Penerbit Akademia Baru

Journal of Advanced Research in Applied Sciences and Engineering Technology

Journal homepage: www.akademiabaru.com/araset.html
ISSN: 2462-1943

Analysis of Parameters which Affects Prediction of Energy Consumption in Buildings using Partial Least Square (PLS) Approach

Muhammad Irfan[1,2,*], Faizir Ramlie[1]

[1] Razak Faculty of Technology and Informatics, Universiti Teknologi Malaysia, Kuala Lumpur, Malaysia
[2] Department of Electrical Engineering, Faculty of Engineering, Universitas Muhammadiyah Malang, Indonesia

ABSTRACT

The development of energy consumption prediction model is an integral part of the management and improvement of building energy efficiency in order to save the energy and to reduce the environmental impact. There are factors that affect energy in buildings which are heating and cooling. Building materials, ventilation, building direction, and building area are the important factors to determine building energy efficiency. The study aims to find the role of input variable (independent) on the output variable (dependent) in the form of Heating Load (HL) and Cooling Load (CL). The study employs Partial Least Square (PLS) analysis method which is a variant-based Structural Equation Modeling analysis known as SEM-PLS. The result of the study indicates that the estimation of inner model of the direct influence of Orientation on Cooling Load and Heating Load is not significant. It means the size of the Orientation value does not significantly affect the increase/decrease in Cooling Load and Heating Load. While the direct effect of Overall Height, Wall Area, and Surface Area on Cooling Load and Heating Load is significant. It means that the value of Overall Height, Wall Area, and Surface Area has a significant effect on increasing/decreasing Cooling Load and Heating Load. The results of this study are expected to be useful to help building designers, especially related to energy efficiency in the buildings. In addition, the development of this model can be used as an alternative in determining the factors that affect comfort in a building

Keywords:
Building energy; strucural Equation Meodelling, Partial Least Square.

Received: 20 October 2021 **Revised: 27 November 2021** **Accepted: 17 December 2021** **Published: 22 December 2021**

1. Introduction

There have been many studies related to the energy-efficient buildings. All energy-efficient building models developed have the same goal, which is to be able to predict future energy needs and save building energy. Buildings are responsible for 40% of total global energy consumption [1] and and 33% of global greenhouse gas emissions [2]. Moreover, the annual energy demand of buildings is expected to increase by an average of 1.5% between 2012 and 2040 resulting in a total increase of 48% over that time period[3]. If no action is taken to improve the energy efficiency of buildings, energy demand is expected to increase by 50% by 2050 [4]. The basic principle of building

* *Corresponding author.*
E-mail address: irfan20@graduate.utm.my or irfan@umm.ac.id

https://doi.org/10.37934/araset.25.1.6168

energy efficiency is to use less energy for operations including heating, cooling, lighting and other appliances, without affecting the health and comfort of its occupants [5]. Meanwhile, Kreith in his book Handbook of Energy Efficiency and Renewable Energy[6] also states that the use of energy from buildings is dominated by climate influences because the heat obtained from direct conduction from heat sources or air infiltration/exfiltration through the building surface reaches 50-80 % of energy consumed.

Heating, ventilation, and air conditioning have important roles in determining indoor climate and play an important role in most of the energy use in buildings [7]. Energy consumption is used to indicate the amount of energy that must supply the building at any given time. Prediction of building energy consumption then plays an important role in energy management, therefore the designers are expected to be wiser in designing new buildings[8]. The characteristics of a building are an important factor to determine the necessities of heating and cooling loads [9].

Building orientation and layout prove to be very important in reducing building energy consumption in cold and hot climates. The design is limited by the specific characteristics of the planned building, the size, shape, and orientation of the building plot. Meanwhile, the calculation of heating load (HL) and cooling load (CL) in terms of building design efficiency is needed to determine the specifications of heating and cooling equipment needed to maintain comfortable indoor air conditions. This shows that the influence of CL and HL greatly affects the energy requirements of the building [10]. Architects and building designers need information about the building, the space required, the climate, and its uses for cooling and heating. Energy-efficient buildings with special designs such as orientation, insulation, and windows can be well adapted to withstand adverse weather conditions [11].

The description above shows that there are many variables that affect the comfort in the building that affect the energy consumption needed in the building. For this reason, it is necessary to conduct research that reveals the variables that affect the heating load and cooling load in a building. Although several techniques have been applied to solve the problem of CL and HL performance as an estimate of energy consumption, but specifically the variables that affect the performance of Cl and HL with the PLS approach are needed as an alternative model. Therefore, this study focuses on the application in determining the effect of input variables on output. This research was started from previous research [7] who designed a group of buildings with the aim of predicting heating and cooling loads on buildings by taking into account the variables: relative density, surface area, wall area, roof area, overall height, orientation, glazing area, and glass distribution area. . However, the existing input variables for each of the existing loads have not been discussed in detail. For this reason, the purpose of this study was to determine the effect of eight input variables on two output variables, namely heating load (HL) and cooling load (CL) using Partial Least Square (PLS) analysis method.

2. Methodology

The study employs regression analysis method. The regression analysis could draw a conclusion by taking into account the validity factor in which the study uses Partial Least Square (PLS) approach. The Partial Least Square is a powerful analytical method because it can be applied to all data scales, does not require many assumptions, and the sample size does not have to be large. Besides being able to be used as a confirmation of theory, PLS can also be used to build relationships for which there is no theoretical basis or to test propositions [12].

After fulfilling one of the measurement models, the data can be estimated. The parameter estimates obtained using PLS are categorized into three which are weight estimates, path estimates, and related to means and parameter locations (regression constant values) for indicators and latent variables. There are also blocks that separate latent variables and their indicators. The path diagram in PLS is used to determine the relationship and value between the variables. This analysis is also used to determine which variables have the most dominant effect on certain latent variables.

In addition to what has been described above, Goodness of Fit analysis is carried out to measure the influence of several construct variables on certain latent variables. The Goodness of Fit model was measured using dependent latent variable R-square with the same interpretation as the Q-square regression predictive relevance for the structural model. This measures how well the observed values generated by the model and also the parameter estimates are.

2.1 Dataset

The data used in this study is a dataset obtained through UCI Machine Learning. These data were also used by previous researchers to improve the estimation of energy efficiency performance by using different algorithms. Extensive simulations in 768 various residential buildings with 8 input variables and 2 output variables were demonstrated using ICOTECT[7].

Table 1
Mathematical representation of parameters

Representative of Parameters	Parameters Name
X1	Relative Compactness
X2	Surface Area
X3	Wall Area
X4	Roof Area
X5	Overall Height
X6	Orientation
X7	Glazing Area
X8	Glazing Area Distribution
Y1	Heating Load
Y2	Cooling Load

Variables Information :

- Relative Compactness
- Surface Area - m^2
- Wall Area - m^2
- Roof Area - m^2
- Overall Height - m
- Orientation - 2:North, 3:East, 4:South, 5:West
- Glazing Area - 0%, 10%, 25%, 40% (of floor area)
- Glazing Area Distribution (Variance) - 1:Uniform, 2:North, 3:East, 4:South, 5:West
- Heating Load - kWh/m^2
- Cooling Load - kWh/m^2

To facilitate the analysis using the SEM-PLS method, it is necessary to determine the identification of variables. The variables used consist of exogenous variables (Independent) and endogenous variables (Dependent). In this case, the categories of exogenous variables are Relative Compactness, Surface Area, Wall Area, Roof Area, Overall Height, Orientation, Glazing Area, and Glazing Area Distribution. While the categories of endogenous variables are Heating Load, and Cooling Load.

2.2 Analysis Stage

The data is analyzed by using Smart PLS software. According to Ghozali [12] PLS is an alternative approach that shifts from a covariance-based SEM approach to a variance-based approach. Covariance-based SEM generally tests causality/theory, while PLS is more of a predictive model. Because PLS does not assume a certain distribution for parameter estimation, then parametric techniques to test the significance of parameters are not needed [13].

The measurement model or outer model with reflexive indicators is evaluated using convergent and discriminant validation of the indicators and composite reliability for indicator blocks. While the outer model with formative indicators is evaluated based on its substantive content, namely by comparing the relative weight and seeing the significance of the weight size. The structural model or inner model is evaluated by looking at the R2 value for the latent response construct and also looking at the magnitude of the structural path coefficient. The stability of this estimate was evaluated using the t-test statistic obtained from the bootstrapping procedure [12].

Evaluation of the model using R-square (R^2) for the dependent construct The R-square value reflects the predictive power of the overall model [14,15] with an R-square value limit greater than 0.10 or greater than 10 percent (or goodness-fit of the model. Goodness of fit in PLS can be seen from the value of Q^2 The value of Q^2 has the same meaning as the coefficient of determination (R-square / R^2) in regression analysis The higher R^2, the model can be said to be more fit with the data The Q-Square value is greater than 0 (zero) indicates that the model has predictive relevance, while the Q-Square value is less than 0 (zero) indicating that the model lacks predictive relevance [12]. The Q^2 value uses the following equation as follows :

$$Q^2 = 1 - (1 - R_{21})(1 - R_{22})(1 - R_{23}) \dots .(1 - R_{2n}) \tag{1}$$

where R_{21} , R_{22} ... R_{2n} is the endogenous variable R-square

To calculate reliability, use composite (construct) reliability with a cut off value of at least 0.7. However, for exploratory research, moderate reliability of 0.5 – 0.6 is sufficient to justify the results of the study[16].

3. Results
3.1 Path Analysis Result

The software used in processing the data in this study is smart PLS version 3, all variables (Relative Compactness, Surface Area, Wall Area, Roof Area, Overall Height, Orientation, Glazing Area, Glazing Area Distribution, Heating Load, dan Cooling Load) used are formative indicator types (arrow direction from the construct variable to the latent variable). This is because the indicators of these constructs are mutually exclusive. The path diagram is illustrated in the following Figure 1.

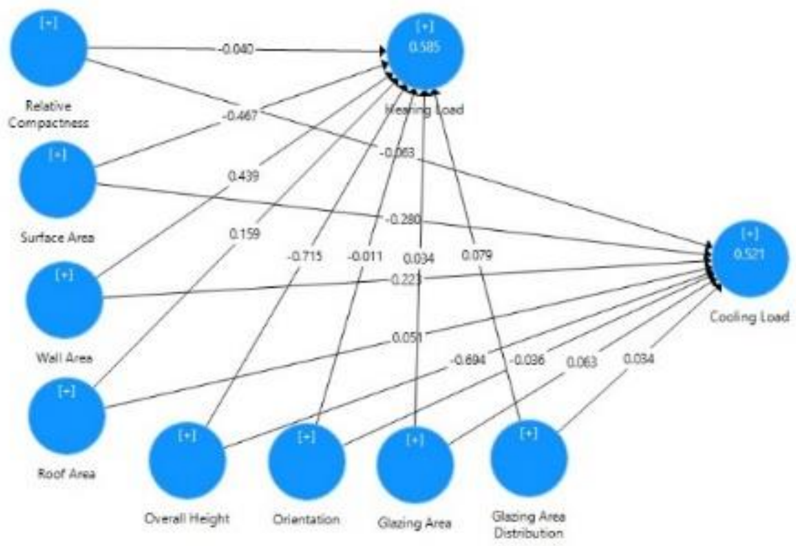

Fig. 1. Structural equation path diagram with partial least square approach

Figure 1 above illustrates the path coefficient values between the input variables and the output variables and the value of R-square (R^2) Heating Load = 0.585; R^2 Cooling Load = 0.521.

3.2 Model Measurement

The output of the measurement model on the variables of Relative Compactness, Surface Area, Wall Area, Roof Area, Overall Height, Orientation, Glass Area, Glass Area Distribution, Heating Load, and Cooling Load each of 1 indicator item, as presented in the following Table 2.

Table 2
Output Model Measurement of variable

Variables	Original Sample (O)
Cooling Load -> Cooling Load	1,000
Glazing Area -> Glazing Area_	1,000
Glazing Area Distribution -> Glazing Area Distribution	1,000
Hearing Load -> Heating Load	1,000
Orientation -> Orientation	1,000
Overall Height -> Overall Height	1,000
Roof Area -> Roof Area	1,000
Surface Area -> Surface Area	1,000
Wall Area -> Wall Area	1,000
Relative Compactness -> Relative Compactness	1,000

Original Sample (O) = loading factor

As can be seen in Table 2, the indicators for Relative Compactness, Surface Area, Wall Area, Roof Area, Overall Height, Orientation, Glazing Area, Glazing Area Distribution, Heating Load, and Cooling Load each has value of 1.0. Because it is a formative indicator, all indicators of these variables have met the criteria with a loading factor coefficient value of >0 7. Therefore all variable items can be used for further testing.

3.3 Model Evaluation

Based on data processing with PLS (Figure 1), the resulting coefficient of determination (R-square) is as shown in Table 3 below.

Tabel 3
Nilai R-Square

Variables	R Square
Cooling Load	0,521
Heating Load	0,585

From Table 3 above, it can be seen that the value of Q^2 is as follows :

$$Q^2 value = 1 - (1 - R_{21})(1 - R_{22})(1 - R_{23}) \ldots .(1 - R_{2n})$$
$$Q^2 value = 1 - (1 - 0.521)(1 - 0.585)$$
$$Q^2 value = 1 - 0.1988 = 0.8012$$

In this research model, the Q-square value generated in the overall model equation is 80.12%, with the number of 768 samples. It means that the structural model has very good predictive relevance and it is suitable for predictions.

3.4 Direct Effect of Exogenous Variable on Endogenous Variable

The direct effect of exogenous variables on endogenous variables can be seen in Table 4 below:

Tabel 4
Direct Effect

Variables	Original Sample (O)	P Values
Glazing Area Distribution -> Cooling Load	0,034	0,163
Glazing Area Distribution -> Heating Load	0,079	0,001
Glazing Area_ -> Cooling Load	0,063	0,009
Glazing Area_ -> Heating Load	0,034	0,159
Orientation -> Cooling Load	-0,036	0,177
Orientation -> Heating Load	-0,011	0,648
Overall Height -> Cooling Load	-0,694	0,000
Overall Height -> Heating Load	-0,715	0,000
Relative Compactness -> Cooling Load	-0,063	0,046
Relative Compactness -> Heating Load	-0,040	0,185
Roof Area -> Cooling Load	0,051	0,286
Roof Area -> Heating Load	0,159	0,003
Surface Area -> Cooling Load	-0,280	0,002
Surface Area -> Heating Load	-0,467	0,000
Wall Area -> Cooling Load	0,223	0,035
Wall Area -> Heating Load	0,439	0,000

Original Sample (O)= koefisien path

The estimation results of the inner model for the direct effect of Glazing Area Distribution on Cooling Load show a path coefficient value of 0.034 with a p-value of 0.163 which is greater than alpha 0.05 (error level =5%). This shows that the direct effect of Glazing Area Distribution on Cooling Load is not significant. This means that the size of the Glazing Area Distribution has no significant effect on increasing or decreasing (good/bad) Cooling Load. While the estimation results of the inner model for the direct effect of Glazing Area Distribution on Heating Load show a path coefficient value of 0.079 with a p-value of 0.001 which is smaller than alpha 0.05 (error level =5%). It shows that the direct influence of Glazing Area Distribution on Heating Load is significant. The resulting effect is positive, which means the higher (good) the Glazing Area Distribution, then the higher the Hearing Load will be. And vice versa, the lower (less good) Glazing Area Distribution, then the Heating Load will be lower. Based on table 4 to get the results of the estimation of other variables, in the same way, the influence of the input variable on the output variable will be obtained.

4. Conclusions

Based on the analysis and testing that has been done, from the 8 input variables, it can be concluded that:

1. The estimation results of the inner model for the direct effect of Orientation on Cooling Load and Heating Load are insignificant, meaning that the size of the Orientation value does not significantly affect the increase/decrease in Cooling Load and Heating Load.
2. The estimation results of the inner model for the direct effect of Overall Height, Wall Area, and Surface Area on Cooling Load, and Heating Load are significant, meaning that the size of the Overall Height, Wall Area, and Surface Area value has a significant effect on increasing/decreasing Cooling Load, and Heating Load.

Acknowledgement

Special thanks to the Muhammadiyah University of Malang for funding and facilitating this research. We also thank Dr. Faizir Ramlie as a supervisor in the Department of Engineering, under the Faculty of Technology and Informatics Razak, Universiti Teknologi Malaysia who has supported this work with the Collaborative Research Grant Project (Project Grant No. Q.K130000.2456.08G38) under the Collaborative Research Grant Program (Program Grant No. Q.K130000.2456.08G27).

References

[1] Sepehr, Mohammad, Reza Eghtedaei, Ali Toolabimoghadam, Younes Noorollahi, and Mohammad Mohammadi. "Modeling the electrical energy consumption profile for residential buildings in Iran." *Sustainable cities and society* 41 (2018): 481-489.

[2] Zhan, Jinyan, Wei Liu, Feng Wu, Zhihui Li, and Chao Wang. "Life cycle energy consumption and greenhouse gas emissions of urban residential buildings in Guangzhou city." *Journal of cleaner production* 194 (2018): 318-326.

[3] EIA, *International Energy Outlook 2016*. Washington DC: U.S. Energy Information Administration, 2016.

[4] International Energy Agency, *Transition to Sustainable Buildings*. 2013.

[5] Duarte, Grasiele Regina, Leonardo Goliatt da Fonseca, Priscila Vanessa Zabala Capriles Goliatt, and Afonso Celso de Castro Lemonge. "Uma comparação de técnicas de aprendizado de máquina para a previsão de cargas energéticas em edifícios." *Ambiente Construído* 17 (2017): 103-115.

[6] Goswami, D. Yogi, and Frank Kreith, eds. *Handbook of energy efficiency and renewable energy*. Crc Press, 2007.

[7] Tsanas, Athanasios, and Angeliki Xifara. "Accurate quantitative estimation of energy performance of residential buildings using statistical machine learning tools." *Energy and Buildings* 49 (2012): 560-567.

[8] Zhao, Hai Xiang, and Frédéric Magoulès. "Parallel support vector machines applied to the prediction of multiple buildings energy consumption." *Journal of Algorithms & Computational Technology* 4, no. 2 (2010): 231-249.

[9] Nebot, Àngela, and Francisco Mugica. "Energy performance forecasting of residential buildings using fuzzy approaches." *Applied Sciences* 10, no. 2 (2020): 720.

[10] T. H. E. E. Parliament, T. H. E. Council, O. F. The, and E. Union, “Directive 2002/65/EC of the European Parliament and of the Council,” *Fundam. Texts Eur. Priv. Law*, pp. 65–71, 2020, doi: 10.5040/9781782258674.0021.

[11] Yu, Zhun, Fariborz Haghighat, Benjamin CM Fung, and Hiroshi Yoshino. "A decision tree method for building energy demand modeling." *Energy and Buildings* 42, no. 10 (2010): 1637-1646.

[12] Ghozali, *Structural Equation Modeling Metode Alternatif Dengan Partial Least Square (PLS)*. Semarang: Universitas Diponegoro Semarang, 2006.
[13] Chin, Wynne W., Barbara L. Marcolin, and Peter R. Newsted. "A partial least squares latent variable modeling approach for measuring interaction effects: Results from a Monte Carlo simulation study and an electronic-mail emotion/adoption study." *Information systems research* 14, no. 2 (2003): 189-217.
[14] Falk, R. Frank, and Nancy B. Miller. *A primer for soft modeling*. University of Akron Press, 1992.
[15] Pirouz, Dante M. "An overview of partial least squares." *Available at SSRN 1631359* (2006).
[16] Augusty, Ferdinand. "Structural Equation Modeling dalam Penelitian Manajemen: Aplikasi Model-Model Rumit dalam Penelitian Untuk Tesis Magister & Disertasi Doktor Edisi 2." *Universitas Diponegoro, Semarang* (2002).