Assessing the baseline energy behaviour of the national single-family building stock: a parametrized modelling approach

Fabien Rouault[a], Constanza Molina[b], Claudia Valderrama-Ulloa[c]

[a] Independant Researcher, Rennes, France

[b] Escuela de Construcción Civil, Facultad de Ingeniería, Pontificia Universidad Católica de Chile, Santiago, Chile

[c] Centro de Investigación en Tecnologías para la Sociedad, Facultad de Ingeniería, Universidad del Desarrollo, Santiago, Chile

Abstract

Building stock energy modeling (BSEM) has become an essential tool for policymakers aiming to achieve carbon-neutral built heritage and adapt to climate change. This study proposes a parameterized energy model developed using Rhino/Grasshopper/Ladybug software to create a BSEM set that accurately represents the current Chilean single-family housing stock. The methodology incorporates 29 different building archetypes and 27 stochastic variables, each defined by its probability distribution. The model is applied nationally and across each administrative region, resulting in the creation of 1,130 building energy models to ensure consistent results. Nationwide, space heating energy use is estimated to range from 11 to 850 kWh/m² (90% CI: 27-285 kWh), while total energy use is estimated between 569 and 53,209 kWh/year (90% CI: 1,400-20,173 kWh/year), aligning well with empirical data. The most influential factors affecting space heating energy use include heating control variables, internal mass, and air permeability; these should be prioritized in future research and national energy surveys. In the near future, these BSEMs can be utilized to evaluate the impacts of various climate change scenarios and occupant behaviour on indoor thermal comfort, as well as to explore retrofitting strategies for transitioning the built heritage to a low-carbon building stock.

Keywords: Latin Hypercube Sampling, EnergyPlus, Single-family building energy performance

1 Introduction

During the UN Climate Change Conference (COP21), almost all countries around the world reaffirmed their commitment to reducing greenhouse gas emissions to limit global warming to below 2ºC by 2050 compared to the pre-industrial period (1850-1900). According to the International Energy Agency's (IEA) report, the building sector accounted for 8% of global greenhouse gas (GHG) emissions in 2020 (2.9 GtCO2e out of a total of 33.9 GtCO2e) [1], with the majority being used for thermal comfort and domestic hot water. The common goal is to reach a carbon-neutral economy by 2050, aiming for a reduction of 1.2 Gt by 2030 and 0.3 Gt by 2050[1]. To achieve this goal, countries are required to implement new energy efficiency policies in buildings, ensuring that reduced energy needs are met by renewable energy sources while maintaining a comfortable indoor environment.

The first thermal regulation of building envelopes was the German standard DIN 4108 in 1952, which explicitly established thermal insulation requirements in terms of maximum U-values, the use of insulation materials, and window panels [2]. Since the first oil crisis in the 1970s, various countries have implemented thermal regulations for new buildings to improve their energy efficiency, such as France in 1974, Japan in 1979, and China in 1986. More recently, a new European Union directive [3] established that new buildings should ensure high energy efficiency to minimize energy use, and that the remaining energy needs should be covered by renewable sources, either from the grid (electricity or heat) or on-site. These buildings are called Net Zero

Energy Buildings (NZEBs). Thus, buildings must achieve enhanced thermal performance of the envelope, and incorporate innovative energy systems, while energy managers must ensure the proper functioning of these systems to maintain occupants' thermal comfort, all of which require a multidisciplinary approach [4].

Although NZEBs are considered a key element in achieving carbon neutrality for the building stock by 2050, existing buildings and historical structures must also be considered. Consequently, each country needs to establish a roadmap of energy programs to gradually improve the energy performance of both new and existing buildings over the next three decades. However, several studies highlight significant discrepancies between the predicted energy consumption calculated for new building programs and the actual energy use. For example, Burman *et al.* [5] compared the measured and predicted energy performance of a secondary school in England. Their analysis found that although their calibrated energy model deviated from the measured energy by only 3.1%, the European calculation method (EPBD) standard initially estimated one-third of the actual energy use. Moreover, Kelly *et al.* [6] emphasized the inability of the British Standard Assessment Procedure (SAP) to accurately estimate the actual energy performance of the housing stock. Nevertheless, discrepancies between the measured and predicted performance have been significantly reduced by incorporating in-situ measurements of U-values and air permeability of the building envelope [7], as well as accounting for occupants' behaviour [8,9].

Studies on the impact of occupant behaviour on the variability of energy consumption have also shown that a minor increase in expenses to enhance occupant comfort can lead to higher energy use, a phenomenon known as the *rebound effect* which is often neglected by energy intervention programs. For example, some authors estimate an increase of 0-30% in the predicted energy use for space heating, 0-50% for space cooling, 10-40% for domestic hot water, and 5-12% for illumination [10]. These variations occur because the availability of energy at a lower price can encourage occupants to improve their environmental comfort by adjusting the space heating setpoint, taking longer showers, or increasing the number and size of luminaires and appliances [10]. Conversely, a *prebound effect* is observed in countries where households consume less energy than estimated by international models and standards [11]. Therefore, governments need decision-making support tools to identify key factors in their current energy use profiles and to assess the potential impacts of new regulations and policies on future energy use and GHG emissions.

Consequently, modeling the existing building stock is crucial for achieving the stated objectives [12]. Langevin *et al.* [14] classify Building Stock Energy Modeling (BSEM) techniques into four categories based on their design (top-down or bottom-up) and their degree of transparency (white-box or black-box). In this classification, the use of a building energy simulation program, a physics-based approach, is categorized as a bottom-up white-box modeling technique. This approach can explicitly model the interaction between energy end-use and building characteristics and operations [13]. Conversely, statistical and machine learning models, which fall under top-down black-box modeling, are primarily based on historical data. As a result, they cannot assess the impact of unobserved input data, such as the implementation of new technology or new thermal regulations.

Considering the diversity of buildings and user profiles on a national scale, several BSEM approaches have implemented the use of representative buildings or archetypes to represent a building stock. For example, Korolija *et al.* [14] defined four building geometries to represent the UK office building stock, parameterizing input variables such as envelope U-values, ventilation and infiltration flow rates, air conditioning setpoints, and thermal loads with associated schedules. Subsequently, the authors generated 1,000 Building Energy Models (BEMs) using Latin hypercube sampling (LHS), considering uniform distributions for each input parameter. This process allowed them to present the distribution of predicted space cooling and heating demands.. Mata *et al.* [15] developed a Matlab-Simulink script to calculate the energy use and CO2 emissions of the national building stock. They validated their model using the BESTEST protocol and by comparing empirical data from two actual buildings. In the USA, the ResStock project generated 550,000 residential and 350,000 commercial models to represent 133 million and 1.8 million buildings, respectively [16]. The authors used input data from different states and cities and calibrated their models using energy metering data.

In Chile, the residential sector represented 7.9% of the total energy consumption in 2019, amounting to 54.7 of 692 TWh/year [17]. Chilean houses mainly use biomass or wood as an energy source (38%), followed by electricity (25%), liquefied petroleum gas (LPG) (23%), natural gas (12%), and coal (3%). Due to Chile's vast geographic extension from north to south, the country experiences a wide range of climate conditions. Consequently, national energy policies must consider this diversity in energy demand and thermal comfort requirements.

This study aims to propose a modeling and simulation framework to generate multiple Building Energy Models (BEMs) that represent the energy performance of a national housing stock. The modeling framework is presented in Section 2, and the simulation results of the generated BEMs are presented and compared to national data in Section 3. Although this approach is applied to the Chilean residential stock, it can be used for other building types, countries, and scales (city, region, country) if data are available to generate the inputs.

2 Method

BEMs are generated using a parameterized model created in Rhino 6/Grasshopper and the Ladybug Tools plugin. EnergyPlus is an open-source, cross-platform energy simulation software that models the physics of buildings and associated energy systems. First, a set of input data is obtained using the Latin hypercube sampling method, following Molina *et al.* [24]. Once generated and exported to an EnergyPlus-compatible format [18], the set of BEMs can be simulated and analyzed. These results show the baseline of the building stock's current thermal and energy performance.

2.1 Available input data

Molina *et al.* [19] proposed 496 archetypes to represent the Chilean residential building stock, along with four thresholds of 2, 8, 29, and 90 archetypes, representing 13%, 35%, 70%, and 90%, respectively. The archetypes were defined mainly using building permits from 1990 to 2016 and 2002 census data. They provide information about primary construction materials, floor area, the number of rooms, bedrooms, and bathrooms, and the number of occupants; thus, they are appropriate for describing the stock in this study.

The archetypes include two construction periods: pre- and post-thermal regulation of 2007, which defined maximum U-values for exterior walls, roofs, overhang floors, and window-to-wall ratios[20]. Given the reported discrepancy between prescribed and actual U-values [21,22], a probability density function for this parameter is implemented based on the literature; see Table 1.

The air permeability of the envelope of each house is described following Molina *et al.* [23] which provides two sets of normalized leakage (NL) distributions—one for each construction period or age of the building. A value of NL is sampled from the corresponding distribution according to the age, region number, and climate zone of the modeled house. The envelope leakage values are calculated by converting the NL values to Q50 values [$m^3\ h^{-1}\ m^{-2}$], following Sherman & Dickerhoff [24].

The national energy survey of Chilean households was conducted by the Corporación de Desarrollo Tecnológico (CDT) [25] and sponsored by the Ministry of Energy and the energy trade associations of electricity, natural gas, and liquefied petroleum gas suppliers. The survey comprised 160 questions and was applied to 3,500 households nationwide. The questionnaire focused on energy use habits, house characteristics, and socioeconomic status. Each household's energy use is broken down using historical energy records and statistical analysis, although the consultant does not provide further details. The results of this energy survey are used to define the distributions of the following inputs: 1) internal heat gains, 2) fuel for cooking, space heating, and domestic hot water, and 3) hours and months of space heating; see Table 1. The default entries include a floor height of 2.4 m and a roof angle of 25°.

Table 1: Input data of the energy model

Input	Unit	Range	Reference
Permeability	Q_{50}	eCDF[0.001; 7]	[26]
Heating setpoint	°C	U[18; 22]	-
Months of heating	-	eCDF[0; 12]	[27]
Hours of heating	-	eCDF[0; 24]	[27]
ΔU_{Wall}	W/m^2K	U[-0.15; 0.43]	[21,22]
ΔU_{Roof}	W/m^2K	U[-0.15; 0.27]	[21]
$U_{windows}$	W/m^2K	eCDF[2.8; 5.8]	[28]
$\Delta U_{windows}$	W/m^2K	U[0.03; 0.08]	[21,22]
Lighting loads	kWh/m^2year	eCDF[0.15; 38]	[27]
Appliance loads	W/m^2	eCDF[1.2; 185]	[27]
Domestic hot water by fuel type (4 types)	W/m^2	eCDF[1.3; 194]	[27]
Kitchen by fuel type (4 types)	W/m^2	eCDF[0.39; 69]	[27]
Form factor	-	U[1; 2]	-
Orientation	°	U[0; 179]	-
Glazing ratio	%	U[5; 32]	[20]
Thermal mass	kg/m^2	U[10; 100]	[29]
Flooring material	-	eCDF[carpet, cement tile, ceramic tile, vinyl, wood deck]	[30]
Wall material	-	eCDF[reinforced concrete, brick, concrete blocks, wood frame]	[30]
Roofing material	-	eCDF[Zinc, fiber cement board]	[30]

eCDF: empirical cumulative distribution function and min-max range. U: Uniform distribution and min-max range.

2.2 Stochastic generation of model inputs

Several sets of input data, including the archetypes' ID and physical properties of the buildings, were generated stochastically using Latin hypercube sampling (LHS) and the optimumLHS function in R, following the method described in Molina *et al.* [26]. This method is preferred over random sampling due to its optimized distribution of parameter values within a hypergrid, which requires a reduced number of simulations. Simulations are added until the convergence criteria are met; see Section 2.5. In this study, the parameter values are described using a known cumulative distribution function, computed using empirical data, or a value within a range according to the literature or the national building code, as appropriate; see Table 1.

2.3 Generation of Building energy model

BEM house models are parametrically generated using Rhino 6/Grasshopper [31] and the Ladybug Tools plugin [32]; see Figure 1. First, the generated matrix of input data is imported into the software. Figure 1 shows the Grasshopper workspace used to generate the BEM models based on the input data.

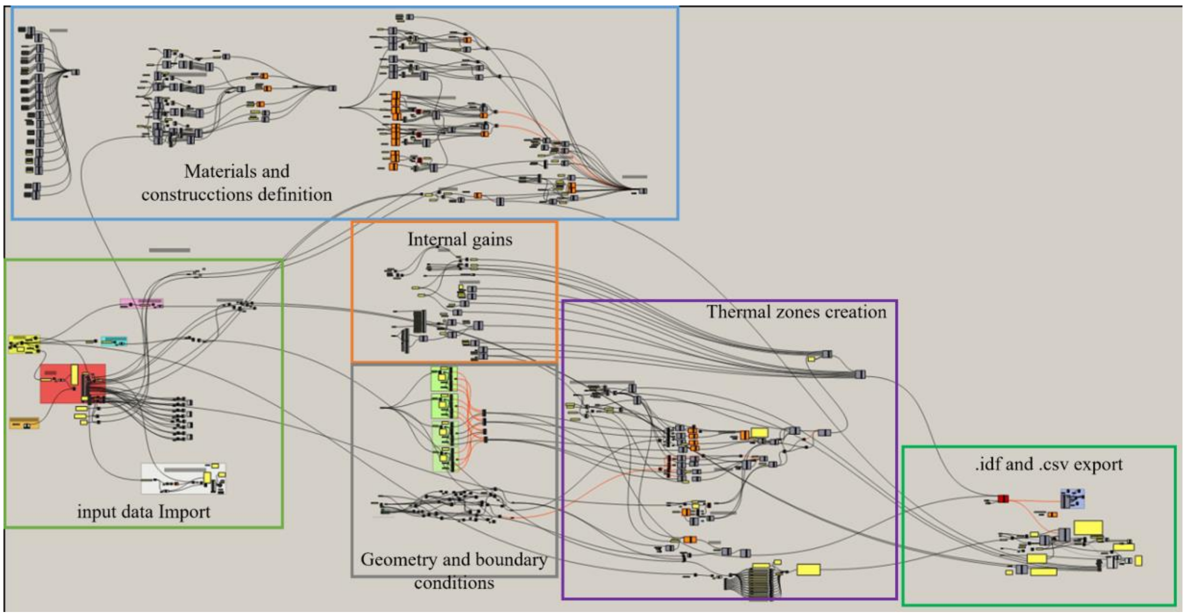


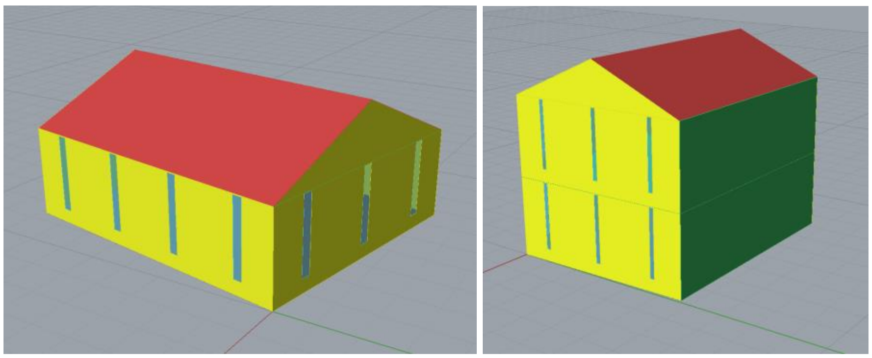
Figure 1: Screenshot of the Grasshopper workspace used to generate BEM models from input data.

For each observation, the U-values of the construction elements are adjusted according to the construction period, climatic zone, and associated variability. To fit the input U-value, the thickness of either the insulation or the structural material is modified. For glazing properties, the simple glazing material component is used, considering the input U-values (with an upper threshold of 5.8 W/m²K, defined by EnergyPlus [18]) and a default solar heat gain coefficient of 0.65.

The 3D geometry of the house is automatically generated in a rectangular floor plan based on the input data on floor area, form factor, and the number of storeys; see Figure 2. The geometry surfaces are then divided to assign the boundary conditions and construction materials. One or two walls are considered adiabatic for mid-terrace and terrace houses, respectively; see the green surface in Figure 2b. Although Molina *et al.* [19] proposed a number of rooms for each archetype, for simplicity, the present thermal models do not consider internal wall partitions. Once the boundary conditions and construction materials are assigned to each surface, they are grouped into thermal zones. The models consist of two types of thermal zones: living spaces and the attic. The attic is uninsulated, with no internal heat gain and a ventilation rate of 3 air changes per hour (ACH). For the living space, a thermal zone is assigned to each storey.

Additionally, thermal loads, air leakage, and thermal mass are included according to the input data. For thermal loads, electric equipment, gas equipment, and lighting are considered from the input data, with load fractions of 0.83, 0.5, and 1, respectively, as defined by Hendron and Engebrecht [33]. The number of occupants is an input of the archetypes, each having a default metabolic rate of 126 W/person [34]. Due to the lack of national data on hourly load schedules, they are defined following Hendron and Engebrecht [33].

Finally, the BEM models are exported and stored as *.idf* files using the simulation software EnergyPlus. The simulations are then run using parallel processing computers to reduce computational time.



(a) (b)

Figure 2: 3D representations of the building energy model using Rhino6/Grasshopper/Honeybee. a) One-storey detached house ID6.idf b) two-storeys mid-terrace house ID10.idf

2.4 Climatic data and zones

The generation of models was organized by the Chilean administrative and political regions. This level of resolution enables national and regional governments to use the results to inform direct region-specific public policies or energy programs. Furthermore, Molina *et al.* [19] provide the archetypes for each of these regions, allowing for the estimation of energy consumption and associated GHG emissions of residential buildings at a regional scale. Table 2 shows the 15 regions[1], their associated climate zones, and the national regulatory maximum U values for walls, roofs, and overhang floors [28]. The current Chilean building code divides the territory into seven thermal zones based on the annual heating degree days (HDD), considering a base temperature of 15 °C. Custom weather files were extracted from Meteonorm v7.2 (1990-2010) for each region[35].

Table 2: Chilean administrative regions from north to south and climate zones with associated regulatory U-values

Region N-S	**City**	**Koppen Climate classification** [36]	**Climate zones**	**U-values [W/m² K]**		
				Walls	**Roofs**	**Overhang floors**
15	Arica	BWn	1	4.00	0.84	3.57
1	Iquique	BWn				

[1] A 16th region, the Ñuble region, was created in 2018 from the division of Biobío Region. This region is not considered for this study because it is not included in any of the existing databases used in this study, and so the 8th region is applicable to it.

2	Antofagasta	BWn				
3	Copiapó	BWk				
4	La Serena	BWn				
5	Valparaíso	Csbn	2	3.03	0.6	0.87
13	Santiago	Csb	3	1.89	0.47	0.7
6	Rancagua	Csb				
7	Talca	Csb	4	1.69	0.38	0.6
8	Concepción	Csbn's				
9	Temuco	Cfb	5	1.59	0.33	0.5
14	Valdivia	Cfb				
10	Puerto Montt	Cfbs	6	1.1	0.28	0.39
11	Coyhaique	Cfc	7	0.6	0.25	0.32
12	Punta Arenas	BSk's				

2.5 Input data generation and simulations

Multiple sets of inputs are generated for each region number. Each iteration generates ten LHS numbers, values between 0 and 1, to obtain the 27 parameters for each region, resulting in ten different BEMs per iteration. As each of the LHS numbers corresponds to the quantile of the corresponding parameter distribution, an optimized combination of both physical and categorical parameters is obtained (see Supplementary material). The batch simulations are carried out using the statistical software *R* and the *eplusr* package [37]. The number of simulations is increased by adding a new set of ten BEMs until a stopping criterion is met (here, the difference in the mean energy use is less than 0.5% between one set of samples and the previous one used).

2.6 Model outputs and post-processing of the simulations results

The space heating and energy use intensity results from each BEM are retained and analyzed regionally. Comparing the simulation indoor temperature distribution to public data on the thermal behavior of the residential stock is relevant for this study. This is carried out using data from the National Monitoring Network of Houses (RENAM), a crowdsourcing program from the Chilean Ministry of Housing that collects and anonymously broadcasts the hourly measurements of indoor environment quality variables from 294 houses distributed across five cities (Antofagasta, Valparaíso, Santiago, Temuco, and Coyhaique) [38]. The sensors measure air temperature, relative humidity, noise, and carbon dioxide concentrations in one room of each house. For the comparison, apartments are discarded because only detached and semi-detached houses are present in the top 29 archetypes selected by Molina *et al.* [19]. The five cities studied by The RENAM are used to compare indoor temperature simulation results by season. To compare with measured indoor temperatures, simulations were carried out using historical weather files representing the years 2017 and 2018, data extracted from Metenorm v7.2, and meteorological data from the national meteorological office repository [39], which provide monthly data on global radiation, temperature, relative humidity, and wind speed.

This verification process will allow us to discuss ways to improve the modelling method and reduce the uncertainty of the model inputs. To compare the agreement between both datasets, the Wasserstein distance is used[40,41], which measures the dissimilarity between two probability distributions. The Wasserstein distance also known as the earth mover's distance quantify the mass transportation required to transform one distribution into the other.

2.7 Sensitivity analyses

The influence of the input data variability on the space heating and temperature statistics is analysed as a guideline for improving the BEM model. A statistical test is used to identify the input parameters that need prioritization in future energy studies. Here, the Spearman's ρ correlation coefficient is reported to show how strongly related the two datasets are (for each input-output combination), as it tests for the strength of linear and monotonic association, giving a correlation coefficient between -1 and 1; weakening the strengths as it approaches zero. Moreover, Spearman's ρ can also be applied to ordinal data, and because it is based on the ranking of variables, it is less sensitive to outliers. Finally, the magnitudes of the test coefficients are used to rank the relative importance of each input to the output. For simplicity, three thermal zones are selected for this comparison: zones 1, 3, and 7.

3 Results and discussion

3.1 Simulation results

Around 75 sets of input data were needed for each region, ranging between 30 for geographical region 10, and 150 for region 1. The convergence analysis on space heating energy use establishes a total of 1,130 archetypes nationwide. Considering a total of 5,167,728 single-family houses occupied in Chile, the BSEM set has a ratio of one building energy model for 4,573 houses. Compared to the ResStock program [16], which established 1 model for 241 houses, this ratio appears reasonable for computational time. Indeed, the BSEM set was simulated in 55 minutes on a virtual machine installed in a calculation server (Xeon(R) CPU E5-2640 v4 @ 2.40GHz) with 64 GB allocated Ram.

Figure 3 shows the energy breakdown by energy source and energy intensity for each end-use by administrative region. Space heating is the only energy-use output from the BEM simulation analyzed here. The other end uses, such as lighting, electric equipment, domestic hot water, and kitchen, are generated by the LHS algorithm. Liquefied petroleum gas (LPG) is the country's most widely used energy source for heating, domestic hot water, and cooking. However, firewood and natural gas are popular alternatives in southern regions. The use of firewood increases in southern regions 7, 8, 9, 10, and 14, which have a more developed agroforestry industry [42].

As expected, the regional median energy use intensity varies from 48 kWh/m².year in the more northerly region (15th) to 275 kWh/m².year in the more southerly region (12th). Space heating is the primary driver of this variation, in line with Chilean climate variability, ranging from Hot Desert (BW), where 63% of the stock uses less than 10% of the total energy for heating, to Cold Semi-arid (BS) climates, where the median energy use for heating is 75% (mean=71%; SD=20%). The energy use intensities for space heating are lower than those simulated by Rouault *et al.* [43] who estimated space heating demands varying from 20 to 600 kWh/m².year using a simplified hourly model based on the ISO 17930 standard. This difference may suggest that Chilean households consume less energy than international models can predict, which can indicate a prebound effect or that they are energy poor. This is an important phenomenon for future research involving the social sciences, building physics, and economics.

Figure 3 also compares the simulation results with the national energy balance [17] provided by the Chilean Ministry of Energy. Although the energy source breakdown (Figure 3a) shows that regions from 5 to 11 have the same energy sources as the national energy balance, they are in slightly different proportions. The simulation, which is based on the national energy survey, discarded firewood as an energy source in regions 15, 1, and 2 and significantly underestimated it in regions 3, 4, 11, and 12. On the other hand, the simulation overestimates the use of oil in most regions. Despite these differences in the energy breakdown, the simulated total energy use by region (Figure 3b) shows good agreement with the national energy balance.

Space cooling is not considered in this study because this energy use has been deemed negligible, according to the national energy survey, in which only 79 of 3,500 (2%) respondents reported owning an air conditioning unit. However, Chile has a young and expanding residential air conditioning market. The national energy surveys found that air cooling in the metropolitan area (Region 13) has climbed from 1% in 2009 to 4.5% in 2018 [44,45]. Additionally, the market saturation in Chile is estimated to be 35% of the residential stock. Furthermore, the market saturation might reach at least 60% by 2050 if the projected GDP per capita is USD $31,500 [46] and the average household size decreases to 2.85 people, as determined by linear regression. Consequently, space cooling should be considered in future work, particularly under these economic and demographic scenarios.

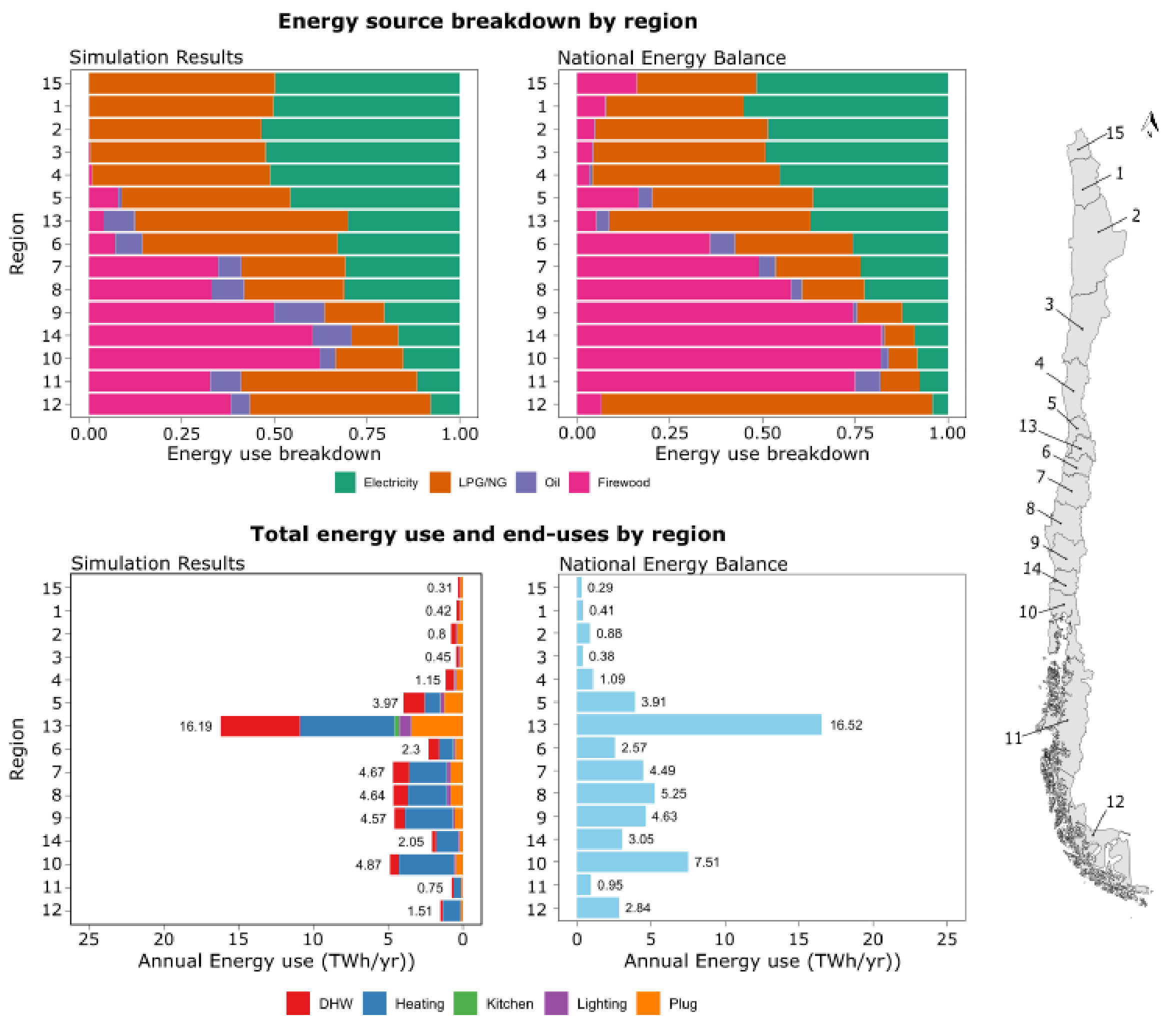


Figure 3: Distribution of energy source and energy end uses by region

3.2 Thermal behaviour

Figure 4 compares the density distributions of indoor dry bulb temperature in the RENAM houses with the simulation results of the BEMs generated for these five cities. The analyses are separated into two periods:

winter (May to October) and summer (November to April). Although the indoor temperature distributions measured by RENAM in 2017 and 2018 differ, the weather variations between these two years have a negligible impact on the distribution of simulated indoor temperatures; see Figure 4.

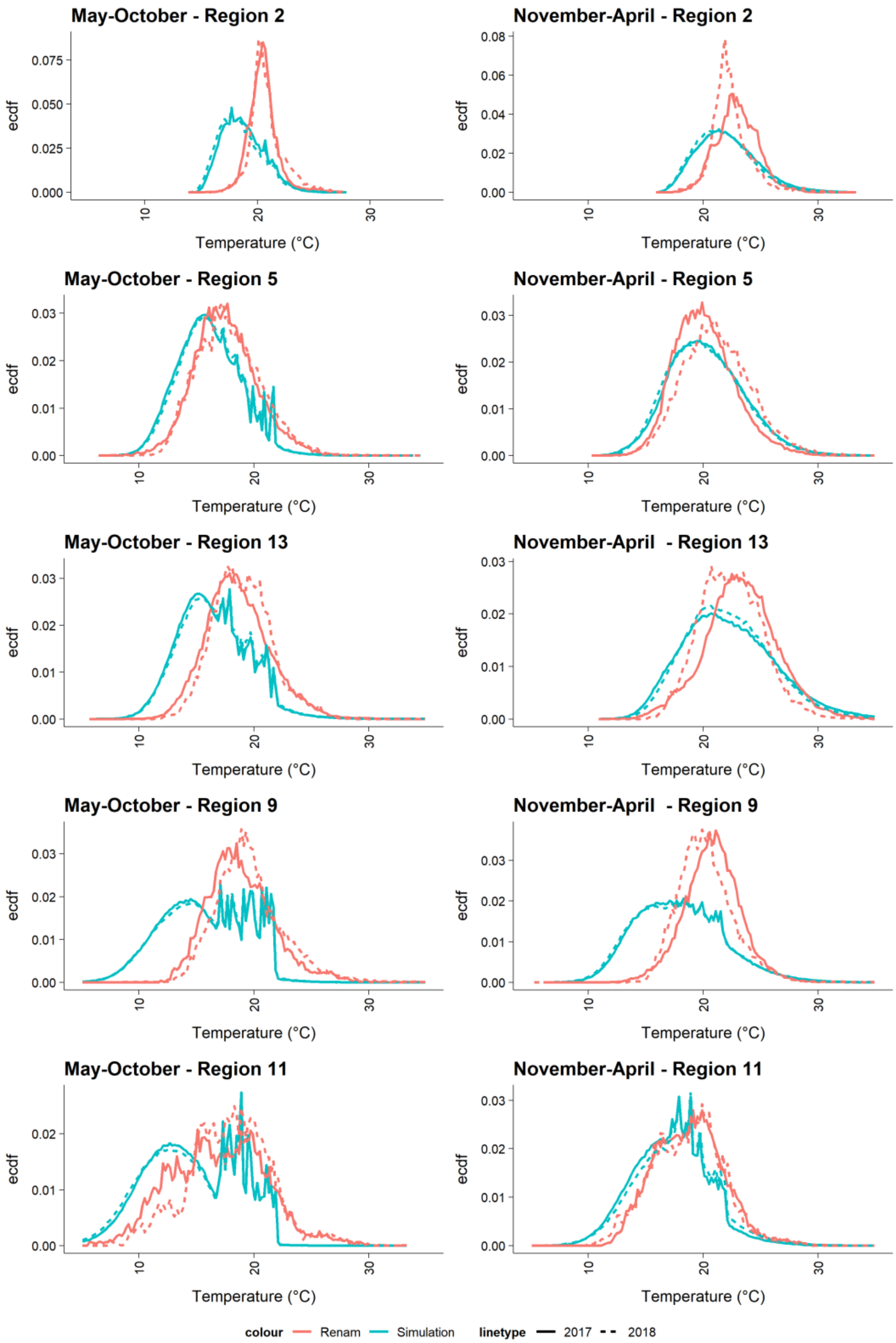


Figure 4:Density distributions of indoor dry bulb temperature measured by RENAM (in red) and simulated (in blue) in 2017 (solid lines) and 2018 (dashed lines).

Figure 4 indicates that the simulated mean indoor temperatures are slightly lower than the measured temperatures across all regions during the winter period. This difference is particularly pronounced in extreme climate regions, such as Antofagasta (Region 2), one of the warmest regions, and Temuco (Region 9), one of the coldest. This discrepancy may stem from an underestimation of the number of heating hours or the space heating temperature setpoint, which requires verification with empirical data. Interestingly, this discrepancy is not observed in Region 11. However, the measured interior temperatures in Region 11 are lower than in other regions, indicating a potential risk of energy poverty.

Table 3 summarizes the results using the Wasserstein distance. For example, a distribution that deviates by 1°C from the mean of the reference distribution (mean = 23°C, standard deviation = 3°C) results in a Wasserstein distance of 2.13×10^{-5} whereas a 1°C deviation in the standard deviation yields a Wasserstein distance of 1.57×10^{-3}. The results show that Regions 2 and 11 exhibit the best agreement with the measured data, indicated by low Wasserstein distances, while Region 2 has the largest distance overall. Although the findings demonstrate a relatively good alignment with the available national data on indoor temperatures, there is potential for further calibration of the energy models through enhanced characterization of household activity data.

Table 3: Results of Wasserstein distance between the density distributions of indoor temperatures measured by RENAM monitoring network and simulation results of this study.

Region N-S	**City**	**Number of monitored houses, RENAM**	**EMD**			
			2017		**2018**	
			Winter	**Summer**	**Winter**	**Summer**
2	Antofagasta	18	7.93×10^{-3}	4.39×10^{-3}	6.81×10^{-3}	7.62×10^{-3}
5	Valparaíso	34	3.14×10^{-4}	1.68×10^{-3}	4.61×10^{-5}	6.68×10^{-4}
13	Santiago	84	6.94×10^{-4}	1.87×10^{-3}	$1{,}40 \times 10^{-3}$	2.21×10^{-3}
9	Temuco	41	1.75×10^{-3}	3.09×10^{-3}	1.90×10^{-3}	3.36×10^{-3}
11	Coyhaique	10	3.49×10^{-4}	3.72×10^{-4}	1.04×10^{-3}	4.21×10^{-4}

3.3 Influential variables

Figure 5 presents Spearman's correlation coefficients between the input variables and selected output variables: (1) energy use intensity (EUI), (2) mean indoor air temperature (T), and (3) standard deviation of indoor air temperatures (σ). The analysis focuses on three climate zones: 1, 3, and 7.

Among the input variables, those related to space heating—specifically, hours of heating, heating setpoint, and months of heating—exhibit the highest correlation coefficients with EUI for space heating and indoor air temperature during winter (see the bottom-left corner of Figure 5). This indicates that they are the most influential factors affecting the outputs in the Chilean residential stock.

In contrast, building envelope parameters, such as air permeability, glazing ratio, heat transfer coefficients, and material types, show a moderate influence on both EUI for space heating and indoor temperatures during summer and winter. Lastly, internal heat gains from lighting, gas, and electric equipment have a negligible impact on space heating and indoor temperatures compared to the other parameters.

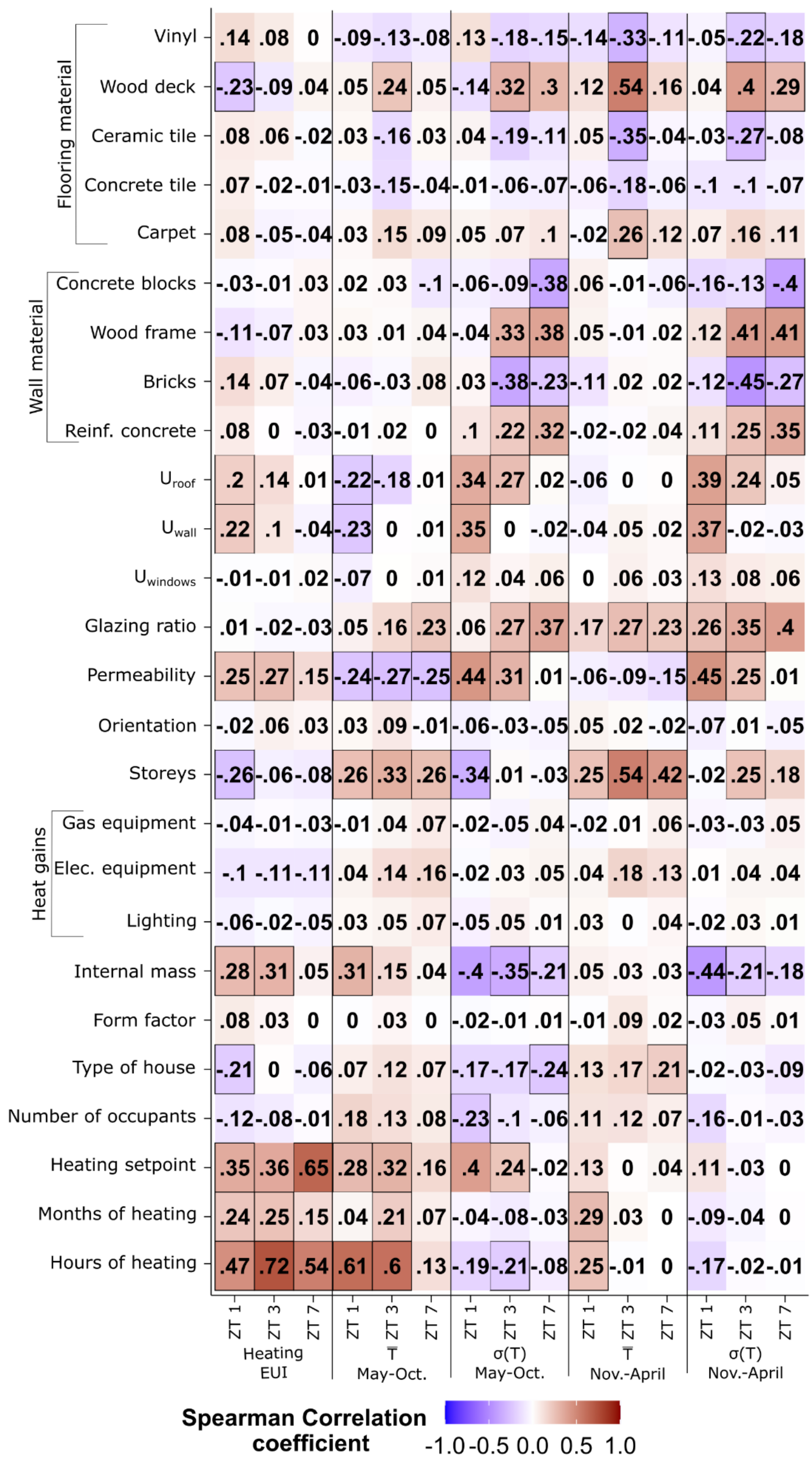


Figure 5: Spearman's correlation coefficient shows the magnitude and direction of the relationships between the input variables and the outputs (the space heating EUI and the mean and standard deviation of indoor temperature for both summer and winter.

Some variables, such as air permeability, wall and flooring materials, and the number of storeys, exhibit correlation coefficients higher than 0.2, indicating they are relatively well defined. Conversely, variables with the highest correlation coefficients, like internal mass and heating setpoint, require better characterization.

Given that space heating control variables are the most influential factors on energy use for space heating, their accurate characterization is essential for developing future energy-saving roadmaps, particularly due to the potential rebound effect. According to the latest national energy survey, half of the respondents heat their homes for less than 5 hours per day. In the coldest climatic zone (ZT 7), half of the respondents reported heating their homes for 6 hours or less. Therefore, enhancing the efficiency of dwellings could lead to improved thermal comfort for occupants while mitigating potential health risks, surpassing mere energy savings and potentially manifesting as a rebound effect. Policymakers should consider this scenario alongside a thorough causal analysis when determining regulatory requirements or subsidies aimed at achieving GHG emission reduction objectives.

The proposed BSEM, serving as a baseline or current scenario, can now be utilized to explore various potential scenarios, such as modifying the empirical cumulative distribution function of heating hours as an input.

Figure 5 also illustrates that, apart from U-values for walls and roofs, the choice of building materials has a moderate impact on indoor temperature conditions. Regarding the indoor temperature spread ($\sigma(T)$), clay bricks and concrete blocks as wall materials moderately reduce temperature variability in ZT3 and ZT7. Both of these heavyweight materials meet current U-value requirements in various climate zones without the need for additional insulation. In contrast, wood frame construction and concrete blocks can increase temperature variability. While reinforced concrete requires insulation to meet efficiency standards, installing it on the interior can diminish the benefits of the concrete's thermal mass.

Additionally, flooring material types can moderately affect both the mean and standard deviation of indoor temperatures during the summer. Vinyl and ceramic tiles, which possess low thermal resistance, enable the ground to function as a heat sink, especially in conjunction with uninsulated slab-on-grade floors, a common feature in most Chilean single-family homes.

4 Conclusion

In this study, a set of 1,130 building energy models (BEMs) was generated to represent the single-family housing stock across the 15 geographic and administrative regions of Chile. This flexible model can be simulated under various future scenarios, including changes in public policies, integration of new technologies, and the impacts of climate change. Unlike the statistical analyses used in the recent national energy survey, the BEM allows for disaggregation of heat transfer, providing deeper insights into the energy needs for space heating. Additionally, it facilitates a detailed examination of indoor thermal comfort during both winter and summer periods, considering current and future climate scenarios.

The following conclusions can be drawn from this study:

- **Methodolog**y: The proposed method of generating BEMs using parametric modeling in Rhino/Grasshopper software, combined with Latin hypercube sampling for input generation, resulted in a BSEM set of 1,130 models. This represents a ratio of one model for every 4,573 houses, allowing for efficient simulations within a limited computational timeframe.
- **Indoor Temperature Distributio**n: The indoor temperature distributions in temperate climate cities, such as Valparaíso and Santiago, show the best agreement with measurements from the RENAM network.
- **Correlation Analysis**: The correlation analysis identified the most influential variables affecting space heating energy use and indoor temperature distribution. Some variables, such as internal mass and space heating setpoints, were assigned uniform distributions due to a lack of available data. Conversely, variables like heating hours and months, provided in the national survey, require better characterization given their significant impact on space heating energy use.

The current BSEM set is limited to simulating the thermal behavior and space heating energy use of single-family homes. While space heating is the primary energy consumption source in temperate and cold climates (from regions 13 to 12), domestic hot water also constitutes a significant portion of energy use. Future models should incorporate a domestic hot water component to enhance the BSEM set.

Future work should focus on two key aspects to improve the proposed energy model set. First, the set should integrate other building types—such as multi-family residences, commercial buildings, and offices—to provide a comprehensive representation of the current building stock. Second, the input data must be enriched with direct measurements, particularly for the most influential variables, including space heating setpoints, internal mass, and permeability. Finally, future research should utilize this BSEM set to assess the adaptability of built heritage in mitigating climate change scenarios and to draft a roadmap for achieving net-zero carbon built heritage by 2050, incorporating retrofitting solutions and renewable energy sources.

5 References

[1] International Energy Agency, Technology Roadmap Energy-efficient Building: : Heating and Cooling Equipment, Paris, France, 2011.

[2] J. Laustsen, Energy Efficiency Requirements in Building Codes, Energy Efficiency Policies for New Buildings, IAE Information Paper (2008).

[3] European Parliament, Council of the European Union, Directive 2010/31/EU on the energy performance of buildings, Official Journal of the European Union (2010).

[4] J.M. Santos-Herrero, J.M. Lopez-Guede, I. Flores-Abascal, Modeling, simulation and control tools for nZEB: A state-of-the-art review, Renewable and Sustainable Energy Reviews 142 (2021) 110851. https://doi.org/10.1016/j.rser.2021.110851.

[5] E. Burman, D. Mumovic, J. Kimpian, Towards measurement and verification of energy performance under the framework of the European directive for energy performance of buildings, Energy 77 (2014) 153–163. https://doi.org/10.1016/J.ENERGY.2014.05.102.

[6] S. Kelly, D. Crawford-Brown, M.G. Pollitt, Building performance evaluation and certification in the UK: Is SAP fit for purpose?, Renewable and Sustainable Energy Reviews 16 (2012) 6861–6878. https://doi.org/10.1016/j.rser.2012.07.018.

[7] A. Marshall, R. Fitton, W. Swan, D. Farmer, D. Johnston, M. Benjaber, Y. Ji, Domestic building fabric performance: Closing the gap between the in situ measured and modelled performance, Energy Build 150 (2017) 307–317. https://doi.org/10.1016/J.ENBUILD.2017.06.028.

[8] O.T. Masoso, L.J. Grobler, The dark side of occupants' behaviour on building energy use, Energy Build 42 (2010) 173–177. https://doi.org/10.1016/J.ENBUILD.2009.08.009.

[9] O. Guerra Santin, L. Itard, H. Visscher, The effect of occupancy and building characteristics on energy use for space and water heating in Dutch residential stock, Energy Build 41 (2009) 1223–1232. https://doi.org/10.1016/J.ENBUILD.2009.07.002.

[10] L.A. Greening, D.L. Greene, C. Difiglio, Energy efficiency and consumption - the rebound effect - a survey, Energy Policy 28 (2000) 389–401. https://doi.org/10.1016/S0301-4215(00)00021-5.

[11] M. Sunikka-Blank, R. Galvin, Introducing the prebound effect: the gap between performance and actual energy consumption, Https://Doi.Org/10.1080/09613218.2012.690952 40 (2012) 260–273. https://doi.org/10.1080/09613218.2012.690952.

[12] N. Kohler, U. Hassler, The building stock as a research object, Building Research and Information 30 (2002) 226–236. https://doi.org/10.1080/09613210110102238.

[13] J. Langevin, J.L. Reyna, S. Ebrahimigharehbaghi, N. Sandberg, P. Fennell, C. Nägeli, J. Laverge, M. Delghust, Mata, M. Van Hove, J. Webster, F. Federico, M. Jakob, C. Camarasa, Developing a common approach for classifying building stock energy models, Renewable and Sustainable Energy Reviews 133 (2020) 110276. https://doi.org/10.1016/j.rser.2020.110276.

[14] I. Korolija, L. Marjanovic-Halburd, Y. Zhang, V.I. Hanby, UK office buildings archetypal model as methodological approach in development of regression models for predicting building energy consumption from heating and cooling demands, Energy Build 60 (2013) 152–162. https://doi.org/10.1016/j.enbuild.2012.12.032.

[15] É. Mata, A.S. Kalagasidis, F. Johnsson, A modelling strategy for energy, carbon, and cost assessments of building stocks, Energy Build 56 (2013) 100–108. https://doi.org/10.1016/j.enbuild.2012.09.037.

[16] National Renewable Energy Laboratory, ResStock - NREL, National Renewable Energy Laboratory (n.d.). https://resstock.nrel.gov/ (accessed June 17, 2022).

[17] Comisión Nacional de Energía, Regional energy balance, (n.d.). https://3b9x.short.gy/C5ufuZ (accessed February 7, 2023).

[18] U.S. Department of Energy- Building Technologies Office, EnergyPlus | EnergyPlus, (n.d.). https://energyplus.net/ (accessed January 26, 2021).

[19] C. Molina, M. Kent, I. Hall, B. Jones, A data analysis of the Chilean housing stock and the development of modelling archetypes, Energy Build 206 (2020) 109568. https://doi.org/10.1016/j.enbuild.2019.109568.

[20] Ministerio de Vivienda y Urbanismo, Dataset of Energy rated houses, (2022).

[21] F. Asdrubali, F. D'Alessandro, G. Baldinelli, F. Bianchi, Evaluating in situ thermal transmittance of green buildings masonries: A case study, Case Studies in Construction Materials 1 (2014) 53–59. https://doi.org/10.1016/j.cscm.2014.04.004.

[22] C. Peng, Z. Wu, In situ measuring and evaluating the thermal resistance of building construction, Energy Build 40 (2008) 2076–2082. https://doi.org/10.1016/j.enbuild.2008.05.012.

[23] C. Molina, B. Jones, I.P. Hall, M.H. Sherman, CHAARM: A model to predict uncertainties in indoor pollutant concentrations, ventilation and infiltration rates, and associated energy demand in Chilean houses, Energy Build 230 (2021). https://doi.org/10.1016/J.ENBUILD.2020.110539.

[24] M.H. Sherman, D. Dickerhoff, Airtightness of U.S. Dwellings, in: The 1998 ASHRAE Annual Meeting, ASHRAE, Toronto, Canada, 1998.

[25] In-Data, Corporación de Desarrollo Tecnológico, Informe final de usos de la energía de los hogares Chile 2018, Santiago de Chile, 2019.

[26] C. Molina, B. Jones, I.P. Hall, M.H. Sherman, CHAARM: A model to predict uncertainties in indoor pollutant concentrations, ventilation and infiltration rates, and associated energy demand in Chilean houses, Energy Build 230 (2021) 110539. https://doi.org/10.1016/j.enbuild.2020.110539.

[27] Corporación de Desarrollo Tecnológico, Corporación de Desarrollo Tecnológico, Encuesta de los usos de la energía de los hogares Chile 2018 - Anexo 1, (2019).

[28] Ministerio de vivienda y urbanismo, Decreto 192 Artículo 4.1.10 de la Ordenanza General de Urbanismo y Construcciones, (2006). https://www.bcn.cl/leychile/navegar?idNorma=8201&idParte=100018401&idVersion= (accessed June 15, 2022).

[29] H. Johra, P. Heiselberg, Influence of internal thermal mass on the indoor thermal dynamics and integration of phase change materials in furniture for building energy storage: A review, Renewable and Sustainable Energy Reviews 69 (2017) 19–32. https://doi.org/10.1016/J.RSER.2016.11.145.

[30] Instituto nacional de estadísticas de Chile (INE)., Censo De Población y Vivienda 2002, (2003).

[31] Robert McNeel & Associates, Rhino - Rhinoceros 3D v.6, (2018). https://www.rhino3d.com/es/ (accessed September 16, 2021).

[32] Ladybug Tools LLC., Legacy Ladybug 0.0.69/ Honeybee 0.0.66, Ladybug Tools LLC (2020).

[33] R. Hendron, C. Engebrecht, Building America Research Benchmark Definition. Technical Report NREL/TP-550-47246, 2010.

[34] U.S. Department of Energy, Group - Internal Gains, in: EnergyPlus™ Input Output Reference, 2021.

[35] Meteotest (ed.), Meteonorm v7.2, (2017).

[36] R. Rioseco Hormazábal, C. Tesser Obregón, Cartografía interactiva de los climas de Chile, Pontificia Universidad Católica de Chile (n.d.). http://www7.uc.cl/sw_educ/geografia/cartografiainteractiva/index.htm (accessed June 15, 2022).

[37] H. Jia, A. Chong, eplusr: A framework for integrating building energy simulation and data-driven analytics, Energy Build 237 (2021) 110757. https://doi.org/10.1016/j.enbuild.2021.110757.

[38] Ministerio de Vivienda y Urbansimo, Kuantum, Red de Monitoreo de Viviendas, (n.d.). https://renam.cl/ (accessed September 19, 2021).

[39] Dirección Meteorológica de Chile, Servicios Climáticos, Dirección General De Aeronáutica Civil (n.d.). https://climatologia.meteochile.gob.cl/ (accessed June 15, 2022).

[40] M. Muskulus, S. Verduyn-Lunel, Wasserstein distances in the analysis of time series and dynamical systems, Physica D 240 (2011) 45–58. https://doi.org/10.1016/J.PHYSD.2010.08.005.

[41] R. Boumaza, P. Santagostini, S. Yousfi, S. Demotes-Mainard, dad: an R Package for Visualisation, Classification and Discrimination of Multivariate Groups Modelled by their Densities, R J 13 (2021) 179–207. https://hal.inrae.fr/hal-03344822 (accessed February 5, 2023).

[42] Inventario Forestal Nacional, Regional maps of native forest and forestry plantation, (n.d.). https://ifn.infor.cl/index.php/informacion-regional (accessed June 15, 2022).

[43] F. Rouault, F. Ossio, P. González-Levín, F. Meza, Impact of Climate Change on the Energy Needs of Houses in Chile, Sustainability 2019, Vol. 11, Page 7068 11 (2019) 7068. https://doi.org/10.3390/SU11247068.

[44] In-Data, Informe final de usos de la energía de los hogares Chile 2018, Santiago de Chile, 2019.

[45] Corporation de Desarrollo Tecnológico, Estudio de Usos Finales y Curva de Oferta de Conservación de la Energía en el Sector Residencial de Chile, Santiago, Chile, 2010.

[46] OECD, Domestic product - GDP long-term forecast, (n.d.). https://data.oecd.org/gdp/gdp-long-term-forecast.htm (accessed January 18, 2022).