



**HAL**  
open science

# The impact of indoor climate and occupancy rates on electricity use in campus buildings Analyse the impact level through the linear regression model

Keovathana Run, Franck Cevaër, Jean-François Dube

## ► To cite this version:

Keovathana Run, Franck Cevaër, Jean-François Dube. The impact of indoor climate and occupancy rates on electricity use in campus buildings Analyse the impact level through the linear regression model. 20th International Conference on Sustainable Energy Technologies - SET2023, Aug 2023, Nottingham, United Kingdom. hal-04793706

**HAL Id: hal-04793706**

**<https://hal.science/hal-04793706v1>**

Submitted on 20 Nov 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

---

# #105: The impact of indoor climate and occupancy rates on electricity use in campus buildings

Analyse the impact level through the linear regression model

---

Keovathana RUN<sup>1</sup>, Franck CEVAËR<sup>2</sup>, Jean-François DUBE<sup>3</sup>

1 LMGC, University of Montpellier, UMR CNRS 5508, cc048, Place E. Bataillon, 34095 Montpellier cedex 5, France, keovathana.run@umontpellier.fr

2 LMGC, University of Montpellier, UMR CNRS 5508, cc048, Place E. Bataillon, 34095 Montpellier cedex 5, France, franck.cevaer@umontpellier.fr

3 LMGC, University of Montpellier, UMR CNRS 5508, cc048, Place E. Bataillon, 34095 Montpellier cedex 5, France, jean-francois.dube@umontpellier.fr

*Abstract: Energy efficiency in buildings can be achieved by many approaches including a high-performance HVAC system, passive design strategies, and accurate energy consumption prediction. Energy evaluation is crucial in existing buildings to improve system control by detecting any unusual pattern of consumption as well as identifying the most influential variables in energy usage for a better target in retrofit projects. This paper aims to identify the significant contributors to energy consumption in both heating and non-heating periods of campus buildings situated in the south of France. The linear regression model is adopted to investigate the relationship between the dependent variable (response) and independent variables (predictors). During the heating period, the analysis focuses on five potential variables, namely, time index (daytime/night-time), day index (weekday/weekend), building size, indoor and outdoor climate, and their confounding. When electric heaters (reversible air-conditioners) are likely to be involved, results have shown that the model with the interactive effect variables achieves a coefficient of determination of 68%. That model consists of CO<sub>2</sub>, time index, day index, building size, indoor humidity, and indoor temperature. During the non-heating period, the occupancy rate is the main target for its impact level on electricity use. The regression model has found that the occupancy rate and day index alone can explain 50% of the consumption variation. Together, this study points out the importance of the occupant's precision to better the predictive model in future studies. This includes a better assumption of occupied space and schedule and a better understanding of the scale of the occupant.*

*Keywords: Electricity consumption, Campus buildings, Linear regression model, Occupancy rate, Indoor climate*

## 1. INTRODUCTION

Studies to date have proposed many approaches and techniques to forecast energy consumption at the design stage or analyses of retrofit options to achieve higher energy efficiency in buildings. Varying from literature reviews, there is no specific structure available on building energy (Fumo, 2014). Data-driven is one of the most frequently discussed due to its ease of use and capacity for handling massive data sets (Afroz et al., 2018). A recent study (Foucquier et al., 2013) categorizes this approach into three categories as follows: (1) Physical models: demand in-depth understanding of how to construct overall information for usage in widely used simulation software. (2) Statistical methods: to create the model, historical data is needed. This approach entails Multiple Linear Regression, Conditional Demand Analysis, Genetic Algorithm, Artificial Neural Network, and Support Vector Machine. (3) Hybrid mode seeks to close the gap between physical and statistical models.

Higher education buildings are designed for various activities which accommodate office, classroom, e-learning, laboratory, workshop, residential and parking (Pérez-Lombard, Ortiz and Pout, 2008). Due to the low building densities on campus and the shorter operating hours during the peak energy demand seasons, campus buildings offer a better possibility for energy reduction (Chung and Rhee, 2014). However, energy consumption is encouraged to decrease in response to the current trend of becoming a sustainable development in education by positively implementing low carbon activities on campus (Luo, Han and Zhou, 2017).

Although there are many different energy consumers in a building, HVAC systems, lighting systems, IT equipment, and server equipment are frequently the biggest consumers (Agarwal, Weng and Gupta, 2009). However, the primary flaw in energy forecasting, particularly in structures with air conditioning systems, is brought on by the outside weather, occupant changes during the day, and the internal loads put in place in the building (Neto and Fiorelli, 2008). In many current simulation tools for simulating building performance, occupant behaviours are viewed as static, deterministic schedules or settings (Hong et al., 2018). The randomization of occupancy has a statistical impact on how efficiently a building uses energy (Carlucci et al., 2016).

The objective of this paper is to identify the most influential variables on electricity consumption during heating and non-heating period. The analyse is done through predicting technique using one of the data-driven approach. As the case study is existing buildings with historical data available both on site and on outdoor climate, multiple linear regression model is adopted for its simple structure, quick and efficient predictive ability.

## 2. METHOD

### 2.1. Multiple linear regression

The technique used in this work is a multivariate regression analysis, which synchronically takes into account changes in the independent and dependent variables (Uyanık and Güler, 2013).

*Equation 1: The multiple linear regression model.*

$$Y_i = \beta_0 + x_{i,1}\beta_1 + x_{i,2}\beta_2 + \dots + x_{i,p}\beta_p + e_i$$

Where:

- $Y$  = the response (dependent variable)
- $x$  = the predictors (independent variables)
- $\beta$  = the unknown regression coefficients
- $e$  = the error to account for the discrepancy between predicted and the observed data
- $i = 1, \dots, n$
- $n$  = the sample sizes

### 2.2. Model selection

This study employs a forward regression, which starts with a model with no predictors, to choose its predictors (the intercept-only model). Next, the variables are added to the model one at a time until none more can improve it by a particular standard. The variable that significantly improves the model is introduced at each step. A variable stay in the model once it is added. The analysis is done using Rstudio version 4.0.3. (2020-10-10) (Verzani, 2011). The train function [caret package] computes the forward regression; the tuning parameter  $nvmax$  specifies the maximum number of predictors to incorporate in the model. Regression analysis uses the Mean absolute error (MAE), Root means squared error (RMSE) metrics to measure the model's prediction error. The better the model, the lower the MAE, and RMSE. The coefficient of determination ( $R^2$ ) denotes the relationship between the values of the desired outcomes and those that the model predicts. The model is better the greater the  $R^2$ .

### 2.3. Model evaluation metrics

The performance of the models is assessed using 10-fold cross-validation after the best models from the combination of chosen parameters have been developed. The evaluation metrics can be calculated as follows (Botchkarev, 2019):

Equation 2: Mean absolute error.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i|$$

Equation 3: Mean square error.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$$

Equation 4: Root mean square error.

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$$

Equation 5: Coefficient of determination.

$$R^2 = \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y}_i)^2}$$

Where:

- $\hat{y}$  = predicted value of  $y$  measured value at the  $i$ th moment
- $\bar{y}$  = mean value of  $y$  measured value at the  $i$ th moment
- $N$  = the number of predictions

### 3. CASE STUDY: IUT DE NÎMES SCHOOL BUILDINGS

IUT de Nîmes campus situates in southern France, at 43°49'N longitude and 4°19'E latitude under Mediterranean weather. Three of five buildings on the campus were studied in this paper. These buildings, which were constructed in 1969, share the same building elements and layout with two sections: teaching and technical. The teaching section has two-storey and the technical has only one. *Figure 1* displays a floor plan for one of the buildings.



Figure 1 Teaching building floor plan, ground floor level

Each building is naturally ventilated in summer and heated in winter. The primary heating source is a substation that supplies the building's radiators. The secondary heating source is aérothermal which is regulated individually but connected to the primary source and therefore operates permanently. These sources are handled by the urban heating network which operates specifically during the heating period from November to April. The electricity system covers the rest of the appliances in the buildings including lighting, the electrical distributor, electrical radiators, reversible air-conditioners, and IT networks, etc. In winter, the user's comfort relies greatly on the primary and secondary heaters. However, the third heating source (reversible air conditioners) is substituted when a disruption occurs in the main heating networks. Consequently, the electricity use intensity could be linked to outdoor and indoor climates.

IUT de Nîmes is equipped with an electric meter that measures the total current used in each building. For a thesis project, two types of climate sensors namely Elsys and Class'Air have been installed on site since 2020. These sensors measure the indoor CO<sub>2</sub> level, indoor relative humidity, and indoor air temperature. In this paper, only Elsys sensor is used as it transmits measurements every 15 minutes, more frequently than the other one.

The three studied buildings are Civil engineering building (GC), Electrical engineering building (GEII), and Material engineering building (GMP) with a total net surface area of 4762m<sup>2</sup>, 3627m<sup>2</sup> and 6357m<sup>2</sup>, respectively. In 2021, GEII was renovated that include the implementation of the building ventilation system, an automated lighting system, and upgrades to the building envelope. These are related to the installation of insulation, high-performance double-glazed windows, and overhangs. In 2022, the same renovation is done on GC from May to November. Long-term monitoring presents common measurement challenges that can lead to potential data gaps. In this study, such challenges emerged, notably in 2020, where technical issues resulted in significant data gaps. The outbreak of COVID-19 further compounded the problem, as it hindered the timely resolution of these technical issues, leaving the data unattended until the end of 2021.

Nonetheless, the data collection process was fully restored in 2022, enabling a comprehensive understanding of the energy consumption evolution throughout the entire year, as presented in *Table 1*.

*Table 1: Monthly electricity consumption in 2022 of each referenced building*

Building	Electricity consumption per month (MWh)											
	Jan	Feb	Mar	Apr	May	Jun	July	Aug	Sep	Oct	Nov	Dec
GC	12.5	11.5	10.8	6.3	3.4	1.9	2.2	2.3	3.2	3.8	5.9	12.2
GEII	12.0	5.0	5.2	2.9	4.1	4.3	3.0	2.5	4.3	4.1	3.1	8.9
GMP	19.7	16.1	19.2	12.4	15.3	17.2	9.9	9.0	13.2	14.8	9.1	17.6

To ensure the accuracy of the analysis, only a period with sufficient data collection was chosen. Therefore, the main subject of this study is on the heating period from November 2021 to April 2022 and on the non-heating period from August to October 2019 and May to October 2022.

## 4. RESULTS AND DISCUSSION

### 4.1. Heating period

The outdoor climate is taken from climate data of Météo France – Nîmes Courbessac. *Table 2* gives the statistical values of the outdoor climate and indoor climate measurements of each building. Although the windows remain closed most of the time to conserve the heat, the level of CO<sub>2</sub> is in the acceptable range on average except in GMP when the maximum reaches 1780 ppm occasionally. IUT de Nîmes's main heating source is a hot water heating system that flows through the pipes of each radiator. To obtain a constant comfortable temperature, the system follows a water law regulation that considers the outdoor temperature by adjusting the slope and parallel graph. Therefore, when the heating system is in a good condition, the indoor temperature is stable at approximately 20°C regardless of the outdoor temperature.

However, failures in the heating system may occur from time to time on the campus, requiring users to activate the heating function of the air-conditioner if necessary. In this case, the electricity consumption may be increased by the indoor temperature and outdoor temperature. Nine potential variables that influence the change in electricity consumption are pre-selected. They are as follows: (1) Outdoor temperature (T<sub>ext</sub>), (2) Outdoor humidity (HR<sub>ext</sub>), (3) Solar radiation (SR), (4) Indoor temperature (T<sub>in</sub>), (5) Indoor humidity (HR<sub>in</sub>), (6) Indoor CO<sub>2</sub> level (CO<sub>2</sub>), (7) Building size, (8) Time index (daytime, night-time) and (9) Day index (weekday, weekend, and vacation). The time index in this context is the occupied hour and unoccupied hour that are referred to as 'daytime' and 'night-time', respectively.

*Table 2: Principal statistics of environment data of referenced building over a heating period*

Building	Variables	Minimum	1 <sup>st</sup> quartile	Mean	3 <sup>rd</sup> quartile	Maximum
Outdoor climate	T <sub>ext</sub> (°C)	-3.5	5.7	9	12.1	27.5
	HR <sub>ext</sub> (%)	18	51	64.9	80	97
	CO <sub>2</sub> (ppm)	379	416	472	493	1084
GC	HR <sub>in</sub> (%)	22	32	38	44	56
	T <sub>in</sub> (°C)	14.6	18.6	19.5	20.3	22.8
GEII	CO <sub>2</sub> (ppm)	369	417	470	478	1043
	HR <sub>in</sub> (%)	18	29.3	38	44	69

	T <sub>in</sub> (°C)	13.5	18.1	19.8	21.9	23.3
	CO <sub>2</sub> (ppm)	380	422.9	549	602	1780
GMP	HR <sub>in</sub> (%)	25	33.5	38.9	44	57
	T <sub>in</sub> (°C)	17.1	20.8	21.2	21.8	23.4

The main heating system on this campus is scheduled to deliver hot water at approximately 40°C during working hours on weekdays. At night-time and weekends, the hot water temperature is decreased to 20°C. In general, these buildings are the most occupied from 8 am to 12:30 am and from 1:30 pm to 5:30 pm daily and from Monday to Friday weekly. This leads to a similar trend of energy use during the occupied and non-occupied periods. Therefore, the time and day index is dominant in the indoor climate which makes a study on their interactive effect necessary. In developing a model, the time index and day index becomes constants, and they are used to couple with indoor CO<sub>2</sub> level, indoor humidity, and indoor temperature to associate the change in electricity consumption.

The model can be written as:

$$\text{Equation 6: Electricity consumption. } P = \beta_0 + \beta_1 \text{size} + \beta_3 T_{\text{ext}} + \beta_4 \text{HR}_{\text{ext}} + \beta_5 \text{IT} + \beta_3 \text{time} * \text{day}(\text{CO}_2 + \text{HR}_{\text{in}} + T_{\text{in}})$$

To avoid biased in the model, the time and day index is transformed into a binary based on 0 as nighttime, 1 as daytime in the time index and 0 as the weekend, 1 as the weekday on the day index. Using 10-fold cross-validation in the forward regression, the first five models are presented in Table 3 where “\*” represents a confounding effect.

Table 3: Developed model equations and performance

Model	Equation	RMSE (kWh)	MAE (kWh)	R <sup>2</sup> (%)
Model 1	0.14 + CO <sub>2</sub> *day*time	10.51	7.53	52
Model 2	0.07+0.13size +0.93 CO <sub>2</sub> *day*time	9.26	5.95	62
Model 3	0.05+0.14size +0.09day*time+0.67CO <sub>2</sub> *day*time	8.71	5.41	67
Model 4	0.05+0.14size+0.14day*time+0.69day*time*CO <sub>2</sub> - 0.13day*time*HR <sub>in</sub>	8.63	5.38	67
Model 5	0.09+0.15size-0.07Tin+0.15day*time+0.71 day*time*CO <sub>2</sub> - 0.15day*time*HR <sub>in</sub>	8.55	5.51	68

Models with SR, HR<sub>ext</sub>, and T<sub>ext</sub> variables have little to no improvement in the real values fitting, so they can be ignored and not present in this paper. Figure 2 compares the weight of the coefficient of the model’s inputs. Apart from the CO<sub>2</sub>\*day\*time, the rest of the variables have roughly the same coefficient weight which is less than the absolute value of 0.25. The most influential parameter is CO<sub>2</sub> that is interacted with the time and day index. The level of CO<sub>2</sub> can testify to the concentration of users in the space; Hence, it is one of the indicators of the user’s presence. But the level of CO<sub>2</sub> can be impacted by many factors such as room size, the state of the window, the mechanical ventilation system, etc. Hence, the interaction effect can strengthen its ability in predicting as the time and day index can contribute to determining the presence of the occupant.

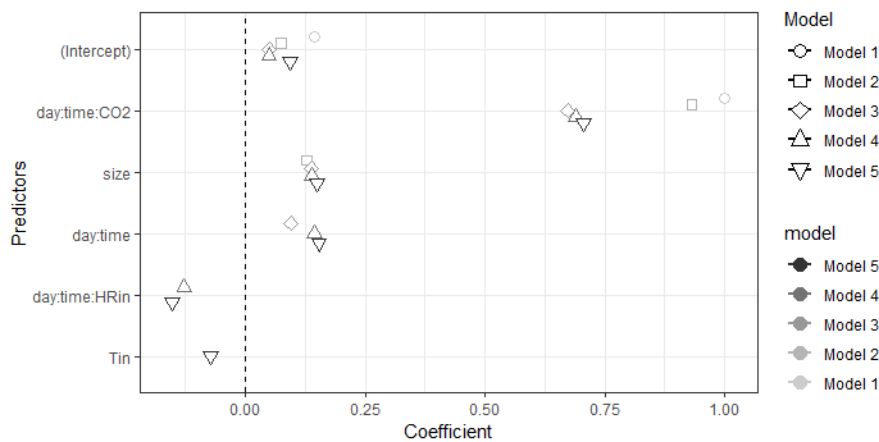


Figure 2 Comparison between coefficients of predictor variables between Model 1, Model 2, Model 3, Model 4, and Model 5. The dashed line defines the negative and positive correlation of the variables when they are on the left and right, respectively. The further the data point from the dashed line, the more influential the variable on the model

Figure 3 illustrates the evolution of real electricity consumption in terms of hours of the day compared to the predicted value of the five models. Globally, the measured consumption is higher than the predicted ones during weekdays but lower during the weekend. The highest peak energy consumption happens at 11 am on a weekday and mostly stable throughout the whole day over the weekend. The models can follow the trend of consumption fairly. However, they struggle to reach the peak consumption in GC and GMP buildings.

GEII's highest peak, on the other hand, can be surpassed by Model 1 and Model 5. With  $R^2$  of 68%, Model 5 has the closest outcomes to real values both when they are stable and when they change for all buildings. Model 1, on the other hand, performs poorly with  $R^2$  of 52% as it noticeably overestimates and underestimates the measured values as to why it can surpass the peak of GEII. Between Model 4 and Model 5 the difference in outcomes is very minimal almost identical which hints that adding Tin to the model is unnecessary.

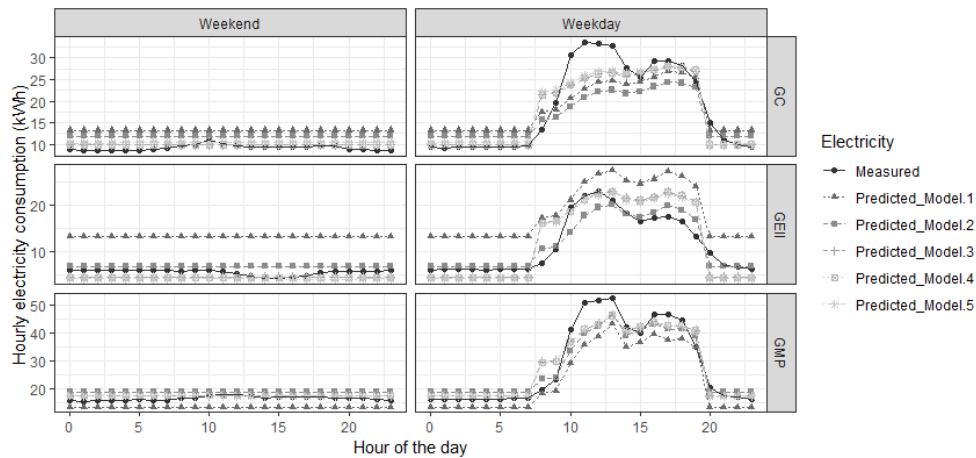


Figure 3 Comparison of the measured values against the predicted values of the heating period, grouped by hour

As the day index is transformed in 0 when it is a weekend, the multiplication with this variable is also 0 which results in constant predicted values on the weekend.

The ideal design of the HVAC system should take into account key outdoor climate factors and main indoor environmental characteristics as they are strongly linked to energy use in buildings (Yang, Clements-Croome and Marson, 2017). However, when the electricity supplies other demands but not the HVAC system, the mentioned key contributors become irrelevant. The case study, IUT de Nîmes, has two separate consumption systems, heating, and electricity.

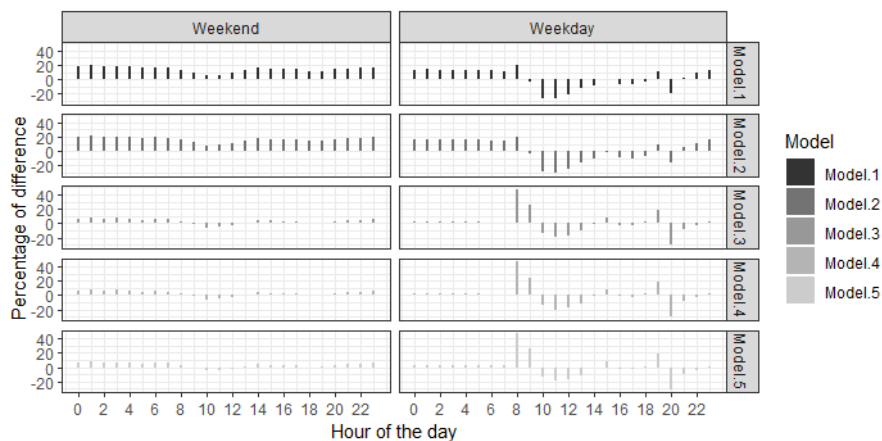


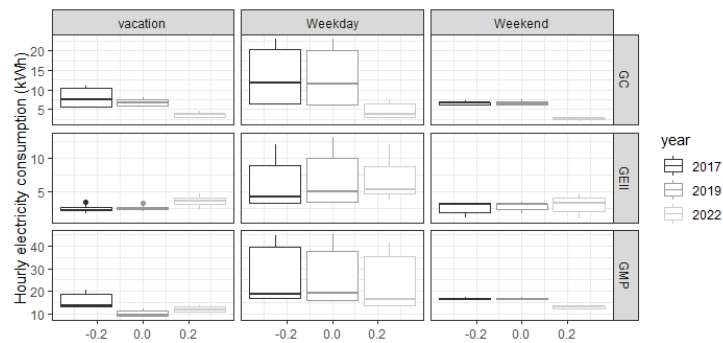
Figure 4 Percentage of difference between measured values and predicted values of each model in the heating period

However, the results show that a relationship between indoor climate and electricity usage has been found, which is likely to be true when the main heating system is disrupted or inefficient. That means the user turns on the electric heater when the radiators are not working due to the disruption in the system or when the heat diffused from the radiator is not to their satisfactory temperature. In this case, the indoor climate especially the level of CO<sub>2</sub> somehow can explain the variation in energy consumption. Based on Figure 4, the biggest error of approximately 40% happens at 8 am and not at 11 am when the electricity consumes the highest. Therefore, these models are manipulated by the sudden increase in CO<sub>2</sub> and the

status change from non-occupied to occupied. *Model 1* and *Model 2* maintain an error percentage of approximately 20% throughout the whole week. *Model 3*, *Model 4*, and *Model 5* have the smallest difference around 5% on the weekend when the measurements are stable, but they also have the biggest difference of 40% maximum when the consumption fluctuates. Overall, these models overestimate electricity consumption on the weekend and underestimate electricity consumption on the weekday.

## 4.2. Non-Heating period

The non-heating period in Mediterranean weather is mostly from either April or May to October which differs from one year to another. *Figure 5* shows a boxplot of hourly electricity consumption non-heating period in 2017, 2019, and 2022. During this period, there are three vacations namely spring break – two weeks in May, summer break – four weeks in August, and Toussaint - one week in May. When a building is vacant for a long period, the number of plugged-in equipment is likely to decrease which leads to a lesser or at least equivalent to the current used over the weekend. The bigger the boxplot, the more significantly different between the maximum and the minimum values which is the case for each building on Weekday.



*Figure 5* Boxplot of electricity consumption during the non-heating period (May to October) of 2017, 2019, and 2022 showing median, inter-quartile range, upper and lower quartile

From May to November 2022, the GC building was undergoing a renovation hence the number of occupied rooms has reduced greatly. Therefore, the electricity consumption of GC in 2022 is four times lower than typical usage in 2017 and 2019. GEII on the other hand was undergoing a renovation from May to November 2021. The implementation of a mechanical ventilation system causes the consumption in 2022 to be marginally higher than it was before the renovation. As the day index and occupancy rate statistically influence the fluctuation of electricity consumption, the explanatory model is employed to unbiasedly study their impact level. During the typical unoccupied period, the electricity consumption of GC and GMP is roughly twice it is of GEII. This is due to the server room operating 24 hours a day, seven days a week in these two buildings.

The occupancy rate is the ratio of occupied rooms to the total amount of available space, as shown in *Table 4*. With the use of the campus’s academic schedule, the occupied room can be identified over the selected academic year to calculate the occupancy rate each hour. However, the academic schedule is only available from August to October 2019 and May to October 2022. Therefore, the analysis is done within this timeframe.

*Table 4: Occupancy profile*

Building	Number of rooms				Occupancy rate		Occupancy profile		
	Classroom	Atelier	Offices	Others	Total	Minimum	Maximum	Schedule	Average number of occupants
GC	19	26	7	7	59	0.13	0.63	8h00 – 18h00, Monday - Friday	Classroom: 30
GEII	14	20	7	4	45	0.18	0.6		Atelier: 15
GMP	27	30	16	11	84	0.2	0.6		Offices: 3

The difference between  $R^2$  in *Model 1* and *Model 2* shows that the day index contributes to making a model better. The occupancy rate is important, but alone, its ability to predict values is rather weak.

*Table 5: Developed model performance*

Model	Input	RMSE (kWh)	MAE (kWh)	$R^2$ (%)
Model 1	Occupancy rate	9.01	6.631	30
Model 2	Occupancy rate, day index	7.58	5.354	50

In summer, the model focuses on the impact of occupancy rate and day index potentially the true indicators of electricity usage. *Figure 6* shows a comparison between measured values and the predicted values in terms of an hour of the day



in each building. The predicted values show a non-uniform electricity trend throughout the studied period in hours of the day. The periods in this analysis are two months in 2019 and six months in 2022 which explains the difference between average hourly measured values between 2019 and 2022. Firstly, in the GC building, the consumption is underestimated in 2019 during typical usage but overestimated in 2022 when the building was undergoing renovation. The entire non-heating period of 2022 is exceptional because the occupancy rate does not reflect the real occupied class. For instance, energy is consumed even though the rooms are programmed to be vacant due to construction work. As for the GEII building, by 2022, the renovation should have better energy efficiency, especially during winter. But in summer, it is understandable that the electricity consumption increases because mechanical ventilation is implemented, which adds up the consumption in return for the user's comfort. Moreover, global GEII has always consumed the least energy on the whole campus for two main reasons. One, whilst the other buildings use of mechanical machines in the atelier that consume a huge amount of energy, most machines in GEII's atelier are electrical which consumes far less. Second, the other two buildings have a constantly running server room that operates every day throughout the entire year. As much as having activities like the other buildings, the amount of electricity used by GEII is less dramatic which leads to the lowest least correlation between occupancy rate and electricity consumption.

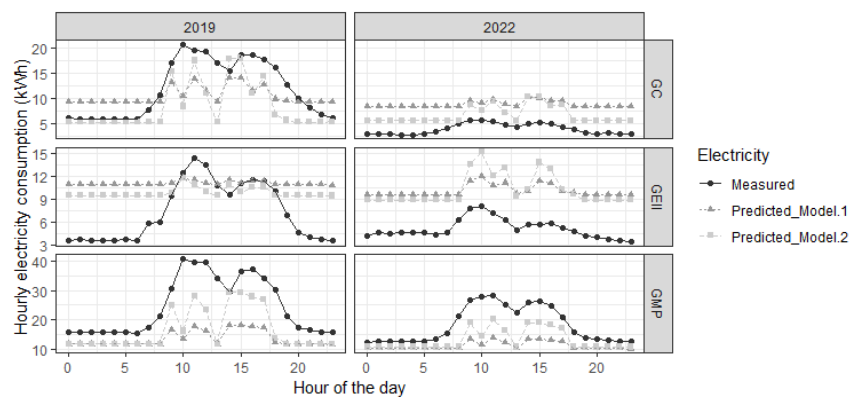


Figure 6 Comparison of the measured values against the predicted values of the non-heating period (August to October 2019 and May to October 2022), grouped by hour

Moreover, during vacations and weekends, the occupancy rate is assumed to be 0 which leads to a constant predicted value. This assumption is not true for all cases. During the Toussaint holidays, students were on break, but the school was still accessible for school staff and researchers. Therefore, it is likely that some of the offices were occupied during that time which contributes to increasing the model's error.

## 5. CONCLUSION

This study aims to analyse the electricity consumption concerning to indoor climate and the occupancy rate of heating and non-heating periods. In a heating period when the HVAC systems are not involved, the indoor climate factors are typically irrelevant to electricity use. An interruption in the system, however, can happen in the energy systems of large-scale buildings. Consequently, users are forced to use another option of heating to achieve a comfort level. For this study case, a higher education building has additional electric heaters (reversible air-conditioners). Thus, a relationship is found through the algorithm between the indoor climate and electricity consumption. Using multiple linear regression, the results have shown an average performance of the developed models, nonetheless, it can prove the impact of the interactive effect of time index, day index, and CO<sub>2</sub> on electricity consumption. This interaction is potential because together, it can better determine the user's presence. For a non-heating period, the models demonstrate a discernible effect of occupancy rate on electricity usage. The overall performance is still limited by the non-uniform situation of the studied buildings. Within the studied period, two of the buildings were renovating to increase their energy efficiency. The occupancy rate does not accurately represent whether a room is occupied or vacant during construction. Additionally, the models return a constant value with the assumption of zero occupancy rate on weekends and vacations. Some part of the building is still in use during the vacation due to the presence of school staff.

A noteworthy limitation of our study lies in the analysis encompassing a post-pandemic period of COVID-19, where occupancy rates were not optimally represented due to the influential impact of the pandemic on building usage patterns. Consequently, the predictive ability of our models remained constrained. However, our findings underscore the crucial of accurate occupancy data for effective energy consumption forecasting. To achieve more precise estimations, future predictive models should incorporate correct data and consider primary contributors such as the main heater's power, equipment loads, lighting intensity, and other relevant factors.

This research highlights the significance of indoor climate conditions and occupancy rates as influential variables impacting electricity consumption. Despite the limitations posed by the post-pandemic period, our study offers valuable insights into the complexities of energy usage in large-scale buildings. The potential for further exploration and elaboration of these findings is substantial, paving the way for advancements in energy efficiency forecasting and management.

## 6. REFERENCES

- Afroz, Z. *et al.* (2018) 'Modeling techniques used in building HVAC control systems: A review', *Renewable and Sustainable Energy Reviews*, 83, pp. 64–84. Available at: <https://doi.org/10.1016/j.rser.2017.10.044>.
- Agarwal, Y., Weng, T. and Gupta, R.K. (2009) 'The energy dashboard: improving the visibility of energy consumption at a campus-wide scale', in *Proceedings of the First ACM Workshop on Embedded Sensing Systems for Energy-Efficiency in Buildings. SenSys09: The 7th ACM Conference on Embedded Network Sensor Systems*, Berkeley California: ACM, pp. 55–60. Available at: <https://doi.org/10.1145/1810279.1810292>.
- Botchkarev, A. (2019) 'Performance Metrics (Error Measures) in Machine Learning Regression, Forecasting and Prognostics: Properties and Typology', *Interdisciplinary Journal of Information, Knowledge, and Management*, 14, pp. 045–076. Available at: <https://doi.org/10.28945/4184>.
- Carlucci, S. *et al.* (2016) 'The effect of spatial and temporal randomness of stochastically generated occupancy schedules on the energy performance of a multiresidential building', *Energy and Buildings*, 127, pp. 279–300. Available at: <https://doi.org/10.1016/j.enbuild.2016.05.023>.
- Chung, M.H. and Rhee, E.K. (2014) 'Potential opportunities for energy conservation in existing buildings on university campus: A field survey in Korea', *Energy and Buildings*, 78, pp. 176–182. Available at: <https://doi.org/10.1016/j.enbuild.2014.04.018>.
- Foucquier, A. *et al.* (2013) 'State of the art in building modelling and energy performances prediction: A review', *Renewable and Sustainable Energy Reviews*, 23, pp. 272–288. Available at: <https://doi.org/10.1016/j.rser.2013.03.004>.
- Fumo, N. (2014) 'A review on the basics of building energy estimation', *Renewable and Sustainable Energy Reviews*, 31, pp. 53–60. Available at: <https://doi.org/10.1016/j.rser.2013.11.040>.
- Hong, T. *et al.* (2018) 'Occupant behavior models: A critical review of implementation and representation approaches in building performance simulation programs', *Building Simulation*, 11(1), pp. 1–14. Available at: <https://doi.org/10.1007/s12273-017-0396-6>.
- Luo, R., Han, Y. and Zhou, X. (2017) 'Characteristics of Campus Energy Consumption in North China University of Science and Technology', *Procedia Engineering*, 205, pp. 3816–3823. Available at: <https://doi.org/10.1016/j.proeng.2017.10.098>.
- Neto, A.H. and Fiorelli, F.A.S. (2008) 'Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption', *Energy and Buildings*, 40(12), pp. 2169–2176. Available at: <https://doi.org/10.1016/j.enbuild.2008.06.013>.
- Pérez-Lombard, L., Ortiz, J. and Pout, C. (2008) 'A review on buildings energy consumption information', *Energy and Buildings*, 40(3), pp. 394–398. Available at: <https://doi.org/10.1016/j.enbuild.2007.03.007>.
- Uyanik, G.K. and Güler, N. (2013) 'A Study on Multiple Linear Regression Analysis', *Procedia - Social and Behavioral Sciences*, 106, pp. 234–240. Available at: <https://doi.org/10.1016/j.sbspro.2013.12.027>.
- Verzani, J. (2011) *Getting Started with RStudio*. O'Reilly Media, Inc.
- Yang, T., Clements-Croome, D. and Marson, M. (2017) 'Building Energy Management Systems', in *Reference Module in Earth Systems and Environmental Sciences*, pp. 291–309. Available at: <https://doi.org/10.1016/B978-0-12-409548-9.10199-X>.