be detected using generalized variance-inflation factor (GVIF). The GVIF indicates the degree to which the confidence interval for that variable regression parameter is expanded relative to a model with uncorrelated predictors. As a general rule, GVIF > 4 indicates a multicollinearity problem. The results presented in Table 3 indicate that the multicollinearity is not a problem in the developed regression models.

## 3.2. Sensitivity analysis

It is important to know how different values of an independent variable affect the energy consumption in the building. Sensitivity analysis is a technique used in predefined boundaries that will show the influence of one or two input variable. In this study, the regression coefficients can be used as the quantitative measure to determine sensitivity of the dependent variable to changes in independent variables. In this study, the developed regression models for two climatic regions in addition to heating and cooling load can be utilize to understand the effect of each parameter in each case. Also, the following information can help us to identify the contradicting effects of individual parameters in different climate zones and on energy demand. Figs. 7 and 8 show regression coefficients for occupancy schedule and exterior wall for heating and cooling loads in two climate regions.

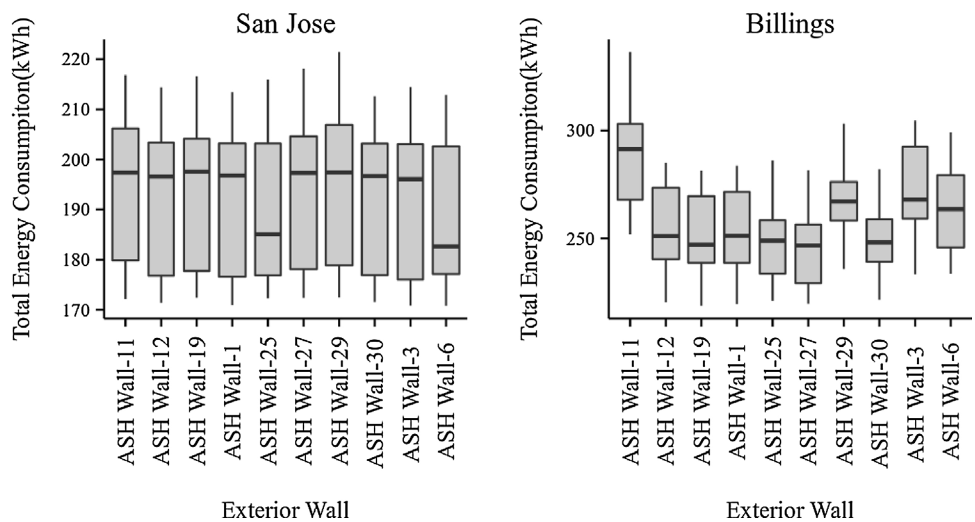


**Fig. 8.** Distribution of total energy consumption vs. exterior wall.

As it can be seen in Fig. 7, the change in occupancy schedule significantly affects the total energy consumption. The occupancy schedule resulted in higher variation in Billings due to compounding effect of sever climate. As expected, the occupancy schedule *b* which represents 24 h HVAC operation has the highest median energy consumption, while occupancy schedule *e*, with operational HVAC during working hours, has the lowest median energy consumption. Similar conclusion can be made in relation to the effect of exterior wall on total energy consumption (Fig. 8). As expected, the variation in energy consumption is higher in Billings, the ASHRAY WALL-11 LAY has the highest mean energy consumption. However, the energy consumption is evenly distributed in San Jose.

a) 08:00:00 AM to 05:00:00 PM (Monday–Friday) +HVAC[1]
b) 08:00:00 AM to 06:00:00 PM (Monday–Thursday) +HVAC [1]
c) 07:00:00 AM to 05:00:00 PM (Monday–Thursday) +HVAC [1]
d) 07:00:00 AM to 04:00:00 PM (Monday–Friday) +HVAC[1]

**Table 3**
Interaction table.

| GVIF | | | | |
|---|---|---|---|---|
| Parameters | San Jose | | Billings | |
| | Cooling load | Heating load | Cooling load | Heating load |
| Occupant schedule | 1.06051 | 1.08365 | 1.06305 | 1.0804 |
| Exterior wall | 1.06685 | 1.07157 | 1.07611 | 1.06639 |
| Exterior wall absorbance | 1.01956 | 1.06453 | 1.01672 | 1.01774 |
| Interior wall | 1.05355 | 1.02867 | 1.06367 | 1.01817 |
| Floor construction | 1.02483 | 1.02171 | 1.02656 | 1.06501 |
| Floor interior finish | 1.01329 | 1.03448 | 1.01868 | 1.02069 |
| Top floor ceiling exterior insulation | 1.02844 | 1.01701 | 1.02287 | 1.02912 |
| Ground floor interior finish | 1.02992 | 1.01689 | 1.02166 | 1.03437 |
| Ceiling insulation | 1.01585 | 1.03065 | 1.02827 | 1.04004 |
| Top floor Batt insulation | 1.0266 | 1.04035 | 1.03293 | 1.02398 |
| Top floor ceiling interior finish | 1.03312 | 1.03237 | 1.0182 | 1.02448 |
| Roof absorbance | 1.01565 | 1.0369 | 1.03975 | 1.01501 |
| Building orientation | 1.02944 | 1.08042 | 1.01562 | 1.02684 |
| Roof absorbance | – | 1.02036 | – | – |
| Roof absorbance | – | 1.02036 | – | – |

e) 08:00:00 AM to 05:00:00 PM (Monday–Friday) +HVAC[2]
f) 08:00:00 AM to 06:00:00 PM (Monday–Thursday) +HVAC[2]
g) 07:00:00 AM to 05:00:00 PM (Monday–Thursday) +HVAC[2]
h) 07:00:00 AM to 04:00:00 PM (Monday–Friday) +HVAC[2]

## 4. Conclusion

The goal of this research was to create multiple regression models for office building in the two climate zones including cold, dry and warm, marine. A total of 150,000-computer simulation was conducted in this study. Stepwise analysis was carried out to correlate the annual energy consumption with the 17 input parameters. Out of 17 building parameters, the stepwise regression algorithm only kept 13 parameters in the cooling load regression model while the heating load regression model retained 14 building parameters and removed the rest. Two design parameters, occupancy schedule and exterior wall construction, were found have the highest influence on both cooling and heating load and the annual energy consumption will be more sensitive to the changes in these two design variables.

The resulting regression models in this study indicated the advantages and potential of this approach to determine the energy performance of the commercial building. What has been offered is a flexible and simple model in order to facilitate the development of regression models to evaluate the building energy consumption and performance. The coefficient of determination $R^2$ varies from 0.95 to 0.98 shows that 95–98% of the variations in annual building consumption use can be explained by changes in few building parameters. The difference between the heating and cooling loads predicted by the regression models were acceptable compared to the result obtained from energy simulations. An accurate, simple and fast way to obtain energy performance of office building is the advantages of the represented regression equation.

## References

[1] Energy Information Administration (EIA), Annual Energy Outlook, in DOE/EIA-03832011 (2011).
[2] S. Asadi, S. Shams Amiri, M. Mottahedi, On the development of multi-linear regression analysis to assess energy consumption in the early stages of building design, Energy Build. 65 (2014) 246–255.
[3] I. Jaffal, C. Ghiaus, Fast method to predict building heating demand based on the design of experiments, Energy Build. 41 (2009) 669–677.
[4] H. Kreiner, A. Passer, H. Wallbaum, A new systemic approach to improve the sustainability performance of office buildings in the early design stage, Energy Build. (2015), http://dx.doi.org/10.1016/j.enbuild.2015.09.040.

[5] V. Corrado, Assessment of building cooling energy need through a quasi-steady state model: simplified correlation for gain-loss mismatch, Energy Build. 39 (5) (2007) 569–579.
[6] C. Ghiaus, Experimental estimation of building energy performance by robust regression, Energy Build. 38 (2006) 582–587.
[7] F. Chlela, Développement d'une méthodologie de conception des bâtiments a basse consommation d'é nergie, Université de la Rochelle, 2008.
[8] E. Gratia, A. De Herde, A simple design tool for the thermal study of an office building, Energy Build. 34 (3) (2002) 279–289.
[9] N. Ghrab-Morcos, CHEOPS: a simplified tool for thermal assessment of Mediterranean residential buildings in hot and cold seasons, Energy Build. 37 (6) (2005) 651–662.
[10] R. Ourghi, A. Al-Anzi, M. Krarti, A simplified analysis method to predict the impact of shape on annual energy use for office buildings, Energy Convers. Manag. 48 (1) (2007) 300–305.
[11] S. Filfli, Optimisation Bâtiment/système pour minimiser les consommations dues à la climatisation, Ecole des Mines de Paris, 2006.
[12] E. Nault, E. Rey, M. Andersen, Towards a predictive model for assessing the passive solar potential of neighborhood designs, in: Presented at the Building Simulation, 2015 (in print), Hyderabad.
[13] A. Mohammadpour, C. Anumba, T. Bulbul, J. Messner, Facilities management interaction with healthcare delivery process, in: Construction Research Congress, 2012, pp. 728–736.
[14] M. Mottahedi, A. Mohammadpour, S.S. Amiri, S. Asadi, D. Riley, A. Shafaghat, Parametric analysis of building elements on building energy use, Jurnal Teknologi 74 (4) (2015).
[15] J. Carlo, R. Lamberts, Development of envelope efficiency labels for commercial buildings: effect of different variables on electricity consumption, Energy Build. 40 (2008) 2002–2008.
[16] N.K. Bansal, A. Bhattacharya, Parametric equations for energy and load estimations for buildings in India, Appl. Thermal Eng. 29 (17) (2009) 3710–3715.
[17] J.C. Lam, K.K.W. Wan, D. Liu, C.L. Tsang, Multiple regression models for energy use in air-conditioned office buildings in different climates, Energy Convers. Manag. 51 (12) (2010) 2692–2697.
[18] J.C. Lam, C.M.S. Hui, A.L.S. Chan, Regression analysis of high-rise fully air-conditioned office buildings, Energy Build. 26 (2) (1997) 189–197.
[19] T. Catalina, J. Virgone, E. Blanco, Development and validation of regression models to predict monthly heating demand for residential buildings, Energy Build. 40 (10) (2008) 1825–1832.
[20] D.H.W. Li, S.L. Wong, K.L. Cheung, Energy performance regression models for office buildings with daylighting controls, J. Power Energy 222 (6) (2008) 557–568.
[21] Python. Available from http://www.python.org/.
[22] American Society of Heating, Refrigerating, and Air-Conditioning Engineers (ASHRAE), 90.1-2007, American Society of Heating, Refrigeration, and Air-Conditioning Engineers, Atlanta, GA, 2007.
[23] T. Hastie, J.H. Friedman, R. Tibshirani, SpringerLink (Online service), The Elements of Statistical Learning: Data Mining, Inference, and Prediction, 2nd ed., Springer, New York, 2009, http://dx.doi.org/10.1007/978-0-387-84858-7.