Revista Hábitat Sustentable Vol. 13, N°. 2. ISSN 0719 - 0700 / Págs. 50 -61
https://doi.org/10.22320/07190700.2023.13.02.04

BUILDING SIMPLE MATHEMATICAL MODELS TO CALCULATE THE ENERGY REQUIREMENTS OF BUILDINGS[1]

Recibido 22/05/2023
Aceptado 16/11/2023

CONSTRUCCIÓN DE MODELOS MATEMÁTICOS SIMPLES PARA EL CÁLCULO DEL REQUERIMIENTO ENERGÉTICO DE EDIFICACIONES

CONSTRUÇÃO DE MODELOS MATEMÁTICOS SIMPLES PARA CALCULAR OS REQUISITOS DE ENERGIA DE EDIFICAÇÕES

María Victoria Mercado
Doctora en Ciencias
Investigadora Adjunta
Instituto de Ambiente, Hábitat y Energía (INAHE) CONICET, Mendoza, Argentina
https://orcid.org/0000-0003-1471-3709
mvmercado@mendoza-conicet.gob.ar

Gustavo Javier Barea-Paci
Doctor en Ciencias
Investigador Adjunto
Instituto de Ambiente, Hábitat y Energía (INAHE) CONICET, Mendoza, Argentina
https://orcid.org/0000-0002-5643-3206
gbarea@mendoza-conicet.gob.ar

Andrés Esteban Aceña
Doctor rerum naturalium
Investigador Adjunto, Facultad de Ciencias Exactas y Naturales
Instituto Interdisciplinario de Ciencias Básicas CONICET-Universidad Nacional de Cuyo , Mendoza, Argentina
https://orcid.org/0000-0002-5790-8103
acena.andres@gmail.com

1 Project supported by CONICET PIP 0777. TRANSITION TO ENERGY-EFFICIENT BUILDINGS: Adjustment of the scale of certification ranges (A-G) applicable to the National Housing Labeling Program by means of predictive metamodeling with AI

RESUMEN

El presente trabajo tiene por objetivo desarrollar un modelo matemático predictivo que otorgue un primer acercamiento al valor de requerimiento energético (RE) de un edificio en un clima templado continental, con el propósito de aportar al conocimiento teórico sobre herramientas de evaluación energética. Se realizaron simulaciones paramétricas procesadas con los programas *EnergyPlus* 9.5 y *JePlus*. Los resultados fueron utilizados como *Dataset* para el armado de diferentes modelos matemáticos, para los cuales se utilizó el programa SageMath a fin de desarrollar ecuaciones que predigan el RE de cada escenario. Se trabajó con modelos escalonando su complejidad en cuanto a métodos utilizados y cantidad de parámetros. Se seleccionó un modelo con bajo nivel de error (0.08) y 15 parámetros. Se advirtió que, si bien el aumentar la cantidad de parámetros acercaba los modelos al error 0.02, se corría el peligro de *overfitting*. El modelo seleccionado busca incorporar la precisión y validez de las simulaciones dinámicas a una herramienta de predicción sencilla y aplicable por profesionales de la construcción.

Palabras clave
modelo matemático, simulaciones, arquitectura sustentable

ABSTRACT

This work looks to build a predictive mathematical model that can provide a first approach to a building's energy requirement (ER) value in a temperate continental climate. The aim is to contribute to the theoretical knowledge of energy assessment tools. To do this, parametric simulations were run and processed using the EnergyPlus 9.5 and JePlus programs. The results were then used as a dataset to build different mathematical models, using the SageMath program to run equations that predicted the ER of each scenario. Work was done with the models, scaling their complexity with the methods and the number of parameters used. Finally, a model with a low error (0.08) and 15 parameters was chosen. It was noted that, although increasing the number of parameters brought the models closer to a 0.02 error, there was a risk of overfitting. The chosen model seeks to incorporate dynamic simulations' accuracy and validity into a simple prediction tool that construction professionals can apply.

Keywords
mathematical modeling, simulations, sustainable architecture

RESUMO

O objetivo deste trabalho é desenvolver um modelo matemático preditivo que possibilite uma primeira abordagem do valor dos requisitos de energia (ER) de um edifício em um clima continental temperado, de forma a contribuir para o conhecimento teórico das ferramentas de avaliação energética. As simulações paramétricas foram realizadas e processadas com os softwares EnergyPlus 9.5 e JePlus. Os resultados foram utilizados como Dataset para a construção de diferentes modelos matemáticos, para os quais foi utilizado o programa SageMath para desenvolver equações que preveem o ER de cada cenário. Trabalhamos com modelos que escalonam sua complexidade em termos de métodos utilizados e número de parâmetros. Foi selecionado um modelo com baixo nível de erro (0,08) e 15 parâmetros. Observou-se que, embora o aumento do número de parâmetros tenha aproximado os modelos ao erro de 0,02, havia o risco de sobreajuste. O modelo selecionado busca incorporar a precisão e a validade das simulações dinâmicas em uma ferramenta de previsão simples que pode ser aplicada por profissionais da construção.

Palavras-chave:
modelo matemático, simulações, arquitetura sustentável.

INTRODUCTION

The construction sector contributes significantly to global energy demand. The energy intensity of buildings has stayed the same in recent years, remaining at 150kWh/m^2. According to the estimates of the International Energy Agency (IEA), to achieve "net zero emissions," it is necessary that the intensity decreases by approximately 35% compared to current levels and remains around 95 kWh/m^2 (International Energy Agency, 2022). Unfortunately, this has remained virtually unchanged since 2019 (United Nations Environment Programme, 2022).

As the global population continues to grow, an increase in the energy demand from buildings is expected. One strategy to mitigate this situation is optimizing their energy efficiency, which can be addressed in their design, construction, and operation stages. Therefore, it is essential to have accurate forecasts of energy requirements, as it is becoming crucial to achieve significant energy savings in the construction sector (Chang et al., 2019).

Timuçin and Wilde (2021) warn that, when designing, more attention should be paid to the holistic investigation of all factors to achieve energy efficiency. To make this possible, it is necessary to consider a series of variables that influence energy consumption and user comfort, such as building orientation, envelope thermal quality, the relationship between opaque and translucent surfaces, and building shape.

Nowadays, professionals frequently resort to computational modeling and simulation (BPS or Building Performance Simulation) to evaluate and analyze different design and operation strategies. This is because the effectiveness of BPS has been documented in the literature and has been used in a wide range of applications (Azar et al., 2021; Raj et al., 2021; Schwartz & Raslan, 2013).

Both modeling and computational simulation are done before the construction or remodeling of a building, considering the variables mentioned above and the outdoor climatic conditions. The results obtained through this process are sufficient and accurate throughout a period and at the same frequency. These results can be the building's energy consumption, maximum loads, and indoor environmental conditions, among others (Seyedzadeh et al., 2019). However, this methodology usually requires many tests and lengthy periods (Papadopoulos et al., 2018) and demands a high level of expertise with powerful computing resources (Catalina et al., 2013). To overcome these limitations, researchers have begun to apply surrogate models that complement the capabilities of BPS. The process consists of training a mathematical model that mimics its performance and testing different building configurations at a low computational cost (Ye et al., 2019; Fang & Cho, 2019).

Substitute models allow users to predict the energy behavior of a building under different conditions. Some of the works that have used mathematical models to predict energy consumption include, for example, the use of neural network models or multiple linear regression models (Chou & Ngo, 2016; González-Vidal et al., 2017; Huang et al., 2021; Jiwon et al., 2022; Kwak et al., 2013; Zhao & Magoulès 2012). However, these are difficult for construction professionals to solve and access.

In this regard, this work aims to contribute to theoretical knowledge about energy assessment tools where it is unnecessary to resort to compelling but complex simulation environments. For this reason, the formulation of a simplified mathematical model based on simple morphological variables is proposed to predictively calculate the annual energy requirement for a building's air conditioning (REC). The argument is that simplifying mathematical models for initial energy assessments in buildings is currently subject to stationary thermal-energy balances, implying a gap between dynamic realities and the answers the BPS can provide. Due to this, the model sought aims to capture the variability of the thermal-energy-dynamic balance using the *EnergyPlus* program, given that the Dataset is built with the parametric simulations this makes. This simplified approach can offer construction professionals a practical and accessible tool for making energy-efficiency decisions.

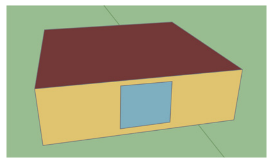
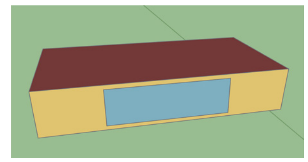
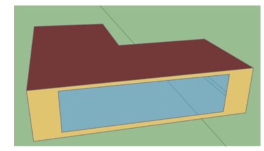
Finally, it should be noted that the context where this model will be applied will be in buildings located in a temperate continental climate environment, specifically in the Mendoza region, Argentina.

METHODOLOGY

The study's work method is applied and divided into three sequential stages: the first is the parametric simulation, the second is creating the Dataset, and

Revista Hábitat Sustentable Vol. 13, N°. 2. ISSN 0719 - 0700 / Págs. 50 -61
https://doi.org/10.22320/07190700.2023.13.02.04

Table 1. Variables considered for the study. Source: Preparation by the authors.

Discrete variable			
	C - square	R - rectangular	L - corner shape
Shape typology			
	Continuous variables		
v- WWR Factor (opaque openings and wall ratio)	20 - 30 - 40 - 50 – 60 – 70 – 80 – 90 - 100		
w- orientation	0 – 15 – 30 – 45 – 60 – 75 – 90 – 105 – 120 – 135 – 150 – 165 – 180 – 195 – 210 -225 – 240 – 255 – 270 – 285 – 300 – 315 – 330 – 345 – 360		
x-Wall - EPS insulation thickness (expanded polystyrene)	0.01 – 0.02 – 0.03 – 0.04 – 0.05 - 0.06 - 0.07 – 0.08 – 0.09 – 0.10 – 0.11 - 0.12 – 0.13 – 0.14 – 0.15 - 0.16 - 0.17 – 0.18 – 0.19 – 0.20		
y-Roof - EPS insulation thickness	0.01 – 0.02 – 0.03 – 0.04 – 0.05 - 0.06 - 0.07 – 0.08 – 0.09 – 0.10 – 0.11 - 0.12 – 0.13 – 0.14 – 0.15 - 0.16 - 0.17 – 0.18 – 0.19 – 0.20		

the third covers constructing the mathematical models.

Each of these consists of sub-stages, which are detailed below. The first parametric simulation stage is based on identifying the entry variables and their ranges to be used as inputs in the simulation model; the second is performing the parametric simulations for the climate of the city of Mendoza, establishing the data of air-conditioning energy requirements as an output.

On the other hand, the sub-stages in the Dataset creation are from processing the data obtained in the simulations with the training data. Finally, once this stage is completed, the third mathematical model creation stage begins, namely their elaboration process. Each of the stages and sub-stages is explained in detail below.

PARAMETRIC SIMULATION

Nowadays, computer simulations have strict validation in the studies and analysis of light, thermal, and energy behaviors, among others, for building projects or already built buildings (Malkawi, 2004). Therefore, the simulation data are considered realistic of the REC values of the building forms used. The parametric simulation methodology allows the systematic combination of all the variables in the same simulation procedure, as it simplifies the actions of the simulator user to execute and program the interaction of the variables one by one in individual simulations.

VARIABLES AND RANGES

INPUTS, Input data of the simulation models

The following variables were used as study variables: the shape, the orientation, the opaque-transparent envelope ratio, and the transmittance values of walls and roofs. Table 1 presents the categorized variables, the ranges used, and their subsequent denomination in the mathematical models.

Variables are divided into discrete and continuous. As discrete variables, the building typologies of houses are determined considering their shape:

square, rectangular, and L, exemplifying different compactness indices (CI), 88.6%, 82.6%, and 75.5%, respectively.

PARAMETRIC COMPUTATIONAL SIMULATIONS

In this phase, the *EnergyPlus* 9.2 and *Jeplus* programs were used to make parametric simulations. The parametric methodology is an exhaustive method that allows evaluating the cross-combination and interrelation of numerous values of ranges entered as inputs, changing one at a time. This results in a total of 270,000 simulations.

The three formal building typologies were modeled. With these models, the parametric simulations were made by modifying the input variables of Table 1. The models worked with a single thermal zone of 80m^2. In the materialization, the traditional and mass construction system of the City of Mendoza was taken as a reference, using for walls: mass ceramic brick, plaster on both sides, and thermal insulation; and for roofs: lightened concrete slabs, mortar for slopes, asphalt membrane, and thermal insulation. The variability of the construction packages was worked on by modifying the thermal insulation thicknesses, using EPS (expanded polystyrene) in both cases.

As an output, the Zone Ideal Loads Heating Energy and Zone Ideal Loads Cooling Energy outputs of the *EnergyPlus* program were used, which provide the heating and cooling energy requirement values. These values are added together to obtain the annual air conditioning total. Using this, thermostats related to the Olgyay comfort ranges were established, i.e., for winter, it was set at 20°C, and for summer, 24°C. The ideal load RECs were used, so specific HVAC systems were not considered.

CONSTRUCTION OF THE DATASET

To set up the Dataset, the results of the simulations of the previous stage were used, whose data were used to train the mathematical models. Unlike what is done with artificial neural network methodologies, in this study, to optimize the models and lower error values, the sample was not divided into a training group and a testing group, as obtaining more data would lead to an increase in the training sample.

The Dataset consists of two columns. The first has the case code name that represents the change of the value of each variable to combine the differences *one at a time* among all of them. The second column has the REC value obtained by the simulation. The outline built is presented in Table 2.

Table 2. Dataset outline built from parametric simulations. Source: Preparation by the authors based on the results.

CASE NAME	REC VALUE [Kw]
EP_G_Pv_0_Pw_0_Px_0_Py_0	x1
EP_G_ Pv_0_Pw_0_Px_0_Py_1	x2
EP_G_ Pv_0_Pw_0_Px_0_Py_2	x3
EP_G_ Pv_0_Pw_0_Px_0_Py_nx	xN
EP_G_ Pv_1_Pw_3_Px_7_Py_0	x5
EP_G_ Pv_1_Pw_3_Px_7_Py_1	x6
EP_G_ Pv_1_Pw_3_Px_7_Py_2	x7
EP_G_ Pv_ nx _Pw_ nx _Px_ nx _Py_ nx	xN

Where:
EP: EnergyPlus;
G: Climate file;
Pv: WRR Variable;
Pw: Orientation variable;
Px: WALL insulation thickness variable;
Py: ROOF insulation thickness variable;
Nx: number of times q changes the value of the range of each variable until the combination of all of them is completed, in this case, the 270,000 cases;
x1-x2-xN: REC values obtained from the simulations for each case.

CONSTRUCTION OF MATHEMATICAL MODELS

Development of mathematical models

Once the Dataset outline stage is over, the third mathematical model development stage begins. For this, the information synthesized in the Dataset was input into the *SageMath* software, whose role is to build mathematical models. In the process, a balance was sought between obtaining a fine model and the statistical model that approached the reference value, which was taken from the results obtained from the simulations (REC). In this way, each mathematical model was developed as an equation where the independent variables are the continuous variables (v: WWR factor, w: orientation, x: Wall-EPS insulation thickness, y: Roof-EPS insulation thickness). The least squares fit sets the parameters of the equations using computational simulations. Once the parameters have been determined in the equation, it can be used to predict the building's REC by replacing the continuous variables with the value of the building in question.

Table 3. Types of models tested and fit measures for each. Source: Preparation by the authors

MODELS	ERROR MEASUREMENTS								
	Square shape			Rectangular shape			L-shape		
	Min.	Max.	STD	Min.	Max.	STD	Min.	Max.	STD
L	-0.123	0.127	0.047	-0.213	0.296	0.099	-0.26	0.224	0.083
Lin	-0.89	0.097	0.033	-0.121	0.219	0.06	-0.199	0.153	0.062
CC/C	0.05	0.064	0.018	-0.094	0.126	0.042	-0.078	0.116	0.033
Cs/C	-0.081	0.116	0.035	-0.154	0.177	0.053	-0.183	0.211	0.063
CIn c/C	-0.044	0.037	0.009	-0.063	0.067	0.022	-0.046	0.061	0.016
CIn s/C	-0.081	0.107	0.032	-0.136	0.154	0.042	-0.17	0.196	0.059
CIn c/R	-0.07	0.056	0.018	-0.086	0.093	0.032	-0.071	0.071	0.023
CIn c/r2	-0.067	0.058	0.018	-0.077	0.095	0.031	-0.074	0.063	0.023
F2t	-0.038	0.031	0.013	-0.101	0.168	0.058	-0.05	0.076	0.022
F3t	-0.03	0.23	0.011	-0.048	0.044	0.018	-0.029	0.053	0.011
F3t/S	-0.036	0.028	0.012	-0.069	0.062	0.032	-0.039	0.061	0.015
F4t	-0.026	0.02	0.006	-0.052	0.015	0.007	-0.028	0.047	0.008

A total of 40 mathematical models were made, trying not to make them unnecessarily complicated. An exploration of options was carried out that began with linear models, which yielded high errors of around 29.6%. For reference, the more parameters a model has, the fewer errors it should have. However, this can be risky if "overfitting" is produced. To avoid this, models with errors that did not decrease substantially by increasing the number of parameters were preferred. Consequently, the improvements were made with quadratic models. In addition, orientation was highlighted as an angular variable, which led to the need to upgrade to trigonometric models.

RESULTS

Different results obtained from the models that allowed an approximation to a simple generic equation were explored. These were analyzed in two senses: one, the equation, its form and development in terms of quantity and representativeness of the parameters involved, and the other, from the error as a diagnostic object of the predictive value and the effectiveness of the model.

Evaluating the relative minimum and maximum errors was considered a goodness-of-fit measure. The tested models and the error measurements compared to each shape are presented in Table 3. The standard deviation value is also presented as a goodness-of-fit measure.

CC/C Quadratic with cross-terms
Cs/C Quadratic without cross-terms
L Linear
CInc/C Inverse quadratic with cross-terms
CIns/C Inverse quadratic without cross-terms
Lin Inverse linear
F4t Fourier 4 terms ⍰ looks like overfitting
F3t Fourier 3 terms
F3t/S Fourier 3 symmetric terms
F2t Fourier 2 terms
CIn c/R Inverse quadratic with clipping
CIn c/R^2 Inverse quadratic with clipping V2

The models were divided by family, linked to the assumed hypotheses. Of the 40 models explored, three stand out due to the following reasons:

a. The errors in the predictions obtained are limited;
b. The number of parameters does not lead to overfitting and
c. They allow for analyzing energy consumption behavior regarding the included variables.

The three models that were considered optimal are presented below:

Model 1 [M1] (model 01, Linear Fourier 2 even terms) (Equation 1)

$$Co(v,w,x,y) = F + F_v v + F_x \frac{1}{x} + F_y \frac{1}{y} \qquad \text{(Equation 1)}$$

Where:

$$F_i = A_i + B_i cos\left(\frac{\pi}{180}w\right)$$

This model has 8 parameters (A, B, Av, BV, Ax, Bx, Ay, and By). The values estimated for each shape are presented in Table 4.

Table 4. Values of the estimates for the parameters of Model 1. Source: Preparation by the authors.

Model 1	C	L	R
A	10633.1	2190.61	5093.01
B	-128.62	-250.7	-442.29
Av	-17.49	-4.49	-5.33
Bv	-14.67	-6.56	-12.09
Ax	14.76	10.7	12.97
Bx	1.79	1.49	2.34
Ay	2.09	4.97	3.9
By	5.94	1.96	3.75

Model 2 [M2] (quadratic angular model 30 even terms) (Equation 2)

$$Co(v,w,x,y) = F + F_w cos\left(\frac{\pi}{180}w\right) + F_v v + F_x\frac{1}{x} + F_y\frac{1}{y}$$
$$+F_{ww} cos^2\left(\frac{\pi}{180}w\right) + F_{vv}v^2 + F_{xx}\frac{1}{x^2} + F_{yy}\frac{1}{y^2}$$
$$+F_{wv} v cos\left(\frac{\pi}{180}w\right) + F_{wx} cos\left(\frac{\pi}{180}w\right)\frac{1}{x} + F_{wy} cos\left(\frac{\pi}{180}w\right)\frac{1}{y}$$
$$+F_{vx}v\frac{1}{x} + F_{vy}v\frac{1}{y} + F_{xy}\frac{1}{xy}$$

(Equation 2)

This model uses 15 parameters. The fit values thereof are presented in Table 5.

Table 5. Values of the estimates of the parameters of the Model 2. Source: Preparation by the authors.

Model 2	C	L	R
F	11226.94	2459.96	5834.87
Fw	-128.62	-250.7	-442.29
Fv	-35.09	-14.09	-29.29
Fx	5.28	8.44	6.76
Fy	-6.28	4.61	2.91
Fww	40.69	-8.89	-148.99
Fvv	0.11	0.06	0.17
Fxx	-0.02	-0.03	-0.05
Fyy	0.01	-0.03	-0.04
Fwv	-14.67	-6.56	-12.09
Fwx	1.79	1.49	2.34
Fwy	5.94	1.96	3.75
Fvx	0.16	0.07	0.17
Fvy	0.09	0.04	0.06
Fxy	0.12	0.07	0.05

Model 3 [M3] (model 15, Quadratic Fourier, 2 even terms) (Equation 3)

$$Co(v,w,x,y) = F + F_v v + F_x\frac{1}{x} + F_y\frac{1}{y}$$
$$+F_{vv}v^2 + F_{xx}\frac{1}{x^2} + F_{yy}\frac{1}{y^2} + F_{vx}\frac{v}{x} + F_{vy}\frac{v}{y} + F_{xy}\frac{v}{xy}$$

(Equation 3)

Where (Equation 4):

$$F_i = A_i + B_i cos\left(\frac{\pi}{180}w\right)$$

(Equation 4)

The values of the parameters are presented in Table 6. In this case, a total of 20 is used.

Table 6. Values of the estimates of Model 3's parameters
Source: Preparation by the authors.

Model 3	C	L	R
A	11247.28	2455.51	5760.38
B	-197.85	-105.05	-16.54
Av	-35.09	-14.09	-29.29
Bv	-13.03	-13.39	-29.96
Ax	5.28	8.44	6.76
Bx	2.33	2.29	2.32
Ay	-6.28	4.61	2.91
By	8.97	3.49	5.35
Avv	0.11	0.06	0.17
Bvv	-0.03	0.05	0.14
Axx	-0.02	-0.03	-0.05
Bxx	-0.02	-0.01	-0.01
Ayy	0.01	-0.03	-0.04
Byy	-0.07	-0.02	-0.04
Avx	0.16	0.07	0.17
Bvx	0.03	0.01	0.03
Avy	0.09	0.04	0.06
Bvy	0.07	0.02	0.04
Axy	0.12	0.07	0.05
Bxy	-0.01	-0.01	-0.01

Revista Hábitat Sustentable Vol. 13, N°. 2. ISSN 0719 - 0700 / Págs. 50 -61
https://doi.org/10.22320/07190700.2023.13.02.04

Table 7. Min, Max, and Std absolute and relative error measurements for the square shape. Source: Preparation by the authors.

S-Shape	Absolute errors			Relative errors		
	Min	Max	Std	Min	Max	Std
M1	-1693	726	183	-0.14	0.07	0.02
M2	-801	897	125	-0.06	0.08	0.01
M3	-521	899	118	-0.05	0.08	0.01

Table 8. Min, Max, and Std absolute and relative error measurements for the rectangle shape. Source: Preparation by the authors.

R-Shape	Absolute errors			Relative errors		
	Min	Max	Std	Min	Max	Std
M1	-1304	735	189	-0.19	0.17	0.04
M2	-551	1038	110	-0.08	0.14	0.02
M3	-597	980	104	-0.09	0.18	0.02

Table 9. Min, Max, and Std absolute and relative error measurements for the corner shape. Source: Preparation by the authors.

L-Shape	Absolute errors			Relative errors		
	Min	Max	Std	Min	Max	Std
M1	-875	352	103	-0.21	0.17	0.05
M2	-362	564	72	-0.16	0.14	0.04
M3	-347	538	67	-0.14	0.14	0.03

The predictive value of each model, the absolute minimum and maximum errors, relative errors, and the standard deviation (Std) between the mathematical model and the prediction of the simulations are presented in Table 7 for the S-shape and Tables 8 and Table 9 for L and R.

These analyses show that, in the S-Shape with the M2, the REC value calculated by the model may have a relative error of 1.2%. The minimum and maximum error values are the percentage of error that can occur when the model's equation gives a lower (Min error) or a higher consumption (Max error) compared to the reference. Following the example of M2, the error can vary by 6%, giving a lower value of REC, and by 8%, giving a higher value of REC.

USE OF THE MODEL IN A DESIGN EXAMPLE

For the following example, a housing project with the characteristics presented in Table 12 was assumed.

Table 10. First variables assumed for a housing project, Case n. Source: Preparation by the authors.

Shape	Square	Rectangular	L
v -WWR	40%	40%	40%
w - orientation	0 (North)	0 (North)	0 (North)
x – wall insulation thickness	0.01 (considering a wall without insulation)	0.01 (considering a wall without insulation)	0.01 (considering a wall without insulation)
y – roof insulation thickness	0.05	0.05	0.05

The REC is obtained with the variables of Table 10 and using Model 1 (considering it as optimal). For this, the first step is to obtain the values of F (Equation 5), Fv (Equation 6), Fx (Equation 7), and Fy (Equation 8) for w=0:

$$F = A + Bcos\left(\frac{\pi}{180}w\right) = 2190.6 - 250.7cos\left(\frac{\pi}{180}0\right) = 1939.9 \quad \text{(Equation 5)}$$

$$F_v = A_v + B_v cos\left(\frac{\pi}{180}w\right) = -4.49 - 6.56cos\left(\frac{\pi}{180}0\right) = -11.05 \quad \text{(Equation 6)}$$

$$F_x = A_x + B_x cos\left(\frac{\pi}{180}w\right) = 10.7 + 1.49cos\left(\frac{\pi}{180}0\right) = 12.19 \quad \text{(Equation 7)}$$

$$F_x = A_x + B_x cos\left(\frac{\pi}{180}w\right) = 10.7 + 1.49cos\left(\frac{\pi}{180}0\right) = 12.19 \quad \text{(Equation 8)}$$

With these values, this is replaced in the model's equation (Equation 9):

$$Co(v,w,x,y) = F + F_v v + F_x \frac{1}{x} + F_y \frac{1}{y} = 1939.9 - 11.05 \times 40 + 12.19\frac{1}{0.01} + 6.93\frac{1}{0.05} = 2855.5$$

(Equation 9)

The values obtained with the mathematical models are compared with the results of the simulations' Dataset to evaluate the margin of error. These values are presented in Table 11, where it is observed that the errors are within the thresholds presented in Table 7, Table 8, and Table 9, where the Max for the square shape is 7% and 17% for the rectangular and L shapes.

Table 11. Results of the air conditioning energy requirement [KW/m2] and relative error. Source: Preparation by the authors.

	Original Square	Original Rectangle	Original L
	Original	Original	Original
Mathematical model	11034	5638	2855
Computational simulation	10646.97	5067.73	2656.52
Relative error	4%	11%	8%

Table 12. Variables corrected to improve the Case n project. Source: Preparation by the authors.

Shape	Square	Rectangular	L
v -WWR	20%	30%	30%
w - orientation	0 (North)	0 (North)	0 (North)
x - wall insulation thickness	0.05	0.05	0.1
y - roof insulation thickness	0.1	0.1	0.1

Table 13. Results of the air conditioning energy requirement [KW/m2] and energy saving percentage. Source: Preparation by the authors.

	Square		Rectangular		L	
	Original	Improved	Original	Improved	Original	Improved
Mathematical model result	11034	10273	5638	4510	2855	1800
REC Reduction		7%		20%		37 %

The values allow considering changes and improvements in the project, such as increasing insulation in walls and ceilings, leading to the values of x=0.1 and y=0.1, and reducing the proportion of openings to 20% in the S shape and 30% in the R and L shapes. (Table 12).

When calculating with the mathematical model, improvements in reducing the energy requirement are observed. These results are presented in Table 13.

Energy reduction and savings are differentiated between building shapes. Up to 37% improvement is achieved in the L shape, 20% for the rectangular shape, and less than 7% for the square shape. This allows differentiating improvement and bioclimatic design strategies for the building's different geometric configurations. In addition, it reveals the importance and potential of the mathematical models.

DISCUSSION

Mathematical modeling is advantageous for predicting the REC. Although using 20 parameters may seem excessive, this contrasts with the time, resources, and expertise needed to perform the 270,000 simulations that fed the model.

In the error analysis, the restricted relative error below 2.1% is considered acceptable. Regarding the number of parameters, a substantial improvement is obtained by moving from M1 (8 parameters) to M2 (15 parameters). However, when moving from M2 to M3 (20 parameters), the same relevance is not observed in the model's improvement, so it can be noted that considering models with more than 20 parameters does not have predictive advantages and runs the risk of overfitting.

As for the models not presented, it is important to emphasize that choosing the model's functional shape strongly affects predictive ability. That is, instead of considering 1/x and 1/y as variables, x and y will be taken into account directly, and the errors grow substantially. The same occurs for the angular dependence in *w*. A "*goodness of fit*" measurement is also made for all models on the calculated values, using the *chi-square over degrees of freedom (x^2/d.o.f.)* methodology. This allows evaluating whether the errors are randomly distributed compared to the prediction, in terms that, if the model fits well, the value of x^2/dof should be the closest to 1.

Table 14 shows models that deviate considerably from the reference measurement (1) and others that are reasonable, although they are not perfect. However, since the Dataset is very numerous and the parameters are few, one can opt for the path of making more models with a greater number of parameters and, in this way, ameliorate the x^2/dof. The problem is that the models become more complex, losing their simplicity.

The improvement observed in moving from model 1 to model 2 shows that switching to non-linear

Table 14. Chi-square over degrees of freedom values. Source: Preparation by the authors.

Model	No of Parameters	Shape		
		C	L	R
01 Linear Fourier 2 even terms [M1]	8	3.4	5	7.4
02 Linear Fourier 2 odd terms	8	51.6	93.5	130.4
03 Linear Fourier 3 terms	12	3.2	5	7.2
04 Linear Fourier 3 even terms	12	3.3.	4.5	6.3
05 Linear Fourier 3 odd terms	12	51.6	93.3	130.4
06 Linear Fourier 4 terms	16	3.2	4.8	6.1
07 Linear Fourier 5 terms	20	3.2	4.3	6.1
08 Fourier Quad STC 2 even terms	14	2.8	3.4	6.1
09 Fourier Quad STC 2 odd terms	14	51.1	92.6	128,5
10 Fourier Quad STC 3 terms	21	2.7	3.4	3.9
11 Fourier Quad STC 3 even terms	21	2.7	3.4	3.9
12 Fourier Quad STC 3 odd terms	21	51	92.4	128.4
13 Fourier Quad STC 4 terms	28	2.6	3.1	2.7
14 Fourier Cuad STC 5 terms	35	2.6	2.7	2.7
15 Fourier Quad 2 even terms [M3]	20	1.4	2.2.	2.1
16 Fourier Quad 2 odd terms	20	1.4	2.2	2.1
17 Fourier Quad 3 terms	30	1.3	2.1	1.9
18 Fourier Quad 3 even terms	30	1.3	1.7	0.9
19 Fourier Quad 3 odd terms	30	49.8	91.4	126.9
20 Fourier Quad 4 terms	40	1.2	1.8	0.7
21 Fourier Quad 5 terms	50	1.2	1.4	0.7
22 Fourier Linear Extra 2 even terms	12	3.3	4.8	7.2
23 Fourier Linear Extra 2 odd terms	12	51.6	93.4	130.3
24 Fourier QuadExtra V1 2 even terms	20	1.4	2.4	2.9
25 Fourier QuadExtra V1 2 even terms	20	49.8	91.7	127.5
26 Fourier Quad lExtra V2 2 even terms	20	1.4	2.4	2.2
27 Fourier Quad Linear Extra V2 2 odd terms	20	49.9	91.8	127
28 Angular Quad STC even	9	11.2	11.5	15.5
29 Angular Quad STC odd	9	51	92.7	129.1
30 Angular Quad even [M2]	15	1.6	2.5	2.4
31 Angular Quad odd	15	49.8	91.7	127.4
35 Linear	5	56.8	58.4	135.6
36 Quadratic STC	9	16	49.4	24.2
37 Quadratic	15	14.5	44.5	22.2
38 Inverse linear	5	51.5	45.9	129.1
39 Inverse quadratic STC	9	13.9	45.1	21.5
40 Inverse quadratic	15	12.6	40.2	18.9

models is the right decision. However, it is necessary to emphasize that achieving reasonable predictions with the consequently less complex linear models is feasible. The use of non-linear models allows conclusions to be reached regarding the importance of the different variables, not only by themselves but also in synergy with the others. M2 is chosen to discuss this point. This seems to be the most successful model of the search carried out in this work.

The standardized M2 parameters, similar to those presented in Table 5, are presented in Table 15, considering the range of the associated variable. That is to say, the change that this variable produces in consumption when it passes from one end of its range to the other makes it possible to see the importance of each term in the REC clearly.

Table 15. Values of the estimates of Model 2's parameters. Source: Preparation by the authors.

Model 2	C	L	R
F	11227	2460	5835
Fw	-257	-501	-885
Fv	-3509	-1409	-2929
Fx	528	844	676
Fy	-628	461	291
Fww	41	-9	-149
Fvv	1099	631	1660
Fxx	-215	-350	-474
Fyy	89	-328	-358
Fwv	-2934	-1312	-2418
Fwx	357	298	468
Fwy	1188	391	750
Fvx	1560	740	1653
Fvy	887	389	596
Fxy	1217	662	477

The first thing observed is that a high degree of consumption is given by parameter F, namely the constant in the model that does not depend on any of the variables. It is also important that the lowest constant is obtained for the L-shape, which highlights that architectural design substantially impacts energy consumption. On this consumption basis, the angular dependence given by *Fw* and *Fww* is not substantial.

Remembering what happens with Fv, whose role contributes to the glazed surface proportion, is also necessary. This term is always important and, moreover, negative, indicating that an increase in the glazed surface leads to reduced consumption. However, it is noted that *Fvv* (glazed surface), in the quadratic term, leads to an increase in consumption, although considerably less than the savings if the term were linear.

Finally, analyzing individual variables shows that the roof and wall insulation thickness have a lower impact than other variables. This means that there is a minor impact on consumption. The most important crossover term is *Fwv*, which, moreover, is negative. The rest of the terms are of intermediate importance, which shows that, although there are factors to highlight, an oversimplification of the model and design considerations is detrimental.

CONCLUSIONS

This research addresses the study of representative equations to obtain the housing's REC value. When making constructive decisions, the energy requirement is one of several factors to consider. Within this analysis, having a predictive model such as this one simplifies decision-making, allowing decisions to be made on quantitative considerations.

Forty mathematical models were run based on the results of parametric simulations, with non-linear models considered more suitable for balancing complexity for non-specialist users with low error levels.

The models considered optimal showed that it is possible to approach a reference value simply. The case used to demonstrate this serves to verify the values using the simulations, considering these as plausible. The results show that these models correctly fit a first analysis and are the basis for accurate energy-efficient decisions in the first steps of the architectural project.

This work can be replicated in other regions of the country by changing the climate archive. The work comprises a diversity of total transmittance values of different walls and roofs, orientations, and WWR ratios. For future work, it is felt that it is both necessary and possible to reduce the data used as a Dataset by using sampling and categorization methodologies such as LHS (Latin Hyper Cube).

Author contributions: Conceptualization, M.V.M. and G.J.B.; Data Curation, A.A. and M.V.M.; Formal Analysis A.A. and M.V.M.; Funding Acquisition, M.V.M.; Research, M.V.M., G.J.B and A.A.; Methodology M.V.M.; Project management, M.V.M.; Resources; Software; Supervision; Validation, M.V.M., G.J.B; and A.A.; Visualization, M.V.M.; Writing - original draft, M.V.M.; Writing - proofreading and editing, M.V.M.; G.J.B and A.A.

BIBLIOGRAPHIC REFERENCES

Azar, E., Alaifan, B., Lin, M., Trepci, E., & Asmar, M. E. (2021). Drivers of energy consumption in Kuwaiti buildings: insights from a hybrid statistical and building performance simulation approach. *Energy Policy*, *150*, 112154. https://doi.org/10.1016/j.enpol.2021.112154

Cătălina, T., Iordache, V., & Caracaleanu, B. (2013). Multiple regression model for fast prediction of the heating energy demand. *Energy and Buildings*, *57*, 302–312. https://doi.org/10.1016/j.enbuild.2012.11.010

Chang, C. C., Shi, W., Mehta, P., & Dauwels, J. (2019). Life cycle energy assessment of university buildings in tropical climate. *Journal of Cleaner Production*, *239*, 117930. https://doi.org/10.1016/j.jclepro.2019.117930

Chou, J., & Ngo, N. (2016). Time series analytics using sliding window metaheuristic optimization-based machine learning system for identifying building energy consumption patterns. *Applied Energy*, *177*, 751–770. https://doi.org/10.1016/j.apenergy.2016.05.074

González-Vidal, A., Ramallo-González, A. P., Terroso-Sáenz, F., & Skarmeta, A. F. (2017). Data driven modeling for energy consumption prediction in smart buildings. *2017 IEEE International Conference on Big Data (Big Data)*. https://doi.org/10.1109/bigdata.2017.8258499

Huang, S., Lin, Y., Chinde, V., Ma, X., & Lian, J. (2021). Simulation-based performance evaluation of model predictive control for building energy systems. *Applied Energy*, *281*, 116027. https://doi.org/10.1016/j.apenergy.2020.116027

INTERNATIONAL ENERGY AGENCY (2022). Tracking Buildings 2022. Paris: International Energy Agency. https://www.iea.org/energy-system/buildings

Kwak, Y. H., Seo, D., Jang, C., & Huh, J. (2013). Feasibility study on a novel methodology for short-term real-time energy demand prediction using weather forecasting data. *Energy and Buildings*, *57*, 250–260. https://doi.org/10.1016/j.enbuild.2012.10.041

Malkawi, A., & Augenbroe, G. (2004). *Advanced Building Simulation*. Routledge. https://doi.org/10.4324/9780203073674

Papadopoulos, S., Azar, E., Woon, W. L., & Kontokosta, C. E. (2017). Evaluation of tree-based ensemble learning algorithms for building energy performance estimation *Journal of Building Performance Simulation*, *11*(3), 322–332. https://doi.org/10.1080/19401493.2017.1354919

Raj, B. P., Meena, C. S., Agarwal, N., Saini, L., Khahro, S. H., Subramaniam, U., & Ghosh, A. (2021). A Review on Numerical Approach to Achieve Building Energy Efficiency for Energy, Economy and Environment (3E) benefit. *Energies*, *14*(15), 4487. https://doi.org/10.3390/en14154487

Schwartz, Y., & Raslan, R. (2013). Variations in results of building energy simulation tools, and their impact on BREEAM and LEED ratings: A case study. *Energy and Buildings*, *62*, 350–359. https://doi.org/10.1016/j.enbuild.2013.03.022

Seyedzadeh, S., Rahimian, F. P., Rastogi, P., & Glesk, I. (2019). Tuning machine learning models for prediction of building energy loads. *Sustainable Cities and Society*, *47*, 101484. https://doi.org/10.1016/j.scs.2019.101484

Harputlugil, T., & De Wilde, P. (2021). The interaction between humans and buildings for energy efficiency: A critical review. *Energy Research & Social Science*, *71*, 101828. https://doi.org/10.1016/j.erss.2020.101828

UNITED NATIONS ENVIRONMENT PROGRAMME (2022). *2022 Global Status Report for Buildings and Construction: Towards a Zero emission, Efficient and Resilient Buildings and Construction Sector*. https://www.unep.org/resources/publication/2022-global-status-report-buildings-and-construction

Fang, Y., & Cho, S. (2019). Design optimization of building geometry and fenestration for daylighting and energy performance. *Solar Energy*, *191*, 7–18. https://doi.org/10.1016/j.solener.2019.08.039

Ye, Y., Hinkelman, K., Zhang, J., Zuo, W., & Wang, G. (2019). A methodology to create prototypical building energy models for existing buildings: A case study on U.S. religious worship buildings. *Energy and Buildings*, *194*, 351–365. https://doi.org/10.1016/j.enbuild.2019.04.037

Zhao, H., & Magoulès, F. (2012). A review on the prediction of building energy consumption. *Renewable & Sustainable Energy Reviews*, *16*(6), 3586–3592. https://doi.org/10.1016/j.rser.2012.02.049