**905\_Kalfountzos\_Dimitrios\_lab08**

**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*

***What problem the paper addresses & Why is the problem important?***

* Building energy consumption depends on many operational and design parameters; large numbers of simulations are needed to generate data for the multiple regression models.

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 thick-ness) and to identify the coefficients of this function by a regression method [11–14].

***What dataset are using & Which machine learning model are using to address the problem?***

* 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 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.

***Results of the Paper***

Generally, the developing model scan be used to estimate the energy consumption of office buildings in early stages of design.

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 R2 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.