DOI: 10.2478/sspjce-2021-0017

Predicting energy demand of residential buildings: A linear regression-based approach for a small sample size

Soufiane Boukarta[1,2]

[1]Institute of Architecture et Urbanism (IAU), university of Blida 1
[2]VUDD laboratory (EPAU)
e-mail: boukarta.soufiane@univ-blida.dz

Abstract

The key design strategies that reduce the energy demand of buildings are not present in most thermal codes in many countries. Therefore, modeling techniques offer an alternative to combine the architects' modus operandi with the energy efficiency in the early stages of architectural design and with higher speed and precision. However, a review of the scientific literature using modeling techniques shows that most researchers use a relatively large sample of thermal simulations. This paper proposes a simplified method based on the linear regression modeling technique and considers a relatively smaller sample of thermal simulations. A total of 6 key building design strategies were identified, related to the urban context, building envelope, and shape factor. A simulation protocol containing 60 possible combinations was designed by random selection. In the present study, the Pleiades software was used to estimate the annual energy demand for heating and cooling for a typical dwelling in a humid climate zone. A parametric study and sensitivity analysis to identify the most efficient parameters was performed in SPSS 21. The resulting model predicts the annual energy demand with an accuracy of 93.7%, a root mean square error (RMSE) of 5.88, and a scatter index (SI) of 8.59%. The models performed could efficiently and quickly assist architects while designing the buildings in the architectural practice.

Key words: linear regression, small sample size, design of experiment, sensitivity analysis, energy demand, humid climate zone structure

1 Introduction

The issue of buildings' energy efficiency has returned to the forefront of many countries' political agendas because of the high cost of energy and even more so because of the climate change that the world is facing, and which is likely to be exacerbated if no action is taken to mitigate it. Energy consumption in residential buildings is significant worldwide and represents more than 43% of final energy consumption in Algeria [1]. The energy demand of the residential sector grew by 8.3% per year in the last decade (2007-2017), driven by population and economic growth. The residential sector emits more than 34 million tonnes of CO_2 or 36% of the total greenhouse gas (GHG) emissions [1]. Because of this relatively large share of a country's energy consumption and resulting GHG emissions, any increase in energy

efficiency in residential buildings would result in significant energy savings and emissions reductions.

In the usual architectural design process, the architectural project is based on several guidelines to satisfy the main design requirements, namely, integration with the site, aesthetics, functionality of the project, and energy efficiency. The latter is partly overlooked by architects because the key design strategies are not accompanied by their importance in explaining energy demand. Therefore, it is at this stage that the designer should consider design alternatives to reduce the energy demand of his building.
In several countries across the world, there is a thermal code that proceeds by static calculation, but this method is not adapted to the architect's modus operandi when designing the project and which, to sum up, is mainly based on drawings developed based on sketches, such as the glazing rate, orientation, prospect and building shape factor. Therefore, modeling techniques predicting the energy performance of buildings could play a role as a tool to support the design and analysis of energy-efficient buildings. These models simulate the effect of different design parameters that architects tend to use without knowing their impact on energy demand [2].

Linear regression is one of the most common methods used to model the energy demand in buildings [2, 3, 4, 38]. It combines the speed of a simple method and the precision of a dynamic thermal simulation [2]. Several studies used the regression technique for several purposes. Semahi et al. [4] used linear regression to assess the correlation between the bioclimatic potential for cooling and heating with the energy demand obtained from dynamic thermal simulation and have found a correlation of R^2 values of 0.88 and 0.7 respectively. Carlo and Lamberts [5] developed a model based on linear regression to assess the electricity demand in regards of the quality of the commercial building envelope in Brazil. In the same track, Bansal and Bhattacharya [6] performed a model to predict the annual energy demand based on the variation of climatic conditions in India. A strong correlation between weather conditions and the energy demand for both cooling and heating was observed (R^2 0.9). Also, Li et al. [7] developed regression models to predict the energy performance on several types of building envelopes and daylighting controls. Regression models could also be used to predict the monthly energy consumption. Catalina et al. [8] have performed in France a regression model based on the variation of 3 key factors, the thermal transmittance value of the building envelope, window to wall ratio, and the building shape factor.
The sample size used is in general relatively high, with 2000 possible combinations for Bouyer [3], 30000 for Amiri et al. [2], 270 for Catalina et al. [8]. In general, the number of possible simulations depends on the number of variables introduced in the study. For example, if we have 6 variables and a range of 6 variations per variable, the possible combination is 66 which gives 46656 possible combinations which represents a heavy campaign of dynamic simulations. Therefore, authors limit the number of possible combinations by choosing a random number of possibilities, but the number of possibilities in several studies remains important. In this paper, we test how well would the models perform within a relatively small sample size.

2 Literature review

The review of the scientific literature on the main driving factors reducing the energy demand in buildings led to classifying the most influential parameters according to 4 registers [9-12].

2.1 The occupant behavior

The occupant behavior is considered as the main source of uncertainty when we should perform thermal and dynamic simulation. Several researchers [13-18] showed that the occupants' behavior could explain a variation of the energy demand by up to 50 to 200% [19]. Hong and Lin [20] classify the occupant behavior according to their impact on the energy demand into three classes: an austerity class, standard or normal, and wasteful class. The occupant impact on the energy demand is still the object of research and its effect is difficult to predict. Some authors prefer studying the variation of the energy demand based on one occupant in a single space [20]. The approaches based on thermal simulations still use simplified schedule occupancy to be able to predict the energy demand more closely to real cases [4]. Abuimara et al. [21] and in a standard commercial building with a floor area of 1728 m², tested 75 occupants' distribution scenarios with an average density of 0.05 person/m², which remained unchanged for the entire building. A variation in occupancy density varied from 0 to 0.21 person/m². The authors found only 9% as a major impact of the variation of the occupancy density on energy demand. To reduce the rate of discomfort due to occupant density, the authors concluded that it would be necessary to deploy adaptive technologies as a function of occupancy density to the HVAC systems.

2.2 The design factors

The design factors are usually studied within two approaches: (i) The engineering approach which is based on the determinist algorithm used via thermal simulation software [22, 9, 23, 3, 24]. The occupancy schedule is usually simplified [4] and considered as unchanged to be able to assess the design factors' impact in explaining the energy demand; (ii) A statistical approach, also called historicist approach [19]. This approach is based on the final energy consumption of the building stock and a set of indicators to perform models of forecasting and monitoring of energy use. This approach could be used within a top-down method to assess the variation of the energy demand of a city for a long time. The factors mainly used within this approach are socio-economic data, climate conditions, and energy cost. In a contrast, the bottom-up approach considers building stock based on a set of building categories and then generalizes the results to the whole building stock based on the similitude between buildings categories. This approach allows considering the design and/or the socio-economic factors to explain the energy performance [12, 24, 25, 10].

2.3 Urban context

Urban context is a relatively new field of research for energy efficiency studies [9, 11, 12]. The impact of the urban context is mediated by solar admittance, ventilation, lighting, and transportation. It could explain the variation of the energy demand by up to 10% [9] and 50% [11] if we consider both buildings and transportation.

2.4 Building energy systems

Building energy systems consider the quality and the technology of the HVAC systems. Dascalaki et al. [10] introduced this register within a statistical approach and Abuimara et al. [21] suggested improving the adaptability of the HVAC technology according to the buildings' occupancy.

The number of variables selected varies relatively from one study to another, 17 for Amiri et al. [2], 6 for Ali-Toudert and Weidhaus [26] and 7 for Bouyer [3], and 3 for Catalina et al. [8]. It is possible to classify the most important variables in the three main components: (i) The environment context. Several studies introduce the climatic conditions, height to width ratio (prospect), vegetal density, the solar orientation, built density, and the amplitude of air temperature which is an interesting variable when we are interested in a comparison between two climate zones; (ii) The design factors. The most important factor representing the design is its building shape factor which is equal to the area to the volume ratio. It's commonly intended that the smaller value of the building shape factor helps to reduce energy demand. In other studies, we find the passive volume which represents the depth of 2 times the ceiling height which is usually equal to 6 meters. This band of 6 meters of depth is naturally lighted, sunny, and ventilated, so its energy demand is smaller than the other spaces of the building considered [9, 11]; (iii) The third component is the quality of the insulation envelope regarding the climatic conditions. In several studies, we found glazing ratio, the thermal transmittance of windows, the thermal resistance of walls, floors, and ceilings, and the thermal mass of materials. The thermal inertia of a material is its ability to provide a temperature phase shift in the heat transfer between the exterior and a building's interior. It is a recommended quality in climates with important thermal amplitude; (iv) Occupancy schedule. The variation of this variable could impact the energy demand from 50% to 200% [19]. In several studies, the occupancy schedule appears as an important variable. Shakouri et al. [17] have found that calibrating the occupancy schedule is more important than the structural approach (reduction of the cooling/heating volume and an increase of the glazed area) with 71% and 33% respectively of impact.

To perform a successful regression model, several parameters shave to be fitted. All the data must satisfy all the assumptions of the linear regression; the normality distribution, independence of independent variables, linearity, and homoscedasticity of residuals which is also known as the homogeneity of the variance [27]. To ensure that the data are normally distributed, a variable with a value larger than 1 for Skewness or out of the range of 2, –2 for the Kurtosis test must be transformed before performing the linear regression [28]. In linear regression, variables considered for the model must be independent without similarity between them. To this end, the variance inflation factor (VIF) larger than 5 and minimum tolerance lesser than 0.2 [27] indicate that there is multicollinearity, and the independent variable has to be removed from the model.

To validate a linear regression model, we must check its *P*-value which has to be less than 0.05 to be considered a significant model with a 95% confidence interval. Then, the homoscedasticity of the standardized residual must also be checked. To do so, the *P*-value of the Kolmogorov-Smirnova and Shapiro-Wilk should be larger than 0.05 to ensure that the model has a homogeneous variance [28].

The quality of the performed linear regression model is checked based on the coefficient of determination R^2, RMSE, and the Scatter Index (SI). The Root Mean Squared Error (RMSE)

has to be closer to zero, and the Scatter Index indicates the percentage error. The SI is the RMSE divided by the mean of the observations and multiplied by 100 [29]. RMSE is calculated based on the formula (1).

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(S_i - O_i)^2} \quad (1)$$

where n is the number of the observations, S_i is the simulated cases and O_i is the observed cases.

The formula of the linear regression model is as follows:

$$Y = C + \beta_1(X_1) + \beta_2(X_2) + \beta_3(X_3) + \beta_4(X_4) + \cdots \beta_i(X_i) \quad (2)$$

with Y as the dependent variable (cooling, heating, or total energy demand), C is constant, β is the coefficient of regression and X_i is the independent variable. To find the best regression model that predicts accurately the energy consumption, stepwise regression [30] is usually performed especially when the number of explanatory variables is large.
In this paper, the engineering approach is chosen, and we focus our attention on the energy efficiency of buildings in humid climate zone, the urban context and building parameters are the main registers considered in this paper.

3 Methods

The method adopted is based on 6 steps as follows (see Figure 1):

(1) Identification of the potential variables from the literature review that could explain the energy demand variation. We have identified six variables: (i) Environmental variables with the building height to street width ratio also called prospect (H/W); (ii)the solar orientation (OR); (iii) The envelope insulation, with the thermal resistance of exterior walls (*Re*); (iv) the thermal transmittance of windows (U_w); (v) glazing ratio which is the ratio of window to room area (Gr). We have not selected the ceiling insulation because the actual thermal code considers already a high thermal resistance; (vi) and the building shape factor (A/V) which is calculated by the ratio of the area of contact with the exterior air and the volume of the dwelling. The lower is the building shape factor, the lesser is the energy demand for cooling and heating because the area of thermal loss is lesser.

(2) The second step is to prepare a protocol of simulation containing all the range of variation for each variable (see Table 1). The occupancy schedule is unchangeable for all the cases to be able to assess the main effect of the 6 variables selected for this paper, see appendix A1. We have set a variation that includes at least a minimum, an average, and a maximum value for each variable so that the possibilities can always be represented within this range of variation. This approach reduces the number of simulations and widens the predictive potential of the models.

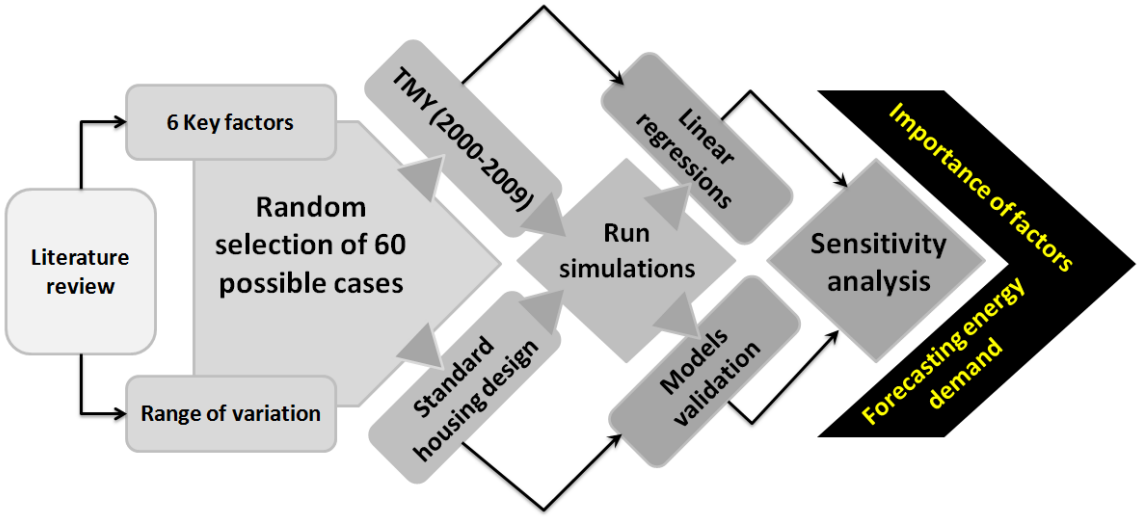


Figure 1: Method framework

(3) The third step is to conceive a simulation plan with a random selection of variables. A set of 60 simulations was prepared.
(4) The fourth step is to conceive a typical dwelling to be considered as a baseline model (see Figure 2).
(5) The fifth step is running the 60 simulations under Pleiades software to obtain the heating and cooling energy demand (kWh/m².y). This software has been validated by the CNERIB (the National Centre for Studies and Research in the Building Industry) which is the authority to issue certificates of conformity in Algeria. The typical Meteorological Year (TMY) used covers the period of ten years (2000-2009). This data is obtained from Meteonorm [31] based on the data of the station of Dar Elbeida (Algiers).
(6) The sixth step is the modeling stage. A total of three regression models are performed for cooling, heating, and total energy demand.

Table 1: Variables range variation

Variables*	Variable's name	Range of variation
H/W	prospect	1, 2, 3, 4, 5, 6
OR	orientation	0, 45, 90, 135, 80, 225, 270, 315
Re (m².K/W)	thermal resistance	Conventional Re (0.63), optimized Re (2.67, 1.75, 3.95)
U_w (W/m².K)	thermal transmittance	(simple glazing) 5, (double glazing) 3.8, 2.79, 1.70, (triple glazing) 0.72
Gr	glazing ratio	10%, 20%, 30%
A/V	shape factor	0.31, 0.6, 0.66, 0.7
TMY	typical meteorological year	2000-2009

* Schedule occupancy (see Appendix A1), Thermal characteristics of the materials used in the simulations (see appendix A2)

Space	Area (m²)
1-Living room	19.25
2-Bedroom 1	15.75
3-Bedroom 2	12.25
4-wc	1.75
5-Bathroom	4.5
6-Kitchen	8.75
7-Hall	7.75
Total	70 m²



Figure 2: Left. Areas of the spaces. Right. Shape of the typical dwelling

4 Results

4.1 Algiers Climate zone

This section presents a short climate analysis of the city of Algiers which belongs to the A climate zone, which is characterized by a humid climate according to the classification of the thermal code C3-2. The humid climate is globally characterized by a hot and humid climate in summer and mild, rainy, and humid winters. According to a comfort range of 20-27 °C [32] the heating (HDD_{20}) and cooling (CDD_{27}) degree days are respectively 1524 and 124 in the city of Algiers. The average monthly temperature in winter is about 10.5 °C with a daily minimum of -0.4 °C. The average monthly temperature in summer is about 25.5 °C with a maximum of 38°C. The Monthly mean data (2000-2009) are presented in Table 2 and give an overview of the climate background.

Table 2: Variables range variation

Month	*Ta**	Ta_{min}	Ta_{max}	*Rh*	*Gh*	*Sh*	*Dh*	*FF*	*Sd*	*RR*	*Tg*
Jan	9.94	5.24	15.74	81.30	65	28	37	2.39	166	77	10.34
Feb	10.99	6.00	16.86	77.24	85	43	42	2.50	170	51	10.16
Mar	13.54	8.15	19.50	75.71	127	67	60	2.69	200	55	13.6
Apr	15.49	9.89	21.22	75.33	161	88	73	2.99	223	39	15.76
May	19.06	13.60	24.53	73.34	196	112	84	2.99	280	46	19.72
Jun	22.90	17.08	28.59	68.95	204	118	86	3.09	300	2	23.98
Jul	26.08	20.63	31.51	67.94	227	144	83	3.10	330	3	26.68
Aug	26.27	20.77	31.87	68.67	193	112	81	2.90	300	10	27.23
Sep	22.99	17.85	28.50	73.14	153	93	60	2.80	241	31	23.7
Oct	20.13	15.00	26.00	74.42	107	53	54	2.29	209	40	19.6
Nov	14.47	9.88	20.04	77.09	71	35	36	2.50	172	98	13.69
Dec	11.80	7.46	17.22	78.04	65	33	32	2.59	161	95	10.94

* *Ta*: Temp Avg Daily (°C), *Rh*: Rel Humidity (%), *Sd*: sunshine duration (hours), Ta_{min}: Temp Avg Daily low (°C), *Gh*,*Sh*, *Dh*: Glo Horiz Rad, direct, diffuse (Wh/m²), *RR*: precipitation (mm), Ta_{max}: Dry Bulb Temp, Avg Daily high (°C), *FF*: wind speed (m/s), *Tg*: Ground temperature (°C)

4.2 Bioclimatic potential of the humid climate zone

The bioclimatic potential is estimated based on Climate Consultant software analysis which is free didactical software developed by the UCLA energy design tools group [33]. Based on the air temperature and air humidity, the psychometric diagram shows how to improve the comfort zone based on 15 strategies. In this paper, we have used the ASHRAE 55-2010 Thermal comfort model which assumes that occupants adapt their clothing to thermal conditions (1.0 to 1.3 met). As output, this software contributes in terms of percentage and number of hours of thermal comfort gained for each proposed design strategy.

The psychometric diagram indicates that the climate of Algiers is 13.3% of the year thermally comfortable. With 33.1% (2897 hours), the internal heat gains represent the most important factor reducing the heating demand, and the passive solar gains and thermal mass could improve further the comfort up to 12.1% (1061 hours) and 2.7% (239 hours) respectively. Sun shading of windows and adaptive comfort ventilation contribute by up to 14.2% (1242 h) and 18.8% (1651 h) respectively in reducing cooling demand. Also, dehumidification is required and could improve the thermal comfort up to 13.2% (1159 h). All the passive design strategies cited above could provide thermal comfort for up to 53.9% per year. Active cooling and heating are required up to 8.3% (724 h) and 24.5% (2143 h) respectively (see Figure 3).

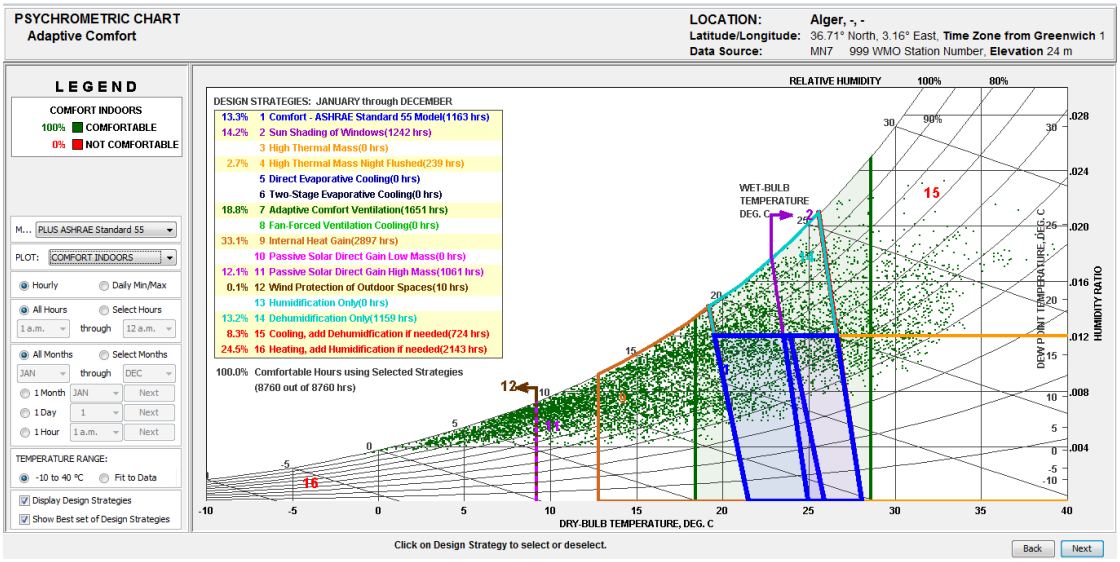


Figure 3: Psychometric charts of the climate of Algiers

4.3 Regression models

A random combination between the six factors selected is presented in Figure 4. Each combination represents a simulation, a total of 60 simulations were performed under Pleiades software. Based on the first stage regression modeling, we have identified 4 cases as outliers and removed them from the final regression models. We have performed three regression models, for cooling, heating, and total energy demand.

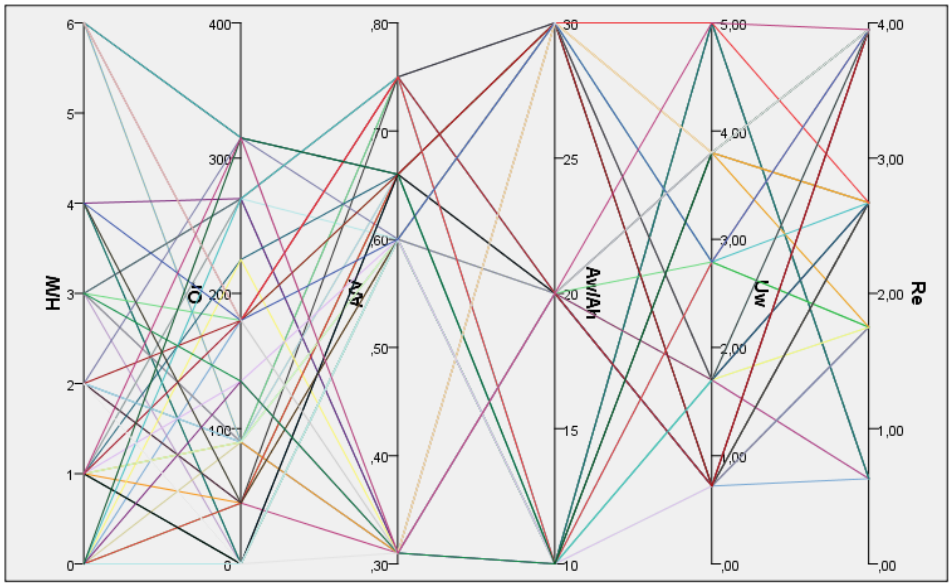


Figure 4: The random combination between variables

4.3.1 Multicollinearity test

The first test to check is the multicollinearity test which indicates that all the independent variables selected for the regression model are not similar. All the VIFs indicate a value closer to 1 and a tolerance value larger than 0.2 (see Table 3).

Table 2: Variables range variation

Variables*	**Total energy demand**		**Heating energy demand**		**Cooling energy demand**	
	Tolerance	VIF	Tolerance	VIF	Tolerance	VIF
log10_U_w	.923	1.083	.861	1.162	.909	1.100
log10_*Re*	.801	1.248	.760	1.316	.920	1.087
Aw/Ah	-	-	.851	1.175	.881	1.135
H/W	.971	1.030	-	-	.955	1.047
A/V	.956	1.046	.940	1.064	-	-
OR	.800	1.249	.766	1.306	-	-

* U_w: Thermal transmittance of windows, *Re*: thermal resistance of exterior walls, Aw/Ah: glazing ratio, H/W: Prospect, A/V: shape factor, OR: orientation

4.3.2 Linearity test of the models

Based on the scatter plot of the standardized predictions and residuals values, it appears that the relationship is linear because the relationship between the results obtained from the simulations and those predicted by models is almost zero since the residuals are randomly scattered around zero [34] (see Figure 5).

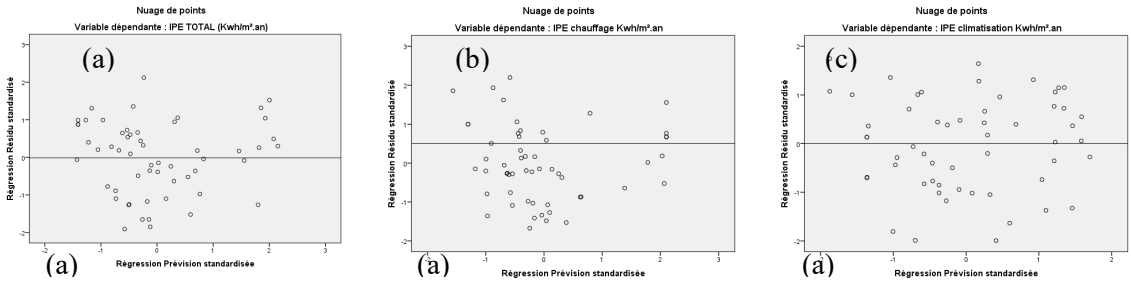


Figure 2: linearity test of residuals. (a) Total energy demand, (b) Heating and (c) Cooling energy demand

4.3.3 Normality test

The normality test is checked by the Skewness and Kurtosis test for both independent variables and for residuals. Tables 4 and 5 indicate clearly that the data fit the normality test.

Table 4: Normality test of the independent variables

	Aw/Ah*	H/W	A/V	U_w	Or	*Re*
Skewness	.139	.846	-.597	.152	.226	-.102
Kurtosis	-1.612	-.092	-1.301	-1.425	-1.292	-1.153

* U_w: Thermal transmittance of windows, *Re*: thermal resistance of exterior walls, Aw/Ah: glazing ratio, H/W: Prospect, A/V: shape factor, OR: orientation

Table 5: Normality test of residuals

	Total energy demand	Heating energy demand	Cooling energy demand
Skewness	-.210	.272	-.240
Kurtosis	-.640	-.501	-.777

Also, the Q-Q plots of standardized residuals show a linear distribution of quantiles which indicate that the models fit the normality assumption.

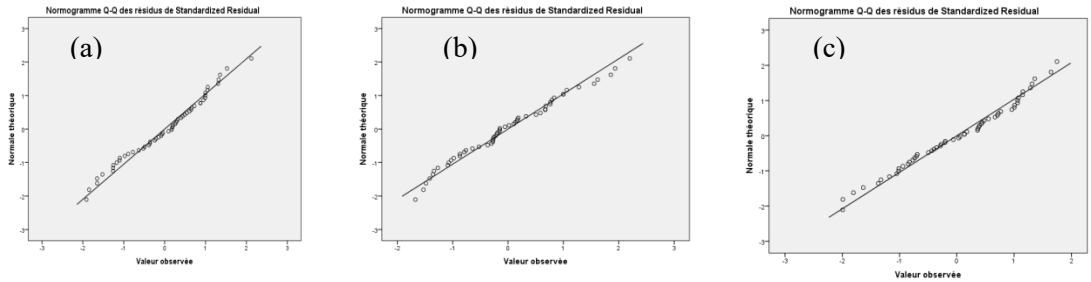


Figure 3: Q-Q plot of residuals. (a) Total energy demand, (b) Heating, and (c) Cooling energy demand

4.3.4 Homoscedasticity test

The homoscedasticity test is satisfied when the error value doesn't vary considerably as the value of the predictor changes. Table 6 shows that residuals do satisfy both the Kolmogorov-Smirnov and the Shapiro-Wilk test, and the regression models performed could predict the energy demand for both heating and cooling.

Table 6: Homoscedasticity test

Models	Kolmogorov-Smirnov		Shapiro-Wilk	
	Statistic	Sig	Statistic	Sig
Total energy demand	.088	**,200***	.977	**.353**
Heating energy demand	.079	**,200***	.976	**.318**
Cooling energy demand	.092	**,200***	.974	**.256**

* Sig>0.05 means that the homoscedasticity test is satisfied

Based on the four tests presented above, the multicollinearity, linearity, normality, and the homoscedasticity tests, the three regression models performed could be considered as a valid model and its accuracy depends on the items presented in the next section.

4.4 The accuracy of the regression models

The accuracy of a regression model is linked to three indices:

(1) The coefficient of determination (R^2) which represents the power of prediction of the models. The R^2 is high for the three regression models performed, 0.937, 0.885, and 0.895 for the total energy demand, heating, and cooling energy demand respectively (see Figures 7 and 8).

(2) The RMSE has the same unit as the energy demand which means that its value is relative to the average value of observed values. Within the three models performed, the RMSE calculated is 5.88, 5.49, and 3.14 (kWh/m².y) for the total, heating and cooling energy demand respectively. These results indicate that the three models indicate quite well the energy demand.

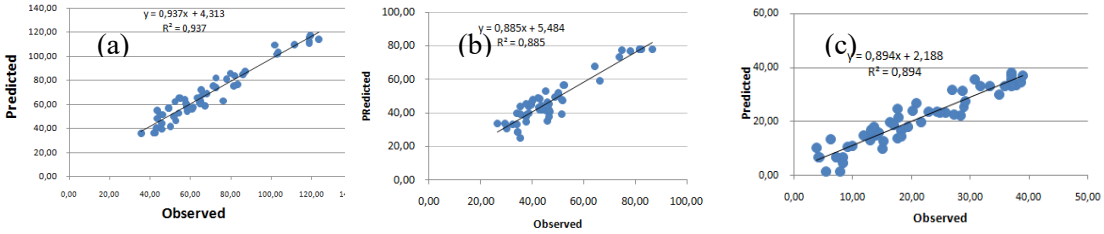


Figure 4: Coefficient of determination of: (a) Total, (b) Heating and (c) Cooling energy demand

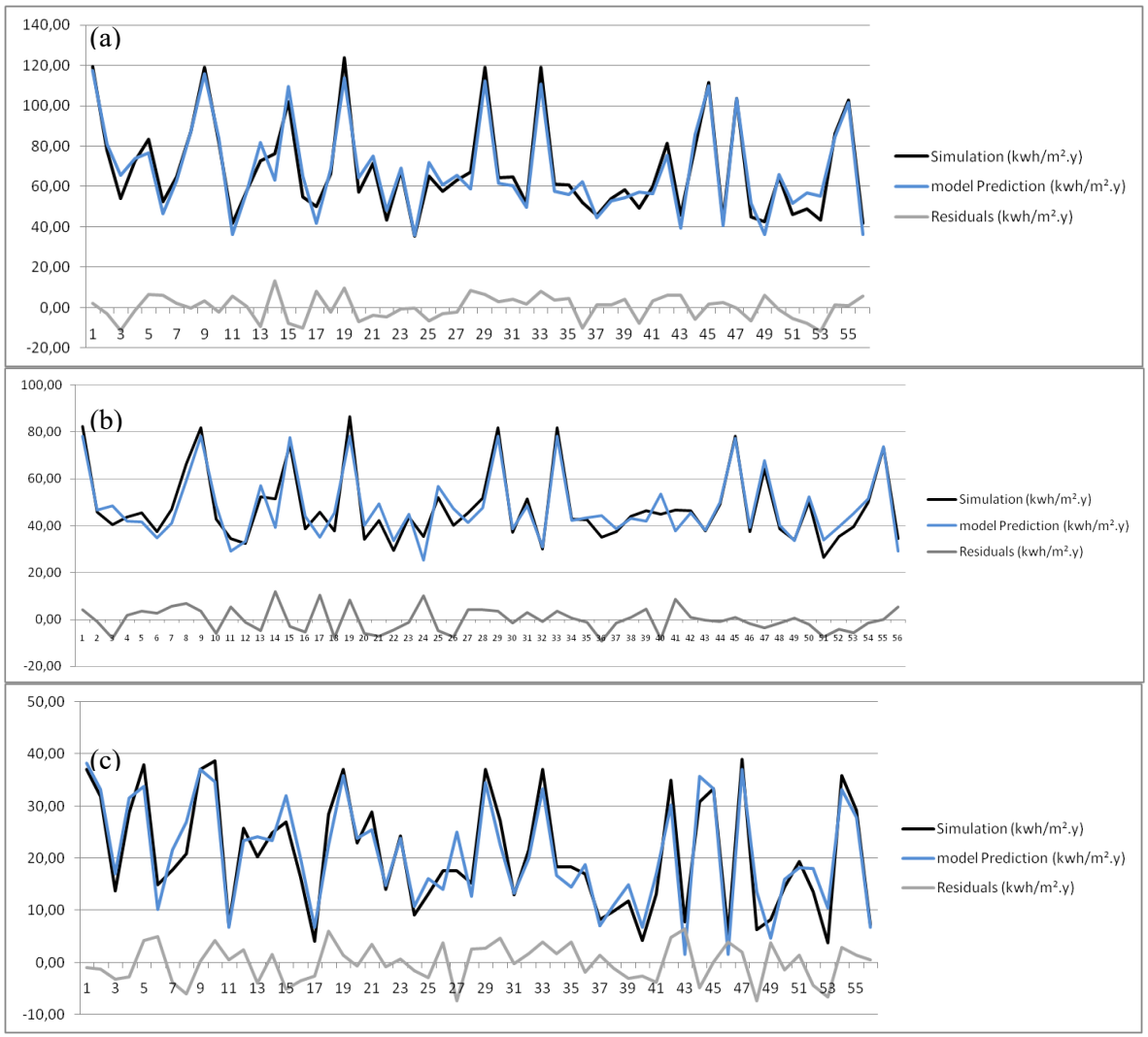


Figure 5: Total model predictions results vs. simulated energy demand. (a) Total, (b) Heating, and (c) Cooling energy demand

(3) The Scatter Index (SI) precisely presents the importance of each model by reporting the RMSE to the average value of the observed one (simulation results). The SI of the three models is 8.49%, 10.74%, and 16.84% for the total, heating, and cooling energy demand (see Figure 9).

These results mean that the most accurate model obtained is the one for the total energy demand.

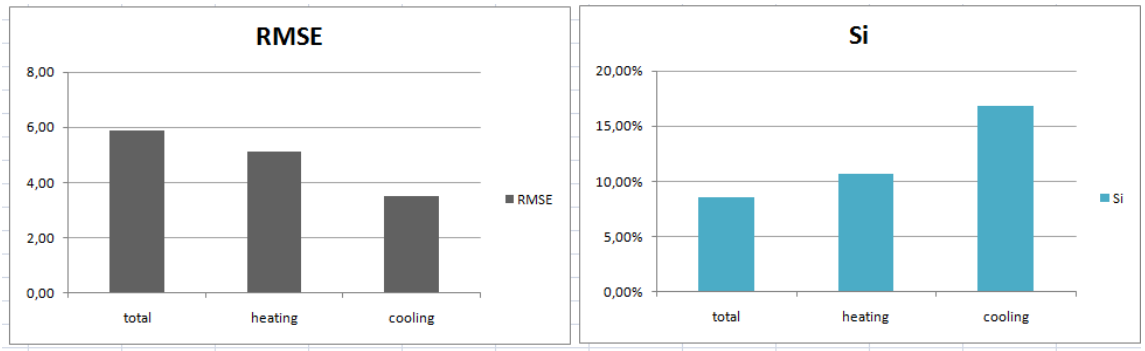


Figure 6: RMSE and SI of the three models

4.5 Sensitivity analysis

The summary of the three regression models is presented in Table 7. All the variables fit the threshold of the *P*-value (<0.05). It means that the variables selected explain well the variance of the energy demand. Also, the *ß* value, which measures the total effect of the predictor variables, indicates that exterior walls' thermal resistance has the greatest impact on reducing energy demand. Indeed, the log10 of the thermal resistance of exterior walls reduces energy consumption by up to -62 kWh/m².y for each standard deviation of the thermal resistance.

Table 7: Summary of the regression models

	Total		Heating		Cooling	
	ß	***P*-value**	***ß***	***P*-value**	***ß***	***P*-value**
constant	93.775	.000	80.619	.000	14.464	.000
log10_U_w	28.362	.000	7.634	.003	19.665	.000
log10_*Re*	-62.009	.000	-37.905	.000	-23.754	.000
H/W	-1.716	.001	-	-	-1.224	.000
A/V	-13.962	.008	-16.960	.000	-	-
Or	-.024	.007	-.026	.001	-	-
Aw/Ah	-	-	-.524	.000	.524	.000
	R^2=0.937		**R^2=0.885**		**R^2=0.895**	

To assess the importance of each variable we have computed the ratio of *ß* to standard deviation values and the results are presented in Figure 10. The most important factors are those related to the envelope of the buildings. For every +1 m².K.W^{-1} in the thermal resistance of exterior walls, +0.2 in the shape factor (A/V), and -1 W/m².K in the thermal transmittance of windows values, the energy demand could be reduced by up to -32.63, -16.4, and -13.56 kWh/m².y respectively. Similar findings were found by Boukarta and Berzowska-Azzag [35]. The effect of the glazing ratio is weak, and it neutralizes itself because the energy demand reduced for the heating demand is required for the cooling demand. It is worth clarifying that the importance of the glazing ratio could be higher with higher U_w values because the thermal loss and the solar gains would be also higher. The impact of the orientation is also very weak, and it could be explained by the fact that the orientation loses its importance with the higher level of envelope insulation.

The same impact was found by Tahmasebi et al. [36]. The height to width ratio (H/W) reduces slightly the energy demand for cooling by increasing the area of sun shading.

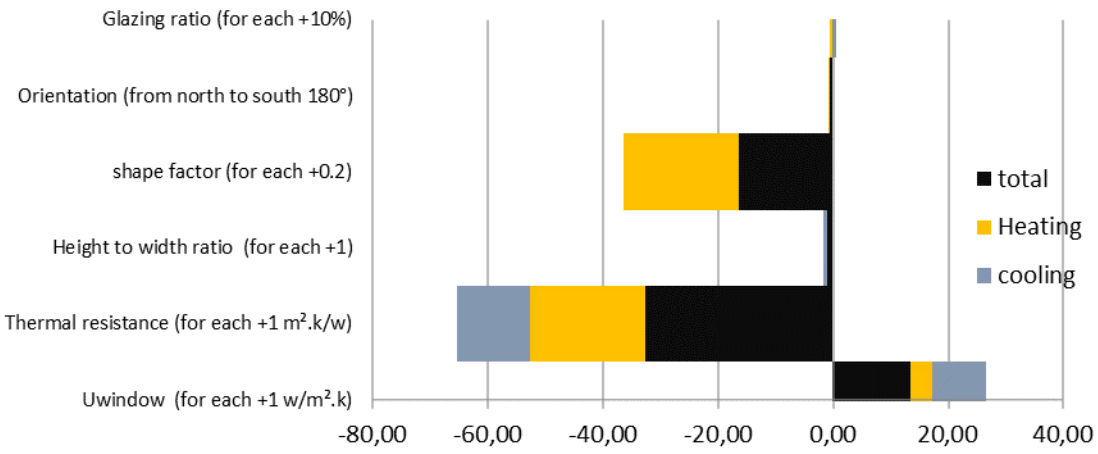


Figure 70: The importance of variables

The best case, observed in our sampling is the case combining the highest thermal resistance of exterior walls, U_w, and an average value of 0.66 for the shape factor (A/V ratio). Also, the best orientation is observed when the living room and room number 2 were oriented directly to the south (180°) and it is worth here to notify that the orientation is mainly depending on the occupancy schedule. The best glazing ratio is 30% because the HDD is higher than the CDD. And the best prospect ratio is observed with the value H/W=2 because it ensures a profitable area of sun shading which could be beneficial for both cooling and heating demand (see Figure 11). Our results fit with the recommendations obtained from the psychometric diagram presented in Figure 5.

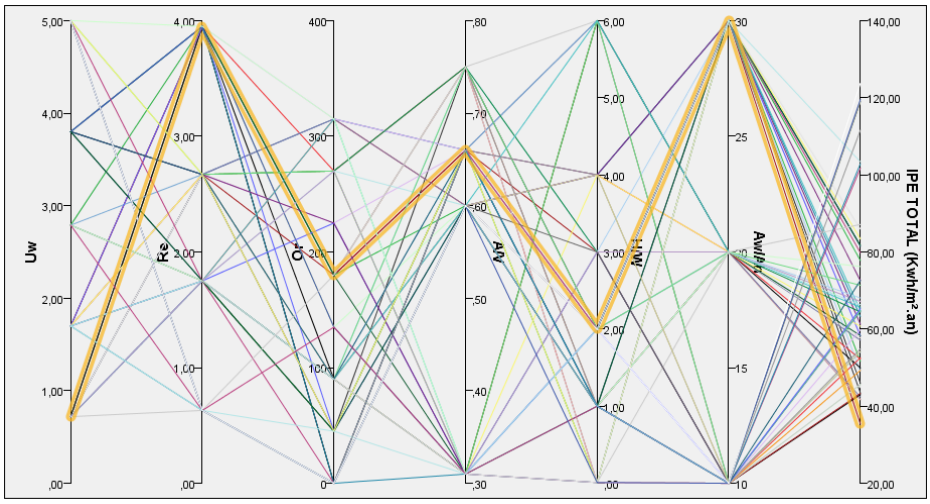


Figure 81: The best combination of the six variables

5 Conclusion

The purpose of this paper is to perform regression models to predict the energy demand of cooling and heating of a residential building in the humid climate zone of Algiers based on the variation of a set of 6 design variables largely used by architects in different stages of design. A small sample size of 60 thermal simulations was performed and used as data to perform regression models. 5 parameters were kept for the total energy demand, 5 also for the heating demand, and 4 for the cooling energy demand. The regression models were tested for the four assumptions to validate the regression models and we have also used the RMSE and Scatter Index (SI) to assess the accuracy of the models performed. The results obtained from the models' prediction show that the errors in the prediction of the annual energy demand compared to the simulated ones were acceptable and that the total energy demand regression model (R^2=0.937 and SI=8.49%) is the best followed by the cooling and heating energy demand models. The sensitivity analysis performed shows that the parameters linked to the shape and the envelope insulation have the highest impact, while the glazing ratio, the height to width ratio, and the orientation of the building have a relatively weak impact on the energy demand. Also, the results obtained confirm that it is possible to utilize the predicting models based on a relatively small sample size.
The performed regression models could be used to predict the annual energy demand of residential buildings and could also be used as a baseline to thermal rehabilitation, and it could also test the different strategies to reduce the energy demand at a larger scale.

References

[1] APRUE. (2019), L'Agence Nationale pour la Promotion et la Rationalisation de l'Utilisation de l'Energie (2005), La situation énergétique nationale, chiffre 2017, édition 2019.

[2] AMIRI, Shideh Shams, MOTTAHEDI, Mohammad, et ASADI, Somayeh. Using multiple regression analysis to develop energy consumption indicators for commercial buildings in the US. Energy and Buildings, 2015, vol. 109, p. 209-216.

[3] BOUYER, Julien. Modélisation et simulation des microclimats urbains-Etude de l'impact de l'aménagement urbain sur les consommations énergétiques des bâtiments. 2009. PHD dissertation. Université de Nantes, France.

[4] SEMAHI, Samir, ZEMMOURI, Noureddine, SINGH, Manoj Kumar, et al. Comparative bioclimatic approach for comfort and passive heating and cooling strategies in Algeria. Building and Environment, 2019, vol. 161, p. 106271.

[5] CARLO, Joyce et LAMBERTS, Roberto. Development of envelope efficiency labels for commercial buildings: Effect of different variables on electricity consumption. Energy and Buildings, 2008, vol. 40, no 11, p. 2002-2008.

[6] BANSAL, Naresh K. BHATTACHARYA, Amitabh. Parametric equations for energy and load estimations for buildings in India. Applied thermal engineering, 2009, vol. 29, no 17-18, p. 3710-3715.

[7] LI, D. H. W., WONG, Sai Li, et CHEUNG, King Lok. Energy performance regression models for office buildings with daylighting controls. Proceedings of the Institution of Mechanical Engineers, Part A: Journal of Power and Energy, 2008, vol. 222, no 6, p. 557-568 568.

[8] CATALINA, Tiberiu, VIRGONE, Joseph, et BLANCO, Eric. Development and validation of regression models to predict monthly heating demand for residential buildings. Energy and buildings, 2008, vol. 40, no 10, p. 1825-1832.

[9] RATTI, Carlo, BAKER, Nick, et STEEMERS, Koen. Energy consumption and urban texture. Energy and buildings, 2005, vol. 37, no 7, p. 762-776.

[10] DASCALAKI, Elena G., DROUTSA, Kaliopi, GAGLIA, Athina G., et al. Data collection and analysis of the building stock and its energy performance—An example for Hellenic buildings. Energy and Buildings, 2010, vol. 42, no 8, p. 1231-1237.

[11] SALAT, Serge. Les villes et les formesurbaines: sur l'urbanisme durable. France, 2011.
[12] BOUKARTA, Soufiane. BEREZOWSKA-AZZAG, Ewa. 'Urban island' as an energy assessment tool: The case of Mouzaia, Algeria. Journal of Applied Engineering Science, 2017, vol. 15, no 2, p. 128-139.
[13] BARTIAUX, Françoise. A socio-anthropological approach to energy-related behaviours and innovations at the household level. ECEEE (European Council for Energy-Efficient Economy): ECEEE 2003 summer study proceedings–Time to turn down energy demand, 2003, p. 1239-1250.
[14] SWAN, Lukas G. et UGURSAL, V. Ismet. Modeling of end-use energy consumption in the residential sector: A review of modeling techniques. Renewable and sustainable energy reviews, 2009, vol. 13, no 8, p. 1819-1835.
[15] SUBRÉMON, H. Pour une anthropologie sensible de la consommation d'énergie. Université de Paris X-Nanterre. 2009. PHD dissertation, France.
[16] YU, Zhun, FUNG, Benjamin CM, HAGHIGHAT, Fariborz, et al. A systematic procedure to study the influence of occupant behavior on building energy consumption. Energy and buildings, 2011, vol. 43, no 6, p. 1409-1417.
[17] SHAKOURI, Mahmoud, BANIHASHEMI NAMINI, S. S., et JAVAHERI, Amin. Analysis and Comparison of Impacts of Design Optimization Approach with Occupant Behaviour on Energy Consumption Reduction in Residential Buildings. International Journal of Engineering and Technology, 2012.
[18] OUYANG, Jinlong et HOKAO, Kazunori. Energy-saving potential by improving occupants' behavior in urban residential sector in Hangzhou City, China. Energy and buildings, 2009, vol. 41, no 7, p. 711-720.
[19] DALL'O', Giuliano, Annalisa GALANTE, Marco TORRI, "A methodology for the energy performance classification of residential building stock on an urban scale". Energy and Buildings 48 (2012) 211–219.
[20] HONG, Tianzhen et LIN, Hung-Wen. Occupant behavior: impact on energy use of private offices. Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States), 2013.
[21] ABUIMARA, Tareq, O'BRIEN, William, et GUNAY, Burak. Quantifying the impact of occupants' spatial distributions on office buildings energy and comfort performance. Energy and Buildings, 2021, vol. 233, p. 110695.
[22] BOUKARTA, Soufiane. BEREZOWSKA-AZZAG, Ewa. Energy demand of occupant's spatial modification in residential buildings. Case study of Médéa, Algeria. Selected Scientific Papers-Journal of Civil Engineering, 2018, vol. 13, p. 15-28.
[23] STEEMERS, Koen. Energy and the city: density, buildings and transport. Energy and buildings, 2003, vol. 35, no 1, p. 3-14.
[24] MAÏZIA, Mindjid, SÈZE, Claire, BERGE, Sébastien, et al. Energy requirements of characteristic urban blocks. Proc. of CISBAT 2009-Renewables in a changing climate-From Nano to urban scale, 2009.
[25] CAPUTO, Paola, COSTA, Gaia, et FERRARI, Simone. A supporting method for defining energy strategies in the building sector at urban scale. Energy Policy, 2013, vol. 55, p. 261-270.
[26] ALI-TOUDERT, Fazia et WEIDHAUS, Juliane. Numerical assessment and optimization of a low-energy residential building for Mediterranean and Saharan climates using a pilot project in Algeria. Renewable Energy, 2017, vol. 101, p. 327-346.
[27] LI, Cheng, LI, Junxiang, et WU, Jianguo. What drives urban growth in China? A multi-scale comparative analysis. Applied geography, 2018, vol. 98, p. 43-51.
[28] KIM, Hae-Young. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restorative dentistry & endodontics, 2013, vol. 38, no 1, p. 52-54.
[29] NEILL, Simon P. et HASHEMI, M. Reza. Ocean modelling for resource characterization. Fundamentals of ocean renewable energy, 2018, p. 193-235.
[30] Hastie, T., Tibshirani, R., & Friedman, J. (2009). Linear methods for regression. In The elements of statistical learning (pp. 43-99). Springer, New York, NY.
[31] METEOTEST, Software METEONORM 7, 2016. https://meteonorm.com/en/.
[32] DTR C3.2, Règlement Thermique des Bâtiments d'Habitation - Règles de calcul des déperditions Calorifiques, Fascicule 1, Document Technique Réglementaire, Ministère de l'habitat et de l'urbanisme, CNERIB, Alger, 2004.
[33] UCLA, Climate Consultant 6.0, 2017. www.energy-design-tools.aud.ucla.edu/climate-consultant/.
[34] SANTOS NOBRE, Juvêncio et DA MOTTA SINGER, Julio. Residual analysis for linear mixed

models. Biometrical Journal: Journal of Mathematical Methods in Biosciences, 2007, vol. 49, no 6, p. 863-875.

[35] BOUKARTA, Soufiane. Déterminants de la forme urbaine générant le potentiel de maitrise de l'énergie en zone semi-aride. 2019. PHD dissertation. Ecole polytechnique d'architecture et d'urbanisme (EPAU).

[36] TAHMASEBI, Mohammad Mahdi, BANIHASHEMI, Saeed, et HASSANABADI, Mahmoud Shakouri. Assessment of the variation impacts of window on energy consumption and carbon footprint. Procedia engineering, 2011, vol. 21, p. 820-828.

Appendix A1: Schedule occupancy

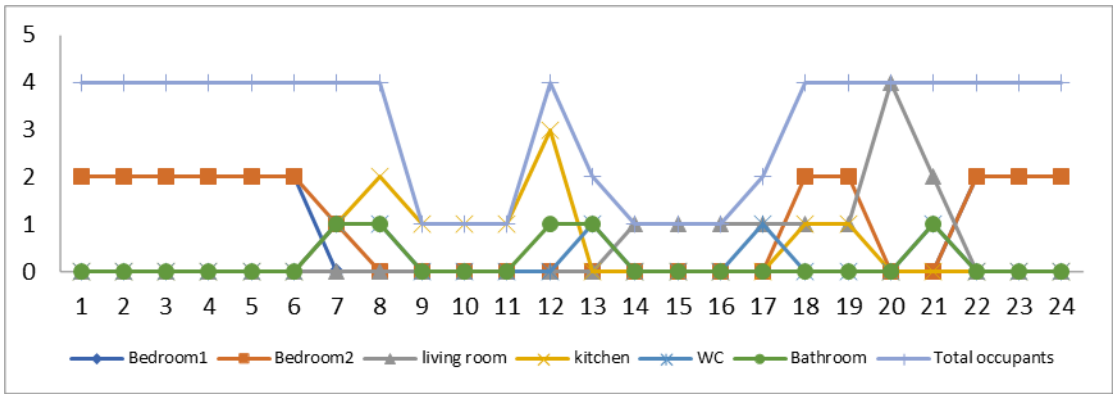


Figure A1- 1 : Schedule occupancy of weekdays

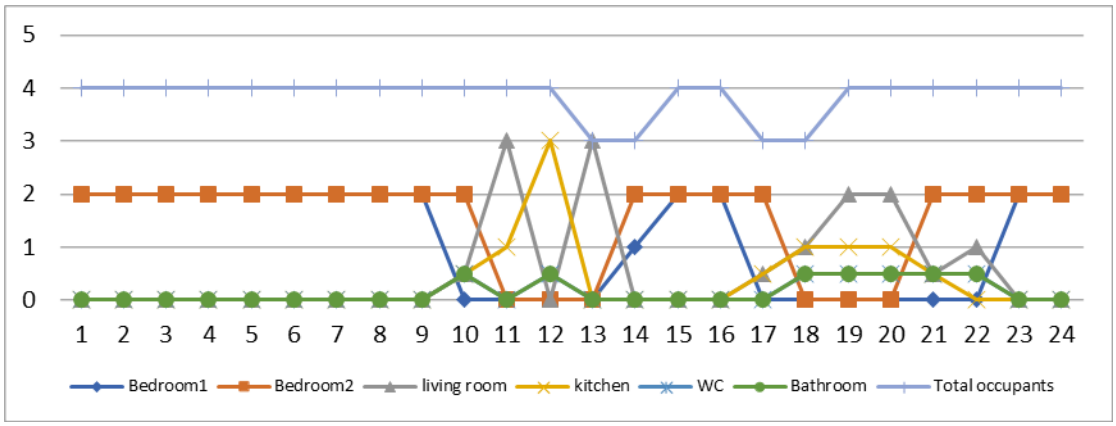


Figure A1- 2 : Schedule occupancy of Fridays

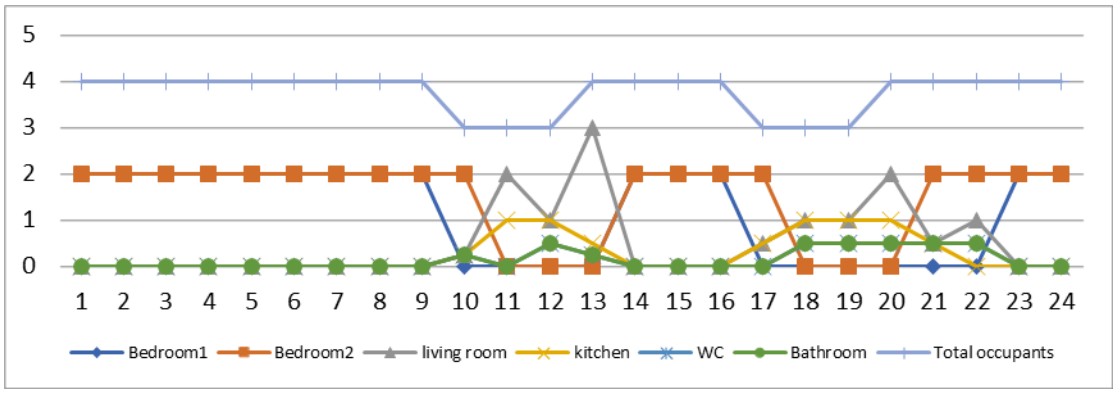


Figure A1- 3 : Schedule occupancy of Saturdays

Appendix A2: Thermal characteristics of the materials used in the simulations

Table A2-1: Thermal characteristics of the materials used in the simulations

Envelope			
Exterior Walls	**Composition**	**Thickness (cm)**	**Thermal Resistance (m²K/W)**
Double walls	Cement mortar Hollow brick Air gap hollow brick Cement mortar	02.0 15.0 05.0 10.0 02.0	0.01 0.27 0.16 0.18 0.01
			Total: 0 .63
Double walls+insulation	Cement mortar Hollow brick Expanded polystyrene Hollow brick Cement mortar	02.0 15.0 05.0 10.0 02.0	0.01 0.27 1.28 0.18 0 .01
			Total: 1.75
Mono Walls (certified)	Cement mortar Mono wall Cement mortar	02.0 30.0 02.0	0.01 2.65 0 .01
			Total: 2.67
Mono Walls +insulation	Cement mortar Mono wall Expanded polystyrene Cement mortar	02.0 30.0 05.0 02.0	0 .01 2.65 1.28 0.01
			Total: 3.95
windows	**Type**	***U*$_W$ (W/m².K)**	
	Single glazing Double glazing A Double glazing B Double glazing C Triple glazing	05.00 03.80 02.79 01.70 0.72	
Exterior door	Wooden door	05.00	
Floor and ceiling			
	composition	**Thickness (cm)**	**Thermal resistance (m²K/W)**
Ceiling	Cement mortar Slabs Heavy Concrete Expanded polystyrene Concrete slope Bituminous felt 36-S Gravel	02.00 16.00 04.00 10.00 03.00 03.00 03.00	0.01 0.13 0.02 2.56 0.08 0.13 0.08
			Total: 3.01
Floor: Slab on ground	Heavy concrete Expanded polystyrene Mortar Tiling	20.00 04.00 05.00 02.00	0.11 1.03 0.04 0.01
			Total: 1.19
Interior			
Interior wall	**Composition**	**Thickness (cm)**	**Thermal resistance (m²K/W)**
Brick partition	Plaster Hollow brick plaster	02.00 05.00 02.00	0.06 0.10 0.06
			Total: 0.22
Interior door	**material**	***U*$_W$ (W/m².K)**	
Door	Wooden door	05.00	