Title

Influence of climate on the creation of multilayer perceptrons to analyse the risk of fuel poverty

Authors:

David Bienvenido-Huertas[1], Alexis Pérez-Fargallo[2], Raúl Alvarado-Amador[2], Carlos Rubio-Bellido*[3]
[1]Department of Graphical Expression and Building Engineering, University of Seville. Seville, Spain
[2]Department of Building Science, University of Bío-Bío, Concepción, Chile
[3]Department of Building Construction II, University of Seville. Seville, Spain
*Author to whom correspondence should be addressed;
Higher Technical School of Building Engineering, Ave. Reina Mercedes 4A, Seville, Spain
E-Mail: carlosrubio@us.es (C.R.B.); Tel.: +34-686-135-595

Highlights:

- Fuel Poverty Potential Risk Index (FPPRI) applied to Santiago, Concepción, and Valparaiso (Chile).
- A total of 116,640 cases were analysed, considering 9 morphological variables per decile.
- Combination of 84 datasets using 2 approaches.
- Performance greater than 96% of the individual models for each climate zone.

Abstract:

Many studies are focused on the diagnosis of fuel poverty. However, its prediction before occupying households is a developing research area. This research studies the feasibility of implementing the Fuel Poverty Potential Risk Index (FPPRI) in different climate zones of Chile by means of regression models based on artificial neural networks (ANNs). A total of 116,640 representative case studies were carried out in the three cities with the largest population in Chile: Santiago, Concepción, and Valparaiso. Apart from energy price (EP) and income (IN), 9 variables related to the morphology of the building were considered in approach 1. Furthermore, approach 2 was developed by including comfort hours (NCH). A total of 84 datasets were combined considering both approaches and the 5 most unfavourable deciles according to the income level of Chilean families. The results of both approaches showed a better performance in the use of individual models for each climate (MLP_C, MLP_S, and MLP_V), and the dataset with all deciles (Full) could be used. Regarding the influence of the input variables on the models, IN was the most determinant, and NCH becomes important in approach 2. The potential of using this methodology to allocate social housing would guarantee the main objective of the country: the reduction of fuel poverty in the roadmap for 2050.

Keywords:

Fuel Poverty; Fuel Poverty Potential Risk Index (FPPRI); Climate zone; Multilayer perceptrons; Social housing; Policymaking

Nomenclature

Symbols

EC: Energy consumption [kWh]
EC_{SC}: Average monthly cooling energy consumption starting from the simulation [kWh]
EC_{SE+L}: Average monthly energy consumption of equipment and lighting from the simulation [kWh]
EC_{SH}: Average monthly heating energy consumption from the simulation [kWh]
EP_C: Electricity price for cooling [$/kWh]
EP_{E+L}: Electricity price for equipment and lighting [$/kWh]
EP_H: Energy of fuel price for heating [$/kWh]
$FPI_{adaptative}$: Fuel poverty index with an adaptive energy consumption
H_c: Occupied hours in thermal comfort applying category III of EN 15251:2007 [h]
H_d: Unoccupied hours of the analysed period [h]
H_t: Total hours of the analysed period [h]
MAE: Mean absolute error

m_i: Multilayer perceptron prediction
n: Number of instances in the training or testing dataset
R^2: Correlation coefficient
$RMSE$: Root mean square error
t_i: Actual value
TI: Threshold income [$]

Greek letters

μ_{D_n}: Average of the income distribution of decile n
$\sigma^2_{D_n}$: Variance of the income distribution of decile n

Abbreviations

ANN: Artificial neural network
C: Concepción
CI: Central internal climate zone (Chile)
CL: Central coastal climate zone (Chile)
D: Shadow's distance [m]
D_1: First decile
D_2: Second decile
D_3: Third decile
D_4: Fourth decile
D_5: Fifth decile
EP: Energy price [$]
FPI: Fuel poverty index
FPPRI: Fuel Poverty Potential Risk Index
FR: Form ratio
Full: Set of the five poorest family deciles (from D_1 to D_5)
H: Shadow's height [m]
IN: Household income [$]
MLP: Multilayer perceptron
MLP_C: Multilayer perceptron trained with data of the climate of Concepción
MLP_S: Multilayer perceptron trained with data of the climate of Santiago
MLP_V: Multilayer perceptron trained with data of the climate of Valparaiso
$MLP_{C\text{-}S}$: Multilayer perceptron trained with data of the climates of Concepción and Santiago
$MLP_{C\text{-}V}$: Multilayer perceptron trained with data of the climates of Concepción and Valparaiso
$MLP_{S\text{-}V}$: Multilayer perceptron trained with data of the climates of Santiago and Valparaiso
$MLP_{C\text{-}S\text{-}V}$: Multilayer perceptron trained with data of the climates of Concepción, Santiago and Valparaiso
NCH: Number of comfort hours [h]
OCDE: Organisation for the Economic Co-operation and Development
OR: Orientation
S: Santiago
SAG: Surface area in contact with the ground [m²]
SAH: Horizontal surface area in contact with other homes [m²]
SAV: Vertical surface area in contact with other homes [m²]
SAR: Roof surface area [m²]
SL: South coastal climate zone (Chile)
V: Valparaiso
VOL: Volume [m³]

1. Introduction

Global warming, the increase of environmental pollution, and Earth's land degradation constitute the main problems faced by state policies [1]. One of the main reasons for such problems are the greenhouse gases emitted to the atmosphere because of the energy consumption of the main industries and sectors of each country. In this sense, the building sector is among those contributing to such situation, with buildings being responsible for between 30 and 40% of the total energy consumption [2,3]. The energy problem of buildings increases therefore the cases of fuel poverty [4].

Generally, poverty has been one of the main concerns in the last fifty years. The United Nations Millennium Development Goals suggest the eradication of extreme poverty as a major objective in order to achieve acceptable conditions for the sustainable development of the various communities [5]. Meeting basic needs (e.g., thermal comfort) by means of energy consumption can be a real challenge for developing countries [6], mainly due to the combination of political and economic factors, as well as to the limitations to generate and distribute the energy [7].

In developed countries, there are no obstacles to generate and distribute energy, so the problem to access to energy is focused on the potential of households to access to such energy. Because of the variations of energy prices, transformations of the energy sector, the high energy consumption of buildings, and the low incomes of families due to the economic crisis, many countries (e.g., United Kingdom or United States) have positioned their energy policies to make possible that households with lower economic resources access to energy [8,9]. If families cannot pay bills, health problems can be caused [10,11], as well as the death of members in extreme cases [12,13].

Such limitations for households to access to energy resources, or even accessing in an inadequate way (thereby limiting inhabitable conditions inside dwellings), is known as energy poverty or fuel poverty [14,15]. Another possible definition was given by Boardman [16] by establishing that households presenting limitations to guarantee an acceptable thermal comfort in their interior as well as expending more than 10% of their incomes in energy consumption are in fuel poverty. The mathematical expression used for this definition is as follows:

$$FPI = \frac{EP \cdot EC}{I} \tag{1}$$

Where FPI [dimensionless] is the fuel poverty index, EP is the energy price used [$/kWh], EC is the energy consumption [kWh], and I is the household income ($).

Boardman's approach has been widely used in several studies from different countries, such as: (i) O'Sullivan et al. [17] analysed the limitations of using the prepayment metering for electricity to identify households in risk of fuel poverty in New Zealand. The results showed that the method could be used to target energy policies in spite of generating a tendency of greater levels of fuel poverty in households; (ii) Legendre and Ricci [18] studied the approach of fuel poverty of 10%, together with other two approaches in French dwellings. The results showed the high probability of cases of fuel poverty in buildings occupied by retired people living alone. Likewise, factors such as the fact of being the owner of the dwelling and the type of heating and energy systems used imply an increase or a decrease of the risk of fuel poverty; and (iii) Santamouris et al. [19] analysed the influence of the economic crisis on the energy consumption of households in Greece. After analysing the surveys and the energy consumption between 2010 and 2012, 2% and 14% of families with high and low incomes, respectively, were found to be in fuel poverty.

On the other hand, other studies have analysed the limitations of Boardman's approach by developing new indexes to be adapted better to the characteristics of each region [20]: (i) Fabbri [21] developed the Building Fuel Poverty Index, which combines the heating energy consumption of the building and the average incomes with a fuel poverty line of 6.9%. The author applied such index to various Italian buildings. The results showed values greater than 50% in buildings built between 1950 and 1980, with a decreasing tendency in the index value since around 90s; (ii) Desiere et al. [22] created the progress out of poverty index, which is based on 10 weighted questions on the characteristics of households (e.g., member's education and conditions of the dwelling). The combination of the values assigned by each question determines the risk of energy poverty; (iii) Nussbaumer et al. [23] suggested the multidimensional energy poverty index, which is based on the privatization of 5 basic energy services using 6 indicators. A household is at risk of energy poverty if the multidimensional energy poverty index exceeds a predefined threshold value; (iv) Wang et al. [24] developed the energy poverty comprehensive evaluation index, which is obtained from the sum of four categories: household energy affordability and energy efficiency, energy management completeness, energy consumption cleanliness, and energy service availability. Such index analyses the energy poverty in China; and (v) Bonatz et al. [25] developed a new index of energy poverty to measure cases of fuel poverty in Germany and China. This new index is based on the multidimensional energy poverty index of Nussbaumer et al. [23] by introducing indicators of access and affordability to compare different developing countries.

However, none of such indexes considered a fundamental aspect affecting the housing energy consumption: the adaptive thermal comfort. The adaptive thermal comfort is the human capacity to adapt to external thermal variations by means of different mechanisms. So, both the heat and the cold thresholds vary according to the tendency of oscillations of the external temperature. This is of great interest when carrying out energy analysis in buildings because there could be variations between 10 and 18% in the energy consumption [26]. Under this circumstance, the authors of this research developed the Fuel Poverty Potential Risk Index (FPPRI) in a previous work [27]. Such index is based on applying models of adaptive thermal comfort to the energy consumption of dwellings by using Boardman's criterion of 10% [16]. To determine the FPPRI, the fuel poverty index is first calculated by using the model of adaptive thermal comfort from EN 15251:2007 [28] for category III:

$$FPI_{adaptative} = \frac{EC_{SC} \cdot \left(1 - \frac{H_d + H_c}{H_t}\right) \cdot EP_C + EC_{SH} \cdot \left(1 - \frac{H_d + H_c}{H_t}\right) \cdot EP_H + EC_{SE+L} \cdot EP_{E+L}}{I} \quad (2)$$

Where $FPI_{adaptative}$ is the fuel poverty index with an adaptive energy consumption [dimensionless], EC_{SC} is the average monthly cooling energy consumption starting from the simulation [kWh], EP_C is the electricity price for cooling [$/kWh], H_t is the total hours of the analysed period [h], H_d is the unoccupied hours of the analysed period [h], H_c is the occupied hours in thermal comfort applying category III of EN 15251:2007 [h], EC_{SH} is the average monthly heating energy consumption from the simulation [kWh], EP_H is the energy of fuel price for heating [$/kWh], EC_{SE+L} is the average monthly energy consumption of equipment and lighting from the simulation [kWh], and EP_{E+L} is the electricity price for equipment and lighting [$/kWh].

Given that $FPI_{adaptative}$ is an index similar to that from Eq. (1) (i.e., an assessment index when the dwelling is already occupied), the authors adapted such approach to be used as a predictive index (i.e., determining the probability that the family allocated to the dwelling suffers from cases of fuel poverty). For this purpose, a threshold income was first defined by using Boardman's threshold of 10% (Eq. (3)). The field of study selected was Chile, which uses data from the survey called Chilean National Socioeconomic Characterization [29], and a normal curve of distribution was adjusted to the risk of fuel poverty, in which the area under such curve would determine the FPPRI for a household (Eq. (4)).

$$TI = 10 \cdot I \cdot FPI_{adaptative} \quad (3)$$

$$FPPRI = \int_{-\infty}^{TI} f(TI, \mu_{D_n}, \sigma_{D_n}^2) d_{TI} \quad (4)$$

Where TI is the threshold income [$], μ_{D_n} is the average of the income distribution of decile n [$], and $\sigma_{D_n}^2$ is the variance of the income distribution of decile n.

Chile was used as a case study because of several aspects. On the one hand, Chile is the first South American country to be joined to the Organisation for the Economic Co-operation and Development (OCDE) [30], thus reducing poverty rates from 29.1 to 14.4% between 2006 and 2013 (these percentage data were obtained by means of surveys using the theory of "capabilities framework" for the fuel poverty [31]) [32]. On the other hand, Chile has a strong policy for the development of social housing addressed to the most vulnerable households since 1936 [33]. The national property register of social dwellings in Chile contains 344,402 units [33]. Such social dwellings are defined in different typologies: (i) typologies of type A (A1, A2, A3, A4, A5, and A6) for individual buildings, and (ii) typologies of type B (B1, B2, B3, and B4) for constructions of several block of flats connected through access systems. Moreover, another aspect to highlight of Chile is the high electricity prices, which usually are higher than in countries such as France, Norway or United Kingdom, with an average price of 211 $/MWh [34]. This aspect becomes more noticeable if prices are adjusted to the purchasing power parity index, thereby increasing the average price to 251 $/MWh, and considering the increasing tendency presented by prices in recent years [34]. Based on such prices, the monthly cost in a Chilean household can reach $500 [34,35]. Consequently, Chilean households spend little money according to the statistical data of the country [36]. This aspect, together with the low internal temperature in winter (below 15ºC in the main cities [37]), stresses the existing cases of fuel poverty in the country. In addition, there is a relationship between the cases of fuel poverty and the increase of firewood in the region of Valdivia, thereby increasing the air pollution in winter as well as contributing to the appearance of health problems [38].

So, the FPPRI is an opportunity to reduce the risk of fuel poverty before allocating social housing, thus achieving the objectives of the roadmap for 2050 to reduce fuel poverty [39]. Given the technical difficulty and the time required to perform simulations in buildings to which families were allocated, the use of artificial neural networks (ANNs) was recently suggested to estimate the FPPRI [40]. The input variables corresponded to parameters to be easily measured by the staff because such variables were morphological aspects of the building (e.g., orientation and surface areas of the envelope), the energy price and the family income. More adjusted results were obtained by using ANNs as a

regression algorithm than using multivariable regressions. The results showed that the ANNs carried out estimations close to the actual values (with a correlation coefficient higher than 99%), while multivariable regressions obtained a correlation coefficient between 81 and 96%. As a result, the advantage of the approach to carry out accurate estimations is shown, like in other applications of energy estimation (estimation of the energy demand [41] and the energy consumption [42] of buildings).

However, some aspects were not studied in such analysis. Firstly, the case studies analysed by the model corresponded to the region of Bío-Bío in Chile, without analysing the possibilities of estimating the ANNs for other climate regions. This aspect is quite important because there are significant differences in the energy consumption of various climate zones of the country [43]. Secondly, and related to the previous issue, the capacity of generalizing the models in new instances of other climates was not analysed. Therefore, the most effective methodology could not be determined to be applied by regional and provincial governments of the country.

For these reasons, this paper analyses the advantages and limitations of applying prediction models of the FPPRI in different climate zones in Chile. To do this, 38,880 case studies were simulated in three various climates, thereby generating a total of 116,640 case studies. Such case studies were used to train and test 84 prediction models. Based on the results, both the possibilities of estimating the models in other climate and the most effective methodology of implementation were determined.

This article is divided as follows: Section 2 explains in detail the typology of artificial neural network, climate zones of Chile and the workflow of this research. Section 3 assesses and compares the results obtained using different approaches, thus establishing the most effective methodology of implementation. Finally, Section 4 provides the main conclusions and political implications of the results.

2. Methodology

2.1. Algorithm used: multilayer perceptrons

The neural networks is a mathematical model inspired by the neurological brain structure to solve problems [44], both of classification [45,46] and regression [47]. One of the typologies of artificial neural networks providing best features are the multilayer perceptrons (MLPs) because they are models supervised with capacities of universal approximation [48–50]. MLPs consist of a series of nodes or neurons divided into three or more layers: an input layer, one or more hidden layers, and an output layer. The prediction or response given by the model corresponds to the output value of the neuron of the last layer. The output value of this neuron corresponds to the sum of the values of the neurons of the previous layers which are weighted by synaptic weights and by using activation, transference, and propagation functions.

The adjustment of the synaptic weights is an essential process for the development of the MLP as such weights allow the accuracy between the predicted value and the actual value of the instance to be adjusted. For this purpose, a learning algorithm is applied to a training dataset. In this research, the MLPs were trained by means of backpropagation [51–53], using the Broyden-Fletcher-Goldfarb-Shanno [54] algorithm (which belongs to the quasi-Newton methods). Also, the models were developed in the R language. As in the previous work, an architecture of only one layer has been used [40] because such architectures generally have a better performance than architectures of several intermediate layers [55]. Moreover, sigmoidal activation functions were used as activation and transference functions.

2.2. Climate zones of Chile

Chile is a country with a wide climate variety. According to the Köppen-Geiger climate classification [56], climates in this country are dry, mild, and cold. This climate variety implies that the climate characteristics of the different regions of Chile are the same as in different geographical points of the planet, such as Spain or Australia [56]. However, the Köppen-Geiger climate classification provides not enough resolutions in some zones because of some aspects, such as the topography of the ground and the distance from the sea. The topography significantly influences the characteristics of a climate (rainfalls, temperatures, solar radiation, etc.), as many studies show [57–60]. Thus, there are various microclimates in Chile because its altitude ranges between 0 and 6,000 meters above mean sea level [61]. This aspects is reflected in the state regulation in the standard NCh 1079:2008 [62], which classifies the climate of the country in 9 zones (Fig. 1).

As indicated above, this study analyses the potential and limitations of using prediction tools of the FPPRI in the climate variations presented by Chile in different zones. Given the variety of existing climate zones, three of them were analysed: south coastal zone (SL), central internal zone (CI), and central coastal zone (CL). These climates are very

different [43] and the feasibility of estimating the models can be applied to other climate zones. The cities selected for each climate as references were Concepción (SL), Santiago (CI), and Valparaiso (CL).

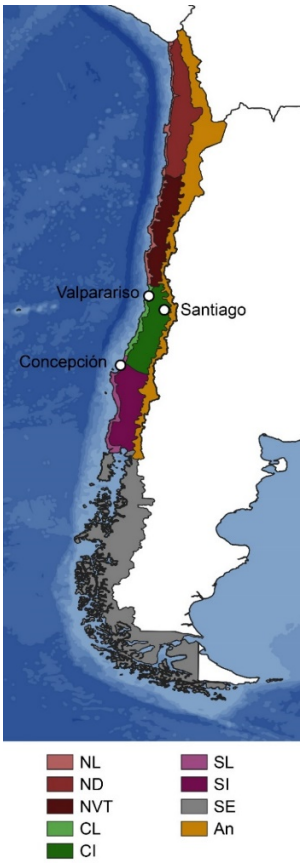

Fig. 1. Climate zones of Chile and the cities analysed.

2.3. Data collection, approaches, and training and testing procedure

Fig. 2 represents the flowchart of this research. Training and testing datasets were generated by making simulations of typologies A1 (individual building with internal vertical access) and B1 (parallel buildings with external vertical access) for projects of social housing developed by the Chilean Ministry of Housing and Urbanism (MINVU in Spanish) [63]. Both typologies were selected because they represent 40.5% of the total existing social housing in the country [64] (the remaining percentage is divided into 8 typologies of social housing). Furthermore, such typologies of social housing are the most built in the country since 2005 [34]. Likewise, the income levels of the families of the dataset were determined according to the poorest deciles. According to the Chilean national socioeconomic survey, Chilean families can be classified in 10 groups (called deciles) depending on their economic incomes (the 10th decile corresponds to the wealthiest families, and the 1st decile to the poorest families) [32]. The five first deciles were therefore considered in this study because they were the most vulnerable (Fig. 3). The maximum and minimum expected household income per decile could therefore be estimated according to the average number of inhabitants per household [32].

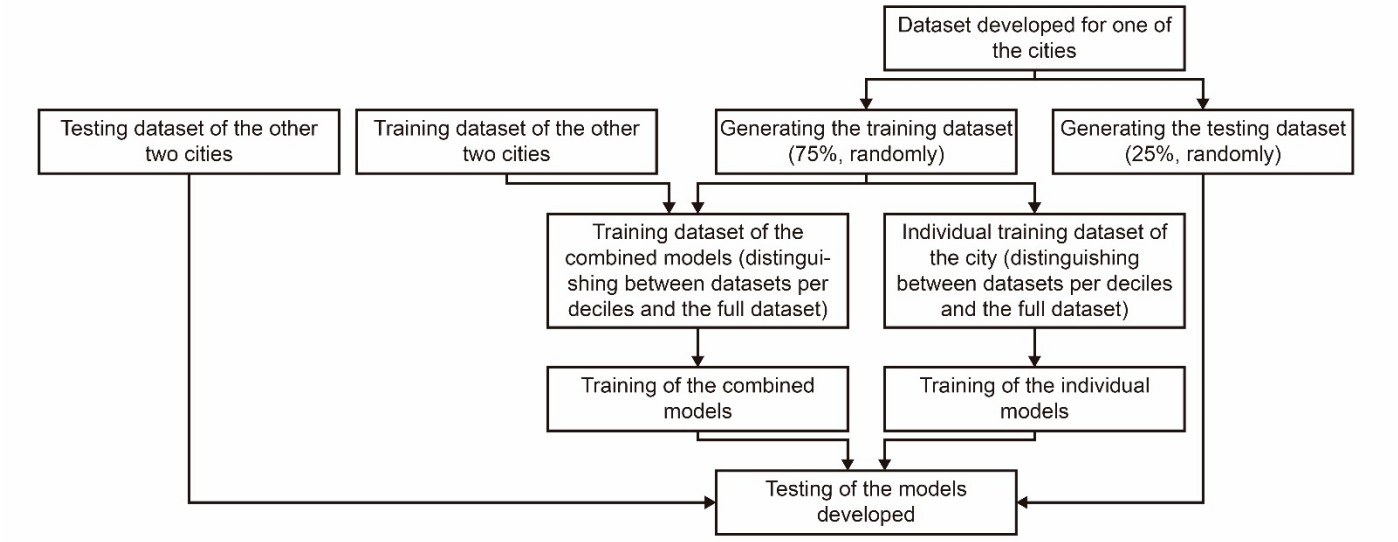

Fig. 2. Flowchart of the training and testing procedure of the models.

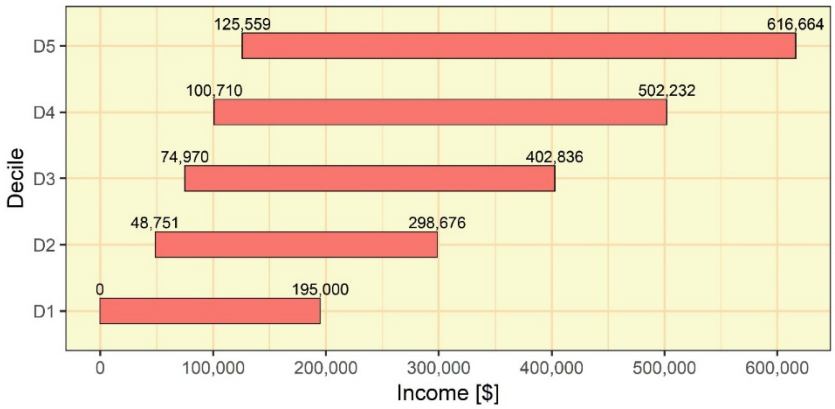

Fig. 3. Upper and lower limit income values of the Chilean deciles.

The thermophysical properties of the envelope's elements were obtained from the standard DS 47 [63]. The minimal values of air tightness, the air change rate, and the thermal loads of occupation, equipment, and lighting were obtained from the Chilean Sustainable Building Code (CCS in Spanish) [65]. Fig. 4. shows the percentage distribution per hour of the internal loads by CCS. Different combinations of surface, form ratio, orientation, etc. were made in the simulations, as in a previous research [34].

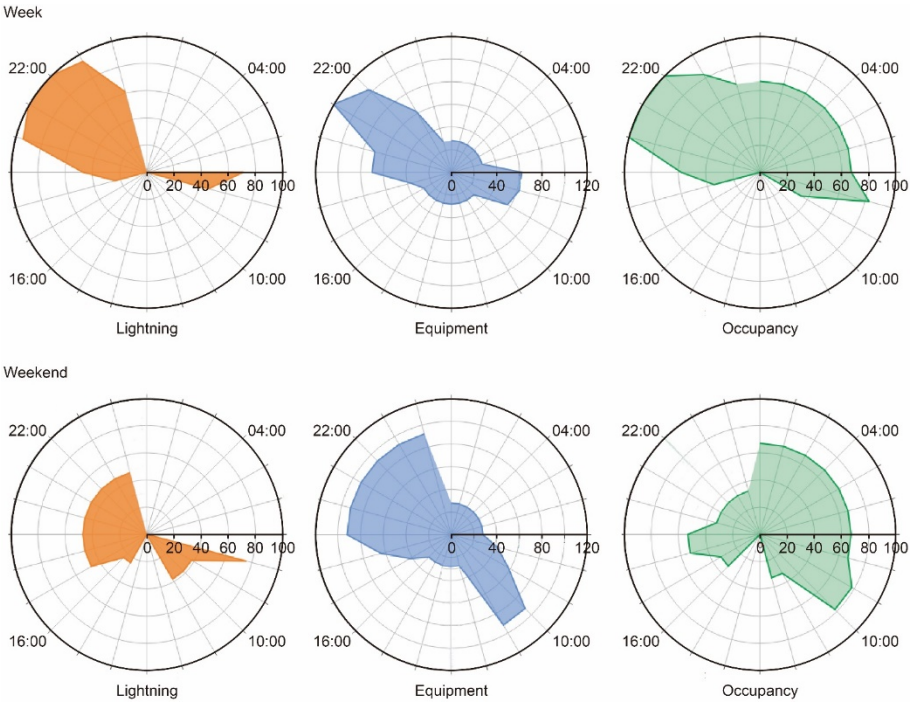

Fig. 4. Distribution of loads for the days of the week and weekends.

Each combination of case study was simulated in the climates of Concepción, Santiago, and Valparaiso. A total of 38,880 case studies were carried out per city. The simulated cases were designed according to Chilean regulations [33,63]. Further information on the parameters used can be found in a previous research [34]. Also, the simulations were undertaken by using Energy Plus software. In such 38,880 case studies, 7,776 case studies corresponded to the poorest deciles (i.e., from D_1 to D_5). The FPPRI was determined to each case study according to their income levels.

After simulating 38,880 case studies per city (i.e., a total of 116,640 case studies), datasets were created. For this purpose, two possibilities for the architecture of the MLP were considered (Fig. 5). The former corresponds to the original approach from the previous research [40] In this approach, the input variables were as follows: (i) orientation (OR); (ii) form ratio (FR); (iii) volume (VOL); (iv) surface area in contact with the ground (SAG); (v) horizontal surface area in contact with other homes (SAH); (vi) roof surface area (SAR); (vii) vertical surface area in contact with other homes (SAV); (viii) shadow's distance (D); (ix) shadow height (H); (x) energy price (EP); and (xi) household income (IN). D and H values have been analysed using the combination with the orientation variables and studying the four orientations with nine combinations of element possibilities that could generate shade over the simulated case. D and H values, which will characterize the position and size of the shade element, have been calculated starting from the winter solstice. Three values of H have been calculated so that the winter solstice projects cast shade over all the ground, intermediate or top floors for a D between 3 and 5.71 m (distances according to Chilean regulations [63]). The output variable was FPPRI. It is worth noting that this variable was logarithmically transformed to obtain a better performance in the estimations.

Given that this study aimed at analysing the capacity of generalization of the models to estimate the FPPRI in other climate regions, a new input variable was introduced in approach 2: the number of thermal comfort hours (NCH). This was due to the lack of a variable considering the variability in the performance of the building because of climate. The number of thermal comfort hours was determined according to the upper and lower limits for category III from EN 15251:2007 [28]. Such limits are in turn determined according to the daily average variation of the external temperature. To verify that there was not a high correlation of the new variable with the previous variables of the model, Pearson's correlation and the p-values were previously analysed. Pearson correlation coefficients between the new predictor variable and the remaining variables were low. Only a low correlation was obtained with VOL, SAG, and SAR, although the p-values obtained for these pairs of variables allowed the null hypothesis to be rejected. Therefore, the relationship between NCH and other variables had a behaviour similar to that of the remaining pairs of variables.

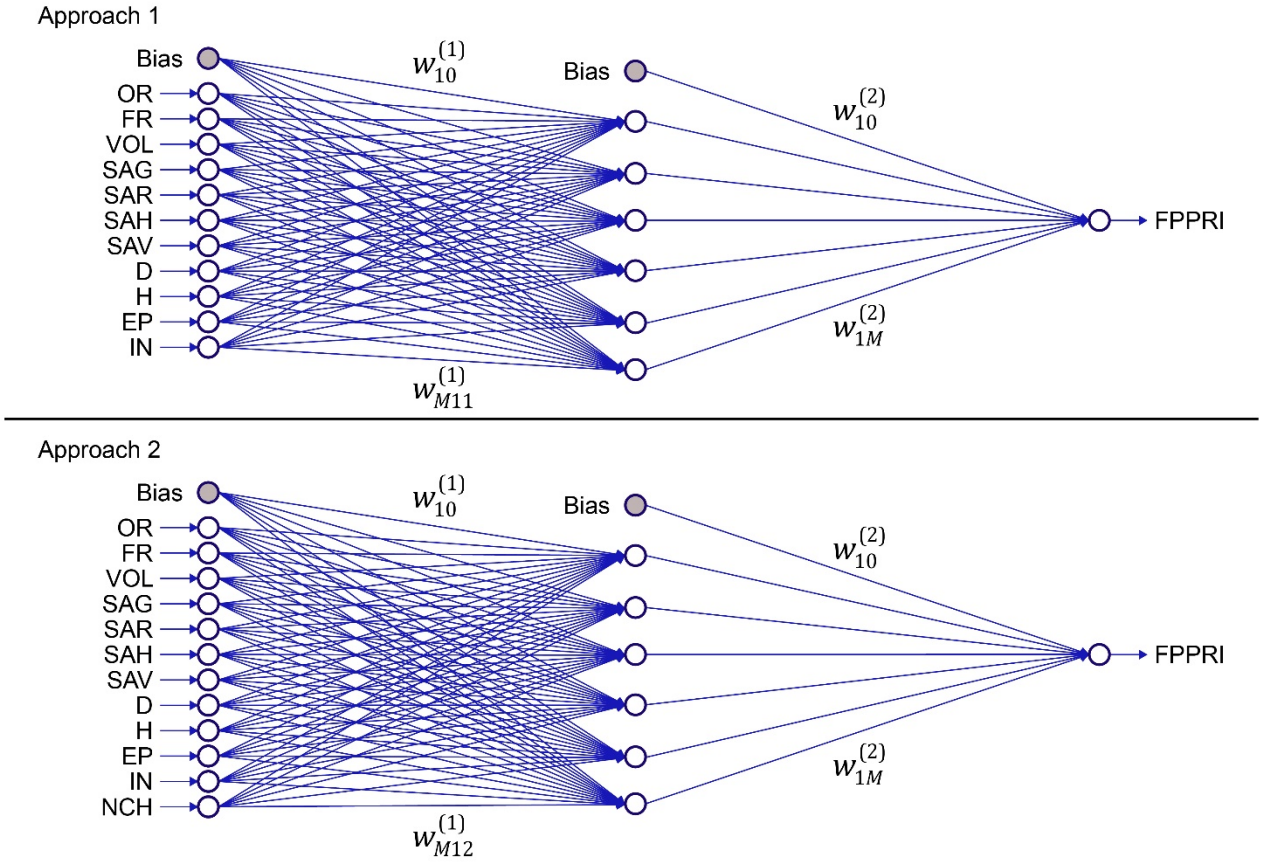

Fig. 5. Scheme of both approaches of MLPs.

Therefore, two types of datasets were developed according to the input variables of each model. At first, 18 datasets were obtained for each approach (5 datasets per decile were carried out for each city, and 1 dataset including all deciles). Each decile was randomly divided into two sets: 75% of instances corresponded to the training dataset, and the remaining 25% to the testing dataset. However, the feasibility of combining data from the different climates was analysed in the training of each model. For this reason, the training datasets of the 3 climates were combining, thereby generating a total of 42 training datasets for each approach (i.e., a total of 84 MLPs were analysed). Likewise, the combined training datasets had all the instances of the training datasets of each climate to perform the training with the same case studies for each climate. For this reason, the number of instances in the training dataset is larger in the combined models (MLP_{C-S}, MLP_{C-V}, MLP_{S-V}, and MLP_{C-S-V}) than in the individual models of each climate (MLP_C, MLP_S, and MLP_V). To provide a clearer explanation of the types of the MLPs developed according to the training data, a further explanation is given in Table 1. It is worth highlighting that a 10-fold cross validation was carried out for the training of the MLPs. The cross validation allows the bias and the variance of the MLP to be reduced [66]. All training datasets were randomly divided into 10 subsets. For each fold, 9 subsets were used for the training, and the remaining subset for the validation of the model. The performance of the model in the training is obtained by the average value of the 10 folds.

Table 1. MLPs and training datasets used.

Model	Submodel	Training dataset	Number of instances in the training dataset
MLP_C	MLP_C (D_1)	Concepción – decile 1	5,832
	MLP_C (D_2)	Concepción – decile 2	5,832
	MLP_C (D_3)	Concepción – decile 3	5,832
	MLP_C (D_4)	Concepción – decile 4	5,832
	MLP_C (D_5)	Concepción – decile 5	5,832
	MLP_C (D1)	Concepción – all deciles	29,160
MLP_S	MLP_S (D_1)	Santiago – decile 1	5,832
	MLP_S (D_2)	Santiago – decile 2	5,832
	MLP_S (D_3)	Santiago – decile 3	5,832
	MLP_S (D_4)	Santiago – decile 4	5,832
	MLP_S (D_5)	Santiago – decile 5	5,832
	MLP_S (Full)	Santiago – all deciles	29,160
MLP_V	MLP_V (D_1)	Valparaiso – decile 1	5,832
	MLP_V (D_2)	Valparaiso – decile 2	5,832
	MLP_V (D_3)	Valparaiso – decile 3	5,832
	MLP_V (D_4)	Valparaiso – decile 4	5,832
	MLP_V (D_5)	Valparaiso – decile 5	5,832
	MLP_V (Full)	Valparaiso – all deciles	29,160
MLP_{C-S}	$MLP_{C:S}$ (D_1)	Concepción and Santiago – decile 1	11,664
	$MLP_{C:S}$ (D_2)	Concepción and Santiago – decile 2	11.664
	$MLP_{C:S}$ (D_3)	Concepción and Santiago – decile 3	11,664
	$MLP_{C:S}$ (D_4)	Concepción and Santiago – decile 4	11,664
	$MLP_{C:S}$ (D_5)	Concepción and Santiago – decile 5	11,664
	MLP_{C-S} (Full)	Concepción and Santiago – all deciles	58,320
MLP_{C-V}	MLP_{C-V} (D_1)	Concepción and Valparaiso – decile 1	11,664
	MLP_{C-V} (D_2)	Concepción and Valparaiso – decile 2	11,664
	MLP_{C-V} (D_3)	Concepción and Valparaiso – decile 3	11,664
	MLP_{C-V} (D_4)	Concepción and Valparaiso – decile 4	11,664
	MLP_{C-V} (D_5)	Concepción and Valparaiso – decile 5	11,664
	MLP_{C-V} (Full)	Concepción and Valparaiso – all deciles	58,320
MLP_{S-V}	$MLP_{S:V}$ (D_1)	Santiago and Valparaiso – decile 1	11,664
	$MLP_{S:V}$ (D_2)	Santiago and Valparaiso – decile 2	11,664
	$MLP_{S:V}$ (D_3)	Santiago and Valparaiso – decile 3	11,664
	$MLP_{S:V}$ (D_4)	Santiago and Valparaiso – decile 4	11,664
	$MLP_{S:V}$ (D_5)	Santiago and Valparaiso – decile 5	11,664
	MLP_{S-V} (Full)	Santiago and Valparaiso – all deciles	58,320
MLP_{C-S-V}	MLP_{C-S-V} (D_1)	Concepción, Santiago and Valparaiso – decile 1	17,496
	MLP_{C-S-V} (D_2)	Concepción, Santiago and Valparaiso – decile 2	17,496
	MLP_{C-S-V} (D_3)	Concepción, Santiago and Valparaiso – decile 3	17,496
	MLP_{C-S-V} (D_4)	Concepción, Santiago and Valparaiso – decile 4	17,496
	MLP_{C-S-V} (D_5)	Concepción, Santiago and Valparaiso – decile 5	17.496
	MLP_{C-S-V} (Full)	Concepción, Santiago and Valparaiso – all deciles	87,480

After training the models, their performances in new instances were assessed. As mentioned above, the testing dataset was obtained from 25% of the instances of the datasets of each climate: such instances were randomly selected. Each MLP was tested in new instances of the 3 climates analysed in this study. The type of testing dataset varied depending on whether the model was trained with separate deciles or the full set. In this sense, it is worth noting that combined testing datasets were not used, unlike in the training of the models: each model was analysed using the same testing subset (deciles or full) as the other models.

To assess the performance of the models developed, three statistical parameters were used: (i) the correlation coefficient (R^2) (Eq. 5), the root mean square error ($RMSE$) (Eq. 6), and the mean absolute error (MAE) (Eq. 7). The use of such parameters allows the performance of the models to be correctly defined. Quality indicators of the model were when the value for R^2 was greater than 0.95, and when $RMSE$ and MAE were as low as possible [40].

$$R^2 = \left(1 - \frac{\sum_{i=1}^{n}(t_i - m_i)^2}{\sum_{i=1}^{n}(t_i - \bar{t}_i)^2}\right) \cdot 100 \quad [\%] \tag{5}$$

$$RMSE = \left(\frac{\sum_{i=1}^{n}(t_i - m_i)^2}{n}\right)^{1/2} \tag{6}$$

$$MAE = \frac{\sum_{i=1}^{n}|t_i - m_i|}{n} \tag{7}$$

Where m_i is the MLP's prediction, t_i is the actual value, and n is the number of instances in the training or testing dataset.

3. Results and discussion

3.1. Results of approach 1 (without NCH as an input variable)

Firstly, the results obtained with the original architecture of the model were analysed. As indicated in Section 2, individual models were trained for each decile, as well as models with all datasets of the deciles. Likewise, the combination of data of the various cities was studied to analyse the possibility of generalization of the MLP. The performance obtained by each model in the training and testing phases are indicated in Tables 2 and 3, respectively. The values obtained in the statistical parameters in the training phase show that the performance obtained by the MLPs with a training dataset of an only climate obtained an acceptable degree of adjustment for training data, whereas the performance obtained by the models with combined training datasets presented limitations (e.g., $MLP_{C\text{-}S}$). In this way, R^2 was in all cases greater than 97.9% in individual models, whereas R^2 presented decreases greater than 10.26% in the combined models with respect to the lowest value of the individual models. Only models of $MLP_{C\text{-}V}$ had values for R^2 greater than 95%. This was due to the similarity presented by climates CL and SL, which contributed to obtain a better performance in the model. However, the dataset of Santiago presented great differences with the other two climates, implying that the combined models trained with data of this climate ($MLP_{C\text{-}S}$, $MLP_{S\text{-}V}$, and $MLP_{C\text{-}S\text{-}V}$) had a low R^2.

Table 2. Results of the training of the models using approach 1.

Model	Submodel and Dataset	R^2 [%]	MAE	$RMSE$
MLP_C	MLP_C (D_1)	98.53	0.027	0.033
	MLP_C (D_2)	98.16	0.015	0.025
	MLP_C (D_3)	98.78	0.009	0.013
	MLP_C (D_4)	98.62	0.008	0.011
	MLP_C (D_5)	98.51	0.006	0.009
	MLP_C (D1)	98.92	0.010	0.017
MLP_S	MLP_S (D_1)	98.16	0.032	0.040
	MLP_S (D_2)	97.91	0.011	0.017
	MLP_S (D_3)	98.13	0.007	0.011
	MLP_S (D_4)	98.29	0.005	0.008
	MLP_S (D_5)	98.33	0.004	0.006
	MLP_S (Full)	98.40	0.008	0.013
MLP_V	MLP_V (D_1)	98.17	0.030	0.037
	MLP_V (D_2)	98.50	0.013	0.020
	MLP_V (D_3)	98.61	0.008	0.013
	MLP_V (D_4)	98.43	0.007	0.011
	MLP_V (D_5)	98.39	0.006	0.009
	MLP_V (Full)	98.67	0.011	0.017
$MLP_{C\text{-}S}$	$MLP_{C:S}$ (D_1)	83.23	0.113	0.125
	$MLP_{C:S}$ (D_2)	79.21	0.049	0.073

	$MLP_{C:S}$ (D_3)	70.13	0.037	0.059
	$MLP_{C:S}$ (D_4)	57.15	0.036	0.054
	$MLP_{C:S}$ (D_5)	73.11	0.023	0.035
	MLP_{C-S} (Full)	84.96	0.036	0.055
MLP_{C-V}	MLP_{C-V} (D_1)	97.81	0.033	0.040
	MLP_{C-V} (D_2)	98.04	0.017	0.025
	MLP_{C-V} (D_3)	97.48	0.013	0.019
	MLP_{C-V} (D_4)	96.86	0.011	0.016
	MLP_{C-V} (D_5)	97.92	0.007	0.011
	MLP_{C-V} (Full)	96.98	0.018	0.027
MLP_{S-V}	$MLP_{S:V}$ (D_1)	91.22	0.070	0.088
	$MLP_{S:V}$ (D_2)	77.87	0.045	0.070
	$MLP_{S:V}$ (D_3)	75.16	0.032	0.051
	$MLP_{S:V}$ (D_4)	80.70	0.023	0.034
	$MLP_{S:V}$ (D_5)	90.02	0.013	0.020
	MLP_{S-V} (Full)	94.92	0.019	0.030
MLP_{C-S-V}	MLP_{C-S-V} (D_1)	86.08	0.100	0.111
	MLP_{C-S-V} (D_2)	86.54	0.037	0.060
	MLP_{C-S-V} (D_3)	76.04	0.032	0.054
	MLP_{C-S-V} (D_4)	87.86	0.020	0.030
	MLP_{C-S-V} (D_5)	85.85	0.017	0.026
	MLP_{C-S-V} (Full)	82.65	0.037	0.060

After analysing the performance obtained by each model in the training, its performance was assessed in new instances. Table 3 indicates the performance in the prediction obtained by the model trained with the total dataset (Full) as well as by the combination of the individual models (D_{1-5}). In the latter, the performance is obtained from the estimation carried out by each model per decile in the testing subset of its decile, combining all estimations in a unique dataset where their statistical parameters are analysed. For further information on the performance in the testing of each decile, Appendix A includes the statistical parameters obtained. Also, the point clouds between the actual values and the predicted values in all the case studies are represented. As can be seen, the estimation of the models trained with a climate presented an acceptable degree of adjustment in estimations carried out in the same climate: in the different models, R^2 was greater than 98.2%, whereas MAE and $RMSE$ were lower than 0.008 and 0.014, respectively. There were not great differences between using individual models per deciles (D_{1-5}) or a full model for all deciles (Full), except a better estimation in the latter (R^2 increased between 0.20 and 0.42% in estimations performed in the same day). However, these models had limitations when estimations of other climates were performed. In this sense, estimations could reach a R^2 of 28.49% (MLP_C - Santiago (D_{1-5})). Only in climates with similar characteristics (Concepción and Valparaiso), the estimation carried out by the model had a good degree of adjustment, with values for R^2 ranging between 94.78 and 97.82. However, because of the limitations presented by the model to estimate the FPPRI in other climates (e.g., NL, SE or AN), it is more advisable to perform estimations only with the models generated by each climate. So, the models generated by combining data of the different climates (MLP_{C-S}, MLP_{C-V}, MLP_{S-V}, and MLP_{C-S-V}) did not improve the estimations generated. The estimations carried out by these models had a tendency similar to the individual models of each climate, with estimations closer to the actual values in climates which were used to train the models. However, except in some models, R^2 in the estimations of the same climates was lower than 95%. Moreover, the estimations in the new climate were not satisfactory, with a very poor performance of the models. Finally, the use of a model trained with data from all climates did not allow an acceptable level of performance to be reached. The purpose of studying this model was due to the potential of having a generic model for all Chile. However, R^2 had values of up to 58.98%, with an MAE of up to 0.056. Therefore, the use of individual models for each climate to perform estimations only in such climate is the best option for approach 1. In this sense, the point clouds between the actual values and the predicted values of the FPPRI presented a greater accuracy in the individual models than in the best option of the combined models (see Appendix A).

Table 3. Results of the testing of the models using approach 1 (the best estimations carried out by each individual model in its climate as well as by the best option of the combined models in each climate are in bold).

Model	Testing	R^2 [%]	MAE	$RMSE$
MLP_C	Concepción (D_{1-5})	98.52	0.012	0.020
	Santiago (D_{1-5})	28.49	0.070	0.088
	Valparaiso (D_{1-5})	97.12	0.018	0.025
	Concepción (Full)	**98.93**	**0.010**	**0.017**
	Santiago (Full)	14.52	0.073	0.100
	Valparaiso (Full)	95.87	0.020	0.031
MLP_S	Concepción (D_{1-5})	63.52	0.075	0.099
	Santiago (D_{1-5})	98.21	0.008	0.014
	Valparaiso (D_{1-5})	71.47	0.062	0.080

	Concepción (Full)	70.65	0.068	0.089
	Santiago (Full)	**98.41**	**0.008**	**0.014**
	Valparaiso (Full)	77.98	0.054	0.071
MLP_V	Concepción (D_{1-5})	94.78	0.023	0.037
	Santiago (D_{1-5})	60.16	0.054	0.066
	Valparaiso (D_{1-5})	98.31	0.011	0.019
	Concepción (Full)	97.82	0.016	0.024
	Santiago (Full)	38.64	0.065	0.085
	Valparaiso (Full)	**98.71**	**0.011**	**0.017**
MLP_{C-S}	Concepción (D_{1-5})	88.73	0.029	0.055
	Santiago (D_{1-5})	53.02	0.054	0.071
	Valparaiso (D_{1-5})	90.29	0.028	0.046
	Concepción (Full)	81.00	0.051	0.072
	Santiago (Full)	92.13	0.022	0.030
	Valparaiso (Full)	86.89	0.037	0.055
MLP_{C-V}	**Concepción (D_{1-5})**	**98.14**	**0.014**	**0.022**
	Santiago (D_{1-5})	35.99	0.065	0.083
	Valparaiso (D_{1-5})	98.15	0.013	0.020
	Concepción (Full)	95.82	0.024	0.034
	Santiago (Full)	62.31	0.049	0.066
	Valparaiso (Full)	**98.38**	**0.013**	**0.019**
MLP_{S-V}	Concepción (D_{1-5})	60.58	0.075	0.102
	Santiago (D_{1-5})	**96.56**	**0.011**	**0.019**
	Valparaiso (D_{1-5})	80.12	0.048	0.066
	Concepción (Full)	68.88	0.066	0.092
	Santiago (Full)	94.87	0.015	0.024
	Valparaiso (Full)	94.75	0.022	0.035
MLP_{C-S-V}	Concepción (D_{1-5})	92.52	0.028	0.045
	Santiago (D_{1-5})	58.98	0.049	0.067
	Valparaiso (D_{1-5})	95.14	0.020	0.033
	Concepción (Full)	74.28	0.056	0.084
	Santiago (Full)	93.63	0.020	0.027
	Valparaiso (Full)	84.83	0.037	0.059

3.2. Results of approach 2 (with NCH as an input variable)

After analysing the performance of the MLPs of approach 1, the performance of the models with the input variable NCH were assessed. As in approach 1, the performance obtained in the training (Table 4) and testing (Table 5) are indicated. The incorporation of the new input variable significantly improved the degree of adjustment obtained in the training phase, the values for R² were modified between -0.8% and 42.62%, and MAE and RMSE were modified up to 0.106 and 0.112 with respect to the models of Table 2. In this sense, individual models per climate obtained a R^2 greater than 99.35%, and MAE and $RMSE$ were lower than 0.012 and 0.015, respectively. Likewise, the models generated with the combinations of data from the different climates had a good performance in the training phase, obtaining the worst values in the MLPs trained with the three climates: the value for R^2 was lower than 96.67% (MLP_{C-S-V} (D_4)), whereas the greatest values for MAE and $RMSE$ were 0.036 and 0.045 (MLP_{C-S-V} (D_1)), respectively.

Table 4. Results of the training of the models using approach 2.

Model	Submodel and Dataset	R^2 [%]	MAE	$RMSE$
MLP_C	MLP_C (D_1)	99.90	0.007	0.009
	MLP_C (D_2)	99.81	0.004	0.008
	MLP_C (D_3)	99.88	0.003	0.004
	MLP_C (D_4)	99.82	0.003	0.004
	MLP_C (D_5)	99.84	0.002	0.003
	MLP_C (D1)	99.92	0.003	0.005
MLP_S	MLP_S (D_1)	99.74	0.012	0.015
	MLP_S (D_2)	99.73	0.003	0.006
	MLP_S (D_3)	99.80	0.002	0.004
	MLP_S (D_4)	99.35	0.003	0.005
	MLP_S (D_5)	99.75	0.001	0.003
	MLP_S (Full)	99.80	0.002	0.005
MLP_V	MLP_V (D_1)	99.90	0.007	0.008
	MLP_V (D_2)	99.88	0.003	0.006
	MLP_V (D_3)	99.72	0.003	0.006
	MLP_V (D_4)	99.79	0.002	0.004

	MLP_V (D_5)	99.85	0.002	0.003
	MLP_V (Full)	99.89	0.003	0.005
MLP_{C-S}	$MLP_{C:S}$ (D_1)	99.82	0.010	0.013
	$MLP_{C:S}$ (D_2)	99.75	0.004	0.008
	$MLP_{C:S}$ (D_3)	99.77	0.003	0.005
	$MLP_{C:S}$ (D_4)	99.77	0.003	0.004
	$MLP_{C:S}$ (D_5)	99.79	0.002	0.003
	MLP_{C-S} (Full)	99.84	0.003	0.006
MLP_{C-V}	MLP_{C-V} (D_1)	99.08	0.020	0.026
	MLP_{C-V} (D_2)	99.03	0.011	0.017
	MLP_{C-V} (D_3)	97.40	0.013	0.019
	MLP_{C-V} (D_4)	99.27	0.005	0.008
	MLP_{C-V} (D_5)	99.17	0.005	0.007
	MLP_{C-V} (Full)	99.13	0.010	0.015
MLP_{S-V}	$MLP_{S:V}$ (D_1)	99.42	0.017	0.023
	$MLP_{S:V}$ (D_2)	98.96	0.008	0.015
	$MLP_{S:V}$ (D_3)	98.25	0.007	0.014
	$MLP_{S:V}$ (D_4)	99.02	0.004	0.008
	$MLP_{S:V}$ (D_5)	98.96	0.004	0.006
	MLP_{S-V} (Full)	98.98	0.007	0.013
MLP_{C-S-V}	MLP_{C-S-V} (D_1)	97.70	0.036	0.045
	MLP_{C-S-V} (D_2)	97.48	0.016	0.026
	MLP_{C-S-V} (D_3)	97.68	0.011	0.017
	MLP_{C-S-V} (D_4)	96.67	0.009	0.016
	MLP_{C-S-V} (D_5)	98.77	0.005	0.008
	MLP_{C-S-V} (Full)	98.85	0.009	0.016

Despite the good performance obtained by the MLPs of approach 2 in the training phase, the performance in testing presented a tendency similar to that obtained in approach 1: the models of each climate generally carried out better estimations than the combined models. In this sense, R^2 oscillated between 99.70 and 99.92 in the estimations performed in the same climate by MLP_C, MLP_S, and MLP_V, whereas MAE and $RMSE$ oscillated between 0.002-0.004 and between 0.005-0.007, respectively. In addition, full models could generalize better than the individual models (D_{1-5}). Also, these models carried out better estimations in the other climates than in the approach 1, thus modifying the tendency of accuracy of the models with respect to approach 1: the MLPs with data of Concepción or Santiago carried out acceptable estimations in the other climate, whereas estimations in Valparaiso presented a lower degree of adjustment. Regarding the models combined with two climates ((MLP_{C-S}, MLP_{C-V}, and MLP_{S-V}), there were limitations to estimate the FPPRI in the climate not included in the training, reaching a correlation of 9.38%, with a MAE of up to 0.117. Under this circumstance, the generation of the model with data of all climates (MLP_{C-S-V}) allowed acceptable estimations in all climates to be performed, with R^2 being greater than 96%. However, the performance obtained by MLP_{C-S-V} was lower than that of individual models of each climate (MLP_C, MLP_S, and MLP_V), with a decrease of R^2 between 0.59 and 3.45%. As in approach 1, the use of individual models is more adequate than the combined models.

Table 5. Results of the testing of the models using approach 2 (the best estimations carried out by each individual model in its climate as well as by the best option of the combined models in each climate are in bold).

Model	Testing	R^2 [%]	MAE	$RMSE$
MLP_C	Concepción (D1-5)	99.81	0.004	0.007
	Santiago (D1-5)	98.05	0.010	0.015
	Valparaiso (D1-5)	91.65	0.032	0.043
	Concepción (Full)	**99.92**	**0.003**	**0.005**
	Santiago (Full)	97.96	0.010	0.015
	Valparaiso (Full)	91.95	0.032	0.043
MLP_S	Concepción (D1-5)	97.75	0.015	0.024
	Santiago (D1-5)	99.70	0.003	0.006
	Valparaiso (D1-5)	91.99	0.028	0.042
	Concepción (Full)	98.90	0.012	0.017
	Santiago (Full)	**99.80**	**0.002**	**0.005**
	Valparaiso (Full)	92.80	0.029	0.041
MLP_V	Concepción (D1-5)	80.18	0.053	0.073
	Santiago (D1-5)	91.87	0.018	0.030
	Valparaiso (D1-5)	99.87	0.003	0.005
	Concepción (Full)	78.81	0.055	0.076
	Santiago (Full)	91.97	0.019	0.031
	Valparaiso (Full)	**99.88**	**0.003**	**0.005**
MLP_{C-S}	Concepción (D1-5)	99.79	0.004	0.008
	Santiago (D1-5)	99.71	0.003	0.006

	Valparaiso (D1-5)	92.20	0.031	0.042
	Concepción (Full)	**99.85**	**0.004**	**0.006**
	Santiago (Full)	**99.78**	**0.003**	**0.005**
	Valparaiso (Full)	92.35	0.031	0.042
MLP_{C-V}	Concepción (D1-5)	99.19	0.009	0.015
	Santiago (D_{1-5})	73.62	0.039	0.054
	Valparaiso (D_{1-5})	98.92	0.010	0.015
	Concepción (Full)	99.43	0.009	0.012
	Santiago (Full)	81.34	0.034	0.047
	Valparaiso (Full)	98.72	0.012	0.017
MLP_{S-V}	Concepción (D_{1-5})	15.19	0.097	0.150
	Santiago (D_{1-5})	98.54	0.006	0.013
	Valparaiso (D_{1-5})	99.11	0.008	0.014
	Concepción (Full)	9.38	0.117	0.173
	Santiago (Full)	97.86	0.008	0.016
	Valparaiso (Full)	**99.44**	**0.006**	**0.011**
MLP_{C-S-V}	Concepción (D_{1-5})	96.56	0.016	0.030
	Santiago (D_{1-5})	96.25	0.009	0.020
	Valparaiso (D_{1-5})	97.18	0.016	0.025
	Concepción (Full)	99.33	0.008	0.014
	Santiago (Full)	98.85	0.007	0.012
	Valparaiso (Full)	98.22	0.011	0.020

3.3. Approach 1 vs approach 2

After analysing their performance, both approaches were compared. For this purpose, Fig. 6 represents radar charts of the values of the statistical parameters obtained in the testing phase. The incorporation of the input variable NCH in the models increased the correlation coefficient of almost all cases, with decreases of MAE and $RMSE$. However, there were exceptions: (i) the estimation of the FPPRI in Valparaiso carried out by MLP_C; and (ii) the estimations of FPPRI in Concepción carried out by MLP_V and MLP_{S-V}. Moreover, the differences between both approaches in the performance of the estimations performed by individual models of each climate in their own climate were low, with an increase of R^2 between 1.00 and 1.59%. There was a significant improvement only in the error parameters, with decreases of up to 0.008 in MAE and 0.013 in $RMSE$. However, the values of the error parameters obtained in these models of approach 1 continued to be low (Table 3). Both approaches could therefore be used as tools to estimate the FPPRI in the different climates. To determine the greatest potential of using some of the approaches, the influence of the input variables on the model was analysed, as well as the cost of obtaining their values.

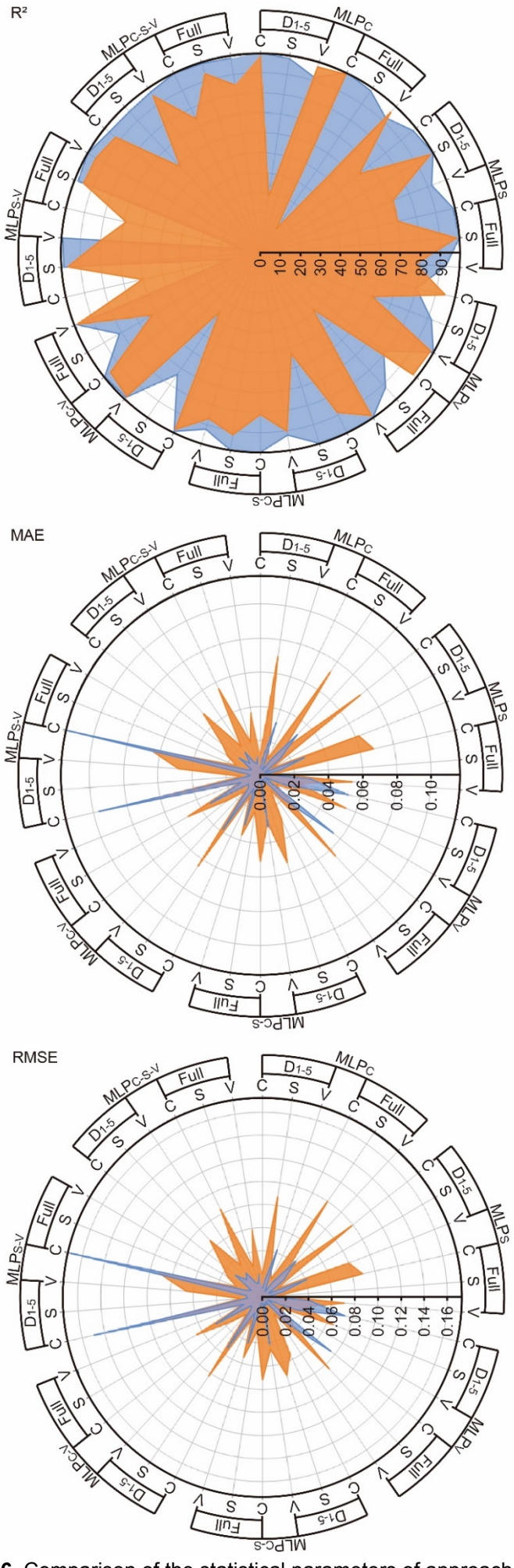

Fig. 6. Comparison of the statistical parameters of approach 1 (orange) and approach 2 (blue).

The influence of the input variables on the models is shown in Table 6. The effect on one of the models is indicated to speed up the reading, since the tendency was the same in the other models. By removing one of the input variables, the effect generated in the estimations carried out by the models can be appreciated, thus decreasing its performance. In approach 1, the variable of income had the highest influence on the performance of the model, with a percentage deviation in R^2 of 89.44% and increases in the error parameters greater than 810%. The remaining input variables presented similar contributions in the performance of the model, highlighting the decrease generated in R^2 by removing SAV, as well as the increase in MAE and $RMSE$ by removing SAG, SAV, and D. Regarding approach 2, the variable of income obtained the highest influence on the performance of the model. NCH (number of comfort hours) influenced the performance of the MLPs similar to that of SAV (variable vertical surface area in contact with other homes). The remaining input variables of approach 2 had a low influence on the performance of the model, although such variables contributed to obtain adjusted results.

As can be seen in both approaches, and except the variable of income, the removal of some input variables did not significantly influence the performance of the model. This aspect becomes important if the time required to obtained values of the input variables is analysed. OR, FR, VOL, SAG, SAR, SAH, SAV, D, and H are morphological aspects of the social building which are easily obtained using both data of the building's project and in situ measurements. Likewise, EP can be obtained based on data from the supplier companies (e.g., the General Electricity Company of Chile) and IN of the family incomes to which a social dwelling should be allocated. The variables of approach 1 are therefore variables easily obtained by the technical personnel responsible for the allocation. NCH also constitutes an input variable whose values can be estimated, although with some limitations with respect to the other input variables. This is because NH is an input variable whose value should be obtained by simulating the building in order to obtain the internal temperature and assess it with respect to the upper and lower limits obtained with the running mean temperature. Hence, the time required to estimate its value is greater than that used in other variables, consequently slowing down the process of determining the FPPRI. Likewise, the possibility of determining such variable in situ would have limitations of use, since the temperatures of dwellings should be monitored for a year before being allocated.

Because of the temporal cost implied by acquiring the NCH and the degree of accuracy obtained by individual models of approach 1, the greatest potential of using the model of approach 1 is therefore determined to carry out estimations of the FPPRI in each climate.

Table 6. Influence of the input variables on the quality indicators (the results obtained in the testing of the data of Concepción using MLP_C).

Model	Input variable removed	R^2 [%]	Deviation percentage of R^2 [%]	MAE	Deviation percentage of MAE [%]	$RMSE$	Deviation percentage of $RMSE$ [%]
Approach 1	OR	97.65	-1.29	0.016	60.00	0.025	47.06
	FR	98.74	-0.19	0.012	20.00	0.018	5.88
	V	98.64	-0.29	0.010	0.00	0.019	11.76
	SAG	98.93	0.00	0.028	180.00	0.035	105.88
	SAR	98.70	-0.23	0.011	10.00	0.019	11.76
	SAH	98.88	-0.05	0.010	0.00	0.017	0.00
	SAV	87.38	-11.67	0.038	280.00	0.059	247.06
	D	98.35	-0.59	0.035	250.00	0.044	158.82
	H	97.51	-1.44	0.017	70.00	0.026	52.94
	EP	98.20	-0.74	0.014	40.00	0.022	29.41
	IN	10.45	-89.44	0.091	810.00	0.156	817.65
Approach 2	OR	99.83	-0.09	0.004	33.33	0.007	40.00
	FR	99.75	-0.17	0.005	66.67	0.008	60.00
	V	99.89	-0.03	0.003	0.00	0.006	20.00
	SAG	99.92	0.00	0.006	100.00	0.008	60.00
	SAR	99.88	-0.04	0.003	0.00	0.006	20.00
	SAH	99.88	-0.04	0.003	0.00	0.006	20.00
	SAV	98.92	-1.00	0.010	233.33	0.017	240.00
	D	99.89	-0.03	0.003	0.00	0.005	0.00
	H	99.87	-0.05	0.003	0.00	0.006	20.00
	EP	99.18	-0.74	0.025	733.33	0.031	520.00
	IN	11.45	-88.54	0.090	2,900.00	0.155	3,000.00
	NCH	98.93	-0.99	0.010	233.33	0.017	240.00

3.4. Estimation methodology of FPPRI for social housing allocation

Given that the models developed with approach 1 presented limitations when FPPRI was estimated in climates not included in its training, it is necessary to carry out models for each type of climate for their correct implementation. Likewise, only a full model for each climate is required because the performance obtained by the individual models per deciles and by models for all of them was almost identical. To obtain the training dataset, 29,160 simulations (instances) should be performed (corresponding to 75% of the full dataset of each climate used in this research). The dataset should include the input variables of the MLPs (i.e., OR, FR, V, SAG, SAR, SAH, SAV, D, H, EP, and IN), as well as the output variable of the models (FPPRI). When assessing the allocation of families in social housing, the input variables corresponding to the morphology of the building should be assessed, as well as the energy price and incomes of the family to be allocated. The input vector generated for each case study will be introduced in the model to estimate the FPPRI. If the estimation is lower than 10%, the combination of the characteristics of the dwelling and the family does not suffer from risk of fuel poverty, so such social dwelling can be allocated to the family. If the value is greater than 10%, the variables of the building should be reassessed until obtain the most adequate configuration with the family. Fig. 7 shows the procedure of generating the model of a climate and its use in a real case of social housing allocation.

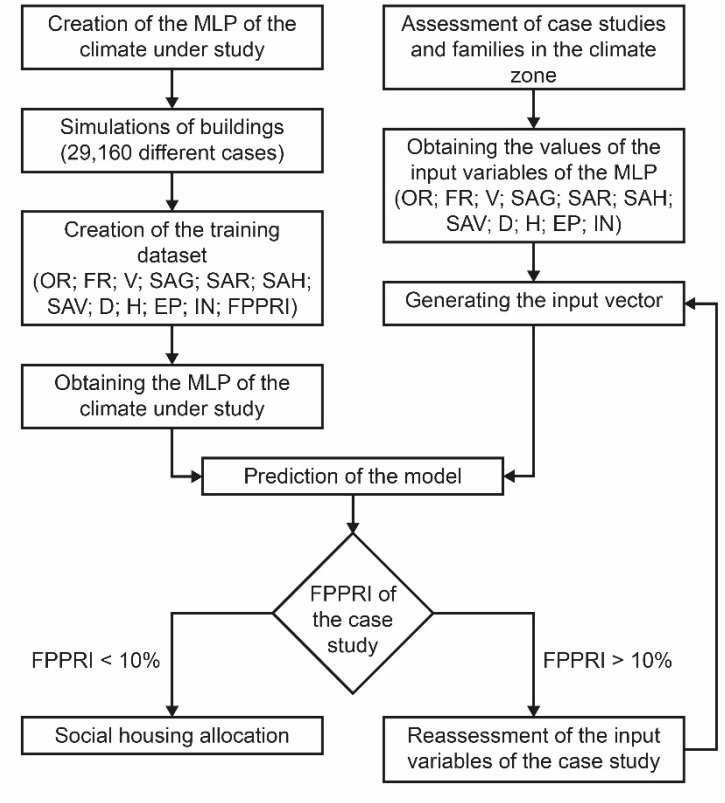

Fig. 7. Flowchart of the methodology proposed for social housing allocation to reduce the risk of fuel poverty.

4. Conclusions

The roadmap of the Chilean government “Energy 2050” establishes as 1 out of the 38 main objectives of the country to promote adequate specific measurements to avoid cases of fuel poverty. One of the possibilities to reduce cases of fuel poverty is the smart allocation of low-income families to households adapted to their limitations. For this purpose, the use of artificial intelligence would provide prediction models assessing the best combination for the different families obtaining the lowest fuel poverty potential risk index (FPPRI). The use of such index has many advantages to predict better the real behaviour of the low-income families in order to have conditions of thermal comfort inside their dwellings as the adaptive comfort is used.

The main objective of this research was therefore the analysis of the potentials and limitations of the multilayer perceptrons to estimate the FPPRI in various climate zones in Chile as well as to allocate social housing to families with low economic resources correctly. A total of 38,880 case studies have been generated in three different climate zones of the country (i.e., a total of 166,640 case studies) Such case studies corresponded to both the 5 deciles of the most vulnerable families of the country and the two most numerous typologies of social housing of new construction since 2005.

The results showed the most adequate methodological approach to use prediction models of FPPRI in social housing located in other climate zones. The use of models typical of each climate allowed estimations adjusted to the

actual values of FPPRI to be carried out in the case studies located in such climate, with a correlation coefficient greater than 98%. Likewise, the use of full models for the five poorest types of families obtained predictions of the same quality as individual models of each type of family. However, there were limitations of generalization of the models for the case studies of other climates. The most effective methodology to estimate the FPPRI of all types of the most vulnerable families is therefore the use of individual models for each climate. Also, the most adequate vector of input variables for the flexible implementation of the methodology was that without the variable of the number of comfort hours.

The potential of using this methodology to allocate social housing would guarantee the main goal of the country, that is, reducing the fuel poverty in the roadmap for 2050. By means of prediction models, regional and provincial governments can have simple and flexible tools of assessment. The use of input variables of architectural characteristics of buildings of social dwellings, the energy price, and family incomes are aspects easily to be analysed by the technical or administrative personnel responsible for social housing allocation. Likewise, technicians responsible for designing new social housing projects can use these prediction models to adapt the morphological characteristics of the building to the poorest families. It is worth highlighting that users should be aware of the need of using adaptively air conditioning systems for the model and the application of FPPRI to work. For this reason, various political estates should promote education and awareness plans for the poorest families before allocating social housing.

To conclude, the results of this research has defined the most effective methodology of implementing prediction models of the FPPRI in all regions of Chile. Likewise, the accuracy achieved by the estimations of the model implies the implementation of the methodology of social housing allocation in other countries.

Acknowledgements

The authors would like to acknowledge “Dirección de Investigación y Creación Artística” and “Dirección de Postgrado” at the University of the Bío-Bío for financing this paper. In addition, the authors would like to acknowledge “Becas Iberoamericanas. Jóvenes Profesionales e Investigadores. Santander Universidades” for financing the international mobility of Alexis Pérez-Fargallo at the University of Cadiz.

References

[1] World Wildlife Fund, Living Planet Report 2014: Species and spaces, people and places, WWF International, Gland, Switzerland, 2014. doi:10.1007/s13398-014-0173-7.2.

[2] The United Nations Environment Programme, Building Design and Construction: Forging Resource Efficiency and Sustainable, Nairobi, Kenya, 2012.

[3] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy Build. 40 (2008) 394–398. doi:10.1016/j.enbuild.2007.03.007.

[4] H. Thomson, C. Snell, C. Liddell, Fuel poverty in the European Union: a concept in need of definition?, People, Place Policy Online 2016. 10/1 (2016) 5–24. doi:10.3351/ppp.0010.0001.0002.

[5] E. Rehfuess, Fuel for life: household energy and health, Geneva World Heal. Organ. (2006) 1–23.

[6] M. González-Eguino, Energy poverty: An overview, Renew. Sustain. Energy Rev. 47 (2015) 377–385. doi:10.1016/j.rser.2015.03.013.

[7] G. Thompson, M. Bazilian, Democratization, energy poverty, and the pursuit of symmetry, Glob. Policy. 5 (2014) 127–131. doi:10.1111/1758-5899.12103.

[8] S.C. Robert, L. Koh, R. Marchand, A. Genovese, A. Brennan, Fuel Poverty. Perspectives from the front line, Sheffield, 2012. https://www.sheffield.ac.uk/polopoly_fs/1.272226!/file/Fuel_Poverty-perspectives_from_the_front_line.pdf.

[9] Department of health and human services (U.S.), LIHEAP Report to Congress for Fiscal Year 2010, Washington, 2010. http://www.acf.hhs.gov/sites/default/files/ocs/fy10_liheap_rtc_final.pdf.

[10] H. Thomson, C. Snell, Quantifying the prevalence of fuel poverty across the European Union, Energy Policy. 52 (2013) 563–572. doi:10.1016/J.ENPOL.2012.10.009.

[11] L. Middlemiss, R. Gillard, Fuel poverty from the bottom-up: Characterising household energy vulnerability through the lived experience of the fuel poor, Energy Res. Soc. Sci. 6 (2015) 146–154.

[12] C. Liddell, C. Morris, H. Thomson, C. Guiney, Excess winter deaths in 30 European countries 1980–2013: a critical review of methods, J. Public Health (Bangkok). 38 (2016) 806–814.

[13] J. Teller-Elsberg, B. Sovacool, T. Smith, E. Laine, Fuel poverty, excess winter deaths, and energy costs in Vermont: Burdensome for whom?, Energy Policy. 90 (2016). doi:10.1016/j.enpol.2015.12.009.

[14] S. Bouzarovski, S. Petrova, A global perspective on domestic energy deprivation: Overcoming the energy poverty–fuel poverty binary, Energy Res. Soc. Sci. 10 (2015) 31–40.

[15] K. Li, B. Lloyd, X.J. Liang, Y.M. Wei, Energy poor or fuel poor: What are the differences?, Energy Policy. 68 (2014) 476–481. doi:10.1016/j.enpol.2013.11.012.

[16] B. Boardman, Fuel Poverty: From Cold Homes to Affordable Warmth, John Wiley & Sons Ltd, London, UK, 1991.

[17] K.C. O'Sullivan, P.L. Howden-Chapman, G.M. Fougere, Fuel poverty, policy, and equity in New Zealand: The promise of prepayment metering, Energy Res. Soc. Sci. 7 (2015) 99–107. doi:10.1016/j.erss.2015.03.008.

[18] B. Legendre, O. Ricci, Measuring fuel poverty in France: Which households are the most fuel vulnerable?, Energy Econ. 49 (2015) 620–628. doi:10.1016/j.eneco.2015.01.022.

[19] M. Santamouris, J.A. Paravantis, D. Founda, D. Kolokotsa, P. Michalakakou, A.M. Papadopoulos, N. Kontoulis, A. Tzavali, E.K. Stigka, Z. Ioannidis, A. Mehilli, A. Matthiessen, E. Servou, Financial crisis and energy consumption: A household survey in Greece, Energy Build. 65 (2013) 477–487. doi:10.1016/j.enbuild.2013.06.024.

[20] J.C. Romero, P. Linares, X. López, The policy implications of energy poverty indicators, Energy Policy. 115 (2018) 98–108. doi:10.1016/j.enpol.2017.12.054.

[21] K. Fabbri, Building and fuel poverty, an index to measure fuel poverty: An Italian case study, Energy. 89 (2015) 244–258. doi:10.1016/j.energy.2015.07.073.

[22] S. Desiere, W. Vellema, M. D'Haese, A validity assessment of the Progress out of Poverty Index (PPI)™, Eval. Program Plann. 49 (2015) 10–18. doi:10.1016/j.evalprogplan.2014.11.002.

[23] P. Nussbaumer, M. Bazilian, V. Modi, Measuring energy poverty: Focusing on what matters, Renew. Sustain. Energy Rev. 16 (2012) 231–243. doi:10.1016/j.rser.2011.07.150.

[24] K. Wang, Y.-X. Wang, K. Li, Y.-M. Wei, Energy poverty in China: An index based comprehensive evaluation, Renew. Sustain. Energy Rev. 47 (2015) 308–323. doi:10.1016/j.rser.2015.03.041.

[25] N. Bonatz, R. Guo, W. Wu, L. Liu, A comparative study of the interlinkages between low carbon development and energy poverty in China and Germany by developing an energy poverty index, Energy Build. 183 (2019) 817–831. doi:https://doi.org/10.1016/j.enbuild.2018.09.042.

[26] S. Attia, S. Carlucci, Impact of different thermal comfort models on zero energy residential buildings in hot climate, Energy Build. 102 (2015) 117–128. doi:10.1016/j.enbuild.2015.05.017.

[27] A. Pérez-Fargallo, C. Rubio-Bellido, J.A. Pulido-Arcas, M. Trebilcock, Development policy in social housing allocation: Fuel poverty potential risk index, Indoor Built Environ. 26 (2017) 980–998. doi:10.1177/1420326X17713071.

[28] European Committee for Standardization, EN 15251:2007 Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor quality, thermal environment, lighting and acoustics, European Committee for Standardization, Brussels, 2007.

[29] Ministerio de Desarrollo Social (Chile), CASEN 2013. Evolución y distribución del ingreso de los hogares (2006-2013), Santiago de Chile, 2015. http://observatorio.ministeriodesarrollosocial.gob.cl/documentos/Casen2013_Evolucion_Distibucion_Ingresos.pdf.

[30] Organisation for Economic Co-operation and Development, Agreement on the terms of accession of The Republic of Chile to the Convention on The Organisation for Economic Co-operation and development, (2010) 66. http://www.oecd.org/chile/44381035.pdf (accessed May 17, 2017).

[31] R. Day, G. Walker, N. Simcock, Conceptualising energy use and energy poverty using a capabilities framework, Energy Policy. 93 (2016) 255–264. doi:10.1016/j.enpol.2016.03.019.

[32] Ministerio de Desarrollo Social (Chile), CASEN 2013. Evolución y distribución del ingreso de los hogares (2006-2013), Santiago de Chile, 2015.

[33] Secretaría Ejecutiva de Desarrollo de Barrios (Chile), Vivienda Social En Copropiedad. Catastro Nacional de condominios sociales., MINVU, Santiago de Chile, 2014.

[34] A. Pérez-Fargallo, C. Rubio-Bellido, J.A. Pulido-Arcas, F. Javier Guevara-García, Fuel Poverty Potential Risk Index in the context of climate change in Chile, Energy Policy. 113 (2018) 157–170. doi:10.1016/j.enpol.2017.10.054.

[35] Organisation for Economic Co-operation and Development, Greening household behaviour. Overview from the 2011 survey - Revised Edition., 2011. doi:Ihttp://dx.doi.org/10.1787/9789264214651-en.

[36] Instituto Nacional de Estadísticas (Chile), Encuesta suplementaria de ingresos 2016, Santiago de Chile, 2016.

[37] S. Palma Sierrra, B. Ebensperger Eguiguren, J. Bustos Salvagno, Energía 2050. Política energética de Chile. Informe de seguimiento 2016., Santiago de Chile, 2016. www.energia2050.cl.

[38] R. Reyes, A. Schueftan, C. Ruiz, A.D. González, Controlling air pollution in a context of high energy poverty levels in southern Chile: Clean air but colder houses?, Energy Policy. 124 (2019) 301–311. doi:10.1016/j.enpol.2018.10.022.

[39] Y. Simsek, Á. Lorca, T. Urmee, P.A. Bahri, R. Escobar, Review and assessment of energy policy developments in Chile, Energy Policy. 127 (2019) 87–101. doi:10.1016/j.enpol.2018.11.058.

[40] R. Pino-Mejías, A. Pérez-Fargallo, C. Rubio-Bellido, J.A. Pulido-Arcas, Artificial neural networks and linear regression prediction models for social housing allocation: Fuel Poverty Potential Risk Index, Energy. 164 (2018) 627–641. doi:10.1016/j.energy.2018.09.056.

[41] S.M.C. Magalhães, V.M.S. Leal, I.M. Horta, Modelling the relationship between heating energy use and indoor temperatures in residential buildings through Artificial Neural Networks considering occupant behavior, Energy Build. 151 (2017) 332–343. doi:10.1016/j.enbuild.2017.06.076.

[42] R. Pino-Mejías, A. Pérez-Fargallo, C. Rubio-Bellido, J.A. Pulido-Arcas, Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO<inf>2</inf> emissions, Energy. 118 (2017). doi:10.1016/j.energy.2016.12.022.

[43] C. Rubio-Bellido, A. Pérez-Fargallo, J.A. Pulido-Arcas, Optimization of annual energy demand in office buildings under the influence of climate change in Chile, Energy. 114 (2016). doi:10.1016/j.energy.2016.08.021.

[44] S.S. Haykin, S.S. Haykin, S.S. Haykin, S.S. Haykin, Neural networks and learning machines, Pearson Upper Saddle River, 2009.

[45] S. Raghu, N. Sriraam, Optimal configuration of multilayer perceptron neural network classifier for recognition of intracranial epileptic seizures, Expert Syst. Appl. 89 (2017) 205–221. doi:10.1016/j.eswa.2017.07.029.

[46] W. Zhou, J. Jia, A learning framework for shape retrieval based on multilayer perceptrons, Pattern Recognit. Lett. 0 (2018) 1–12. doi:10.1016/j.patrec.2018.09.005.

[47] D. Bienvenido-Huertas, J. Moyano, C.E. Rodríguez-Jiménez, D. Marín, Applying an artificial neural network to assess thermal transmittance in walls by means of the thermometric method, Appl. Energy. 233–234 (2019) 1–14. doi:10.1016/j.apenergy.2018.10.052.

[48] A.R. Barron, Universal approximation bounds for superpositions of a sigmoidal function, IEEE Trans. Inf. Theory. 39 (1993) 930–945.

[49] G. Cybenko, Approximation by superpositions of a sigmoidal function, Math. Control. Signals Syst. 2 (1989) 303–314.

[50] K. Hornik, M. Stinchcombe, H. White, Multilayer feedforward networks are universal approximators, Neural Networks. 2 (1989) 359–366. doi:10.1016/0893-6080(89)90020-8.

[51] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning representations by back-propagating errors, Nature. 323 (1986) 533–536. doi:10.1038/323533a0.

[52] Y.N. Wang, A neural network adaptive control based on rapid learning method and application, Adv. Molding Anal. 46 (1994) 27–34.

[53] P. Werbos, Beyond Regression: New Tools for Prediction and Analysis in the Behavior Science, Harvard University, 1974.

[54] R. Fletcher, Practical methods of optimization, John Wiley&Sons, Chichester - New York - Brisbane - Toronto, United States, 1980.

[55] R. Kumar, R.K. Aggarwal, J.D. Sharma, Energy analysis of a building using artificial neural network: A review, Energy Build. 65 (2013) 352–358. doi:10.1016/j.enbuild.2013.06.007.

[56] F. Rubel, M. Kottek, Observed and projected climate shifts 1901-2100 depicted by world maps of the Köppen-Geiger climate classification, Meteorol. Zeitschrift. 19 (2010) 135–141. doi:10.1127/0941-2948/2010/0430.

[57] R.A. Dahlgren, J.L. Boettinger, G.L. Huntington, R.G. Amundson, Soil development along an elevational transect in the western Sierra Nevada, California, Geoderma. 78 (1997) 207–236. doi:10.1016/S0016-7061(97)00034-7.

[58] D.P. Franzmeier, E.J. Pedersen, T.J. Longwell, J.G. Byrne, C.K. Losche, Properties of Some Soils in the Cumberland Plateau as Related to Slope Aspect and Position1, Soil Sci. Soc. Am. J. 33 (1969) 755–761. doi:10.2136/sssaj1969.03615995003300050037x.

[59] C.C. Tsui, Z.S. Chen, C.F. Hsieh, Relationships between soil properties and slope position in a lowland rain forest of southern Taiwan, Geoderma. 123 (2004) 131–142. doi:10.1016/j.geoderma.2004.01.031.

[60] F. Yimer, S. Ledin, A. Abdelkadir, Soil property variations in relation to topographic aspect and vegetation community in the south-eastern highlands of Ethiopia, For. Ecol. Manage. 232 (2006) 90–99. doi:10.1016/j.foreco.2006.05.055.

[61] A. Zurita, A. Castillejo-Cuberos, M. García, C. Mata-Torres, Y. Simsek, R. García, F. Antonanzas-Torres, R.A. Escobar, State of the art and future prospects for solar PV development in Chile, Renew. Sustain. Energy Rev. 92 (2018) 701–727. doi:10.1016/j.rser.2018.04.096.

[62] Instituto Nacional de Normalización (Chile), NCh 1079:2008. Arquitectura y construcción - Zonificación climática habitacional para Chile y recomendaciones para diseño arquitectónico, Santiago de Chile, 2008.

[63] Ministerio de Vivienda y Urbanismo (Chile), DS 47 - ordenanza general de la Ley general de Urbanismo y construcciones, 1992.

[64] Ministerio de Vivienda y Urbanismo (Chile), Base de datos de catastro, Santiago de Chile, 2013. www.minvu.cl/incjs/download.aspx?glb_cod_nodo=20160405115049&hdd_nom_archivo=MINVU - Base Catastro CVS 2013 (simplificada).xlsx (accessed March 1, 2018).

[65] Ministerio de Vivienda y Urbanismo (Chile), Código de construcción sustentable para viviendas en Chile, Santiago de Chile, 2016.

[66] R. Kohavi, A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection, in: Int. Jt. Conf. Artif. Intell., 1995. doi:10.1067/mod.2000.109031.

Appendix A

Table A1. Results of the testing of the MLP_C per deciles (approach 1).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	98.39	0.028	0.035
Santiago (D_1)	7.88	0.115	0.131
Valparaiso (D_1)	96.21	0.029	0.037
Concepción (D_2)	98.15	0.015	0.025
Santiago (D_2)	28.21	0.079	0.097
Valparaiso (D_2)	97.69	0.018	0.025
Concepción (D_3)	98.72	0.009	0.014
Santiago (D_3)	16.38	0.060	0.075
Valparaiso (D_3)	96.72	0.015	0.021
Concepción (D_4)	98.69	0.008	0.011
Santiago (D_4)	2.38	0.053	0.063
Valparaiso (D_4)	94.20	0.016	0.021
Concepción (D_5)	98.22	0.007	0.011
Santiago (D_5)	15.53	0.043	0.052
Valparaiso (D_5)	93.50	0.013	0.018

Table A2. Results of the testing of the MLP_S per deciles (approach 1).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	23.22	0.237	0.241
Santiago (D_1)	98.06	0.013	0.019
Valparaiso (D_1)	63.30	0.098	0.114
Concepción (D_2)	57.03	0.097	0.123
Santiago (D_2)	97.77	0.011	0.017
Valparaiso (D_2)	64.00	0.080	0.098
Concepción (D_3)	55.03	0.067	0.083
Santiago (D_3)	96.91	0.007	0.014
Valparaiso (D_3)	65.16	0.056	0.070
Concepción (D_4)	59.90	0.051	0.061
Santiago (D_4)	98.21	0.005	0.008
Valparaiso (D_4)	68.74	0.041	0.050
Concepción (D_5)	60.40	0.041	0.050
Santiago (D_5)	98.24	0.004	0.006
Valparaiso (D_5)	66.77	0.033	0.040

Table A3. Results of the testing of the MLP_V per deciles (approach 1).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	93.01	0.062	0.073
Santiago (D_1)	56.68	0.079	0.090
Valparaiso (D_1)	97.06	0.021	0.032
Concepción (D_2)	94.98	0.028	0.042
Santiago (D_2)	52.18	0.066	0.079
Valparaiso (D_2)	98.56	0.012	0.020
Concepción (D_3)	95.56	0.017	0.026
Santiago (D_3)	47.58	0.049	0.059
Valparaiso (D_3)	98.49	0.009	0.014
Concepción (D_4)	94.97	0.015	0.022
Santiago (D_4)	50.80	0.038	0.045
Valparaiso (D_4)	97.99	0.007	0.013
Concepción (D_5)	97.32	0.009	0.013
Santiago (D_5)	23.15	0.035	0.043
Valparaiso (D_5)	98.41	0.006	0.009

Table A4. Results of the testing of the $MLP_{C\text{-}S}$ per deciles (approach 1).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	69.91	0.144	0.151
Santiago (D_1)	91.20	0.033	0.041
Valparaiso (D_1)	78.94	0.065	0.087
Concepción (D_2)	97.14	0.019	0.032
Santiago (D_2)	28.80	0.079	0.096
Valparaiso (D_2)	95.91	0.023	0.033

Concepción (D_3)	96.71	0.014	0.022
Santiago (D_3)	1.60	0.060	0.081
Valparaiso (D_3)	93.91	0.018	0.029
Concepción (D_4)	95.07	0.013	0.022
Santiago (D_4)	41.56	0.060	0.076
Valparaiso (D_4)	86.56	0.022	0.032
Concepción (D_5)	96.91	0.008	0.014
Santiago (D_5)	1.69	0.038	0.048
Valparaiso (D_5)	94.36	0.010	0.016

Table A5. Results of the testing of the $MLP_{C\text{-}V}$ per deciles (approach 1).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	97.33	0.037	0.045
Santiago (D_1)	24.40	0.102	0.119
Valparaiso (D_1)	98.08	0.019	0.026
Concepción (D_2)	98.73	0.014	0.021
Santiago (D_2)	16.30	0.085	0.105
Valparaiso (D_2)	97.52	0.018	0.026
Concepción (D_3)	96.62	0.016	0.023
Santiago (D_3)	45.02	0.048	0.060
Valparaiso (D_3)	98.60	0.010	0.014
Concepción (D_4)	98.40	0.008	0.012
Santiago (D_4)	1.15	0.053	0.064
Valparaiso (D_4)	94.96	0.014	0.020
Concepción (D_5)	98.03	0.008	0.011
Santiago (D_5)	17.43	0.036	0.044
Valparaiso (D_5)	97.74	0.007	0.010

Table A6. Results of the testing of the $MLP_{S\text{-}V}$ per deciles (approach 1).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	25.49	0.227	0.238
Santiago (D_1)	95.45	0.019	0.029
Valparaiso (D_1)	87.78	0.050	0.066
Concepción (D_2)	54.58	0.097	0.126
Santiago (D_2)	96.13	0.014	0.023
Valparaiso (D_2)	65.77	0.076	0.096
Concepción (D_3)	49.95	0.068	0.087
Santiago (D_3)	95.48	0.010	0.017
Valparaiso (D_3)	60.79	0.057	0.074
Concepción (D_4)	61.30	0.049	0.060
Santiago (D_4)	97.37	0.006	0.010
Valparaiso (D_4)	70.08	0.039	0.048
Concepción (D_5)	54.00	0.042	0.054
Santiago (D_5)	95.50	0.007	0.010
Valparaiso (D_5)	85.88	0.019	0.026

Table A7. Results of the testing of the $MLP_{C\text{-}S\text{-}V}$ per deciles (approach 1).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	81.82	0.109	0.117
Santiago (D_1)	79.09	0.050	0.063
Valparaiso (D_1)	91.02	0.044	0.057
Concepción (D_2)	97.10	0.024	0.032
Santiago (D_2)	37.25	0.069	0.091
Valparaiso (D_2)	97.72	0.017	0.025
Concepción (D_3)	93.99	0.021	0.030
Santiago (D_3)	0.04	0.055	0.082
Valparaiso (D_3)	91.91	0.019	0.034
Concepción (D_4)	93.22	0.016	0.025
Santiago (D_4)	56.11	0.036	0.042
Valparaiso (D_4)	96.05	0.010	0.018
Concepción (D_5)	96.88	0.010	0.014
Santiago (D_5)	25.49	0.033	0.042
Valparaiso (D_5)	97.10	0.008	0.012

Table A8. Results of the testing of the MLP_C per deciles (approach 2).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	99.88	0.007	0.010
Santiago (D_1)	97.34	0.017	0.022
Valparaiso (D_1)	88.28	0.053	0.065
Concepción (D_2)	99.79	0.005	0.009
Santiago (D_2)	98.51	0.008	0.014
Valparaiso (D_2)	90.77	0.039	0.050
Concepción (D_3)	99.86	0.003	0.005
Santiago (D_3)	97.48	0.009	0.013
Valparaiso (D_3)	90.22	0.029	0.037
Concepción (D_4)	99.81	0.003	0.004
Santiago (D_4)	96.88	0.008	0.011
Valparaiso (D_4)	90.76	0.021	0.027
Concepción (D_5)	99.55	0.002	0.005
Santiago (D_5)	96.92	0.006	0.009
Valparaiso (D_5)	89.24	0.018	0.023

Table A9. Results of the testing of the MLP_S per deciles (approach 2).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	97.99	0.033	0.039
Santiago (D_1)	99.57	0.005	0.009
Valparaiso (D_1)	86.85	0.052	0.068
Concepción (D_2)	97.01	0.022	0.032
Santiago (D_2)	99.74	0.003	0.006
Valparaiso (D_2)	92.30	0.032	0.046
Concepción (D_3)	98.35	0.011	0.016
Santiago (D_3)	99.75	0.002	0.004
Valparaiso (D_3)	91.15	0.026	0.035
Concepción (D_4)	97.90	0.011	0.014
Santiago (D_4)	99.35	0.003	0.005
Valparaiso (D_4)	93.82	0.017	0.022
Concepción (D_5)	97.92	0.008	0.012
Santiago (D_5)	99.76	0.001	0.002
Valparaiso (D_5)	91.45	0.015	0.020

Table A10. Results of the testing of the MLP_V per deciles (approach 2).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	79.01	0.115	0.126
Santiago (D_1)	90.11	0.029	0.043
Valparaiso (D_1)	99.86	0.004	0.007
Concepción (D_2)	79.34	0.065	0.085
Santiago (D_2)	90.99	0.023	0.034
Valparaiso (D_2)	99.89	0.003	0.005
Concepción (D_3)	71.07	0.050	0.066
Santiago (D_3)	87.80	0.018	0.028
Valparaiso (D_3)	99.74	0.003	0.006
Concepción (D_4)	78.11	0.035	0.045
Santiago (D_4)	91.88	0.012	0.018
Valparaiso (D_4)	99.79	0.002	0.004
Concepción (D_5)	76.03	0.030	0.039
Santiago (D_5)	89.56	0.010	0.016
Valparaiso (D_5)	99.86	0.002	0.003

Table A11. Results of the testing of the $MLP_{C\text{-}S}$ per deciles (approach 2).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	99.79	0.010	0.013
Santiago (D_1)	99.63	0.005	0.008
Valparaiso (D_1)	88.47	0.052	0.064
Concepción (D_2)	99.80	0.005	0.008
Santiago (D_2)	99.65	0.004	0.007
Valparaiso (D_2)	91.65	0.038	0.047
Concepción (D_3)	99.77	0.004	0.006
Santiago (D_3)	99.61	0.003	0.005
Valparaiso (D_3)	91.49	0.027	0.034
Concepción (D_4)	99.80	0.003	0.004

Santiago (D_4)	99.69	0.002	0.004
Valparaiso (D_4)	91.52	0.020	0.026
Concepción (D_5)	99.70	0.002	0.004
Santiago (D_5)	99.70	0.002	0.003
Valparaiso (D_5)	90.69	0.017	0.021

Table A12. Results of the testing of the $MLP_{C\text{-}V}$ per deciles (approach 2).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	98.88	0.022	0.029
Santiago (D_1)	59.52	0.072	0.087
Valparaiso (D_1)	99.27	0.011	0.016
Concepción (D_2)	99.35	0.010	0.015
Santiago (D_2)	72.21	0.048	0.060
Valparaiso (D_2)	98.73	0.012	0.019
Concepción (D_3)	97.93	0.011	0.018
Santiago (D_3)	81.27	0.028	0.035
Valparaiso (D_3)	96.59	0.015	0.022
Concepción (D_4)	99.44	0.005	0.007
Santiago (D_4)	65.72	0.031	0.037
Valparaiso (D_4)	99.12	0.005	0.008
Concepción (D_5)	99.30	0.005	0.007
Santiago (D_5)	79.01	0.018	0.022
Valparaiso (D_5)	98.99	0.005	0.007

Table A13. Results of the testing of the $MLP_{S\text{-}V}$ per deciles (approach 2).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	46.63	0.185	0.201
Santiago (D_1)	98.17	0.009	0.018
Valparaiso (D_1)	99.03	0.011	0.019
Concepción (D_2)	0.04	0.129	0.187
Santiago (D_2)	97.86	0.009	0.017
Valparaiso (D_2)	99.33	0.008	0.013
Concepción (D_3)	28.77	0.072	0.104
Santiago (D_3)	98.81	0.004	0.009
Valparaiso (D_3)	97.46	0.011	0.019
Concepción (D_4)	8.80	0.066	0.092
Santiago (D_4)	98.26	0.004	0.008
Valparaiso (D_4)	99.19	0.005	0.008
Concepción (D_5)	21.54	0.050	0.071
Santiago (D_5)	99.05	0.003	0.005
Valparaiso (D_5)	98.95	0.004	0.007

Table A14. Results of the testing of the $MLP_{C\text{-}S\text{-}V}$ per deciles (approach 2).

Testing	R^2 [%]	MAE	$RMSE$
Concepción (D_1)	96.14	0.046	0.054
Santiago (D_1)	90.74	0.022	0.042
Valparaiso (D_1)	98.24	0.016	0.025
Concepción (D_2)	98.38	0.016	0.024
Santiago (D_2)	98.99	0.008	0.012
Valparaiso (D_2)	95.14	0.026	0.036
Concepción (D_3)	98.63	0.010	0.014
Santiago (D_3)	98.72	0.006	0.009
Valparaiso (D_3)	95.32	0.019	0.026
Concepción (D_4)	97.08	0.010	0.017
Santiago (D_4)	98.99	0.004	0.006
Valparaiso (D_4)	94.00	0.013	0.022
Concepción (D_5)	98.50	0.005	0.010
Santiago (D_5)	98.16	0.004	0.007
Valparaiso (D_5)	98.54	0.006	0.008

Fig. A1. Cloud points between the actual values and the predicted values for MLP_C (approach 1). The testing dataset used is indicated in brackets.

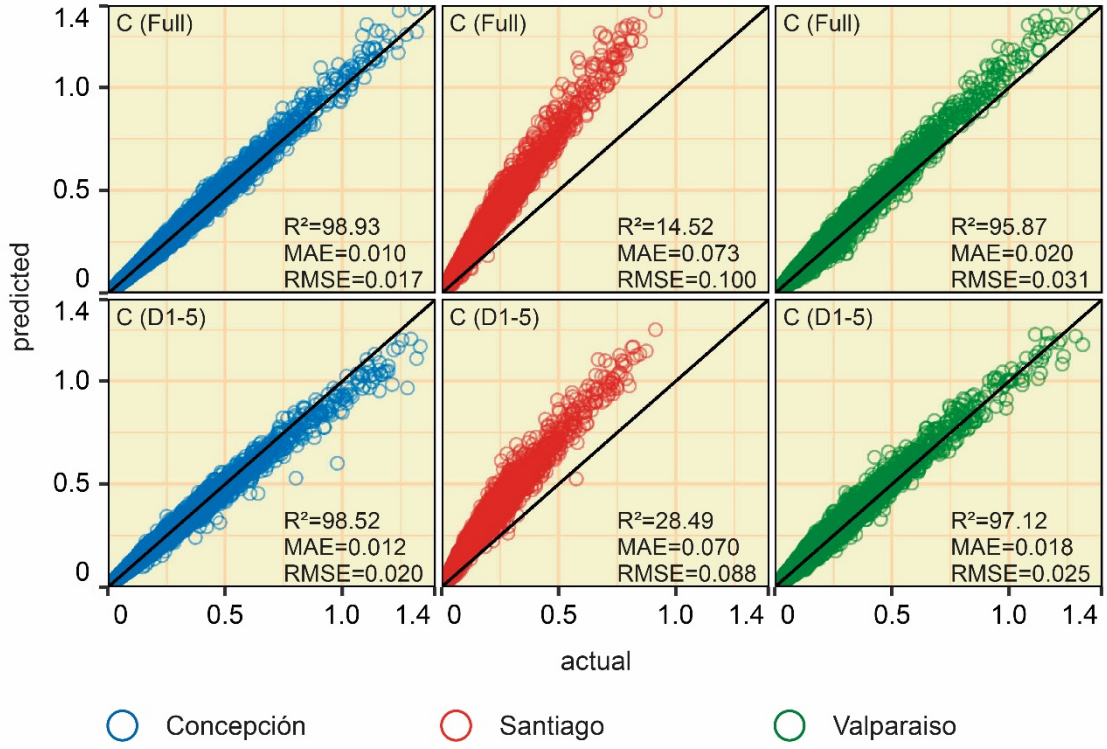

Fig. A2. Cloud points between the actual values and the predicted values for MLP_S (approach 1). The testing dataset used is indicated in brackets.

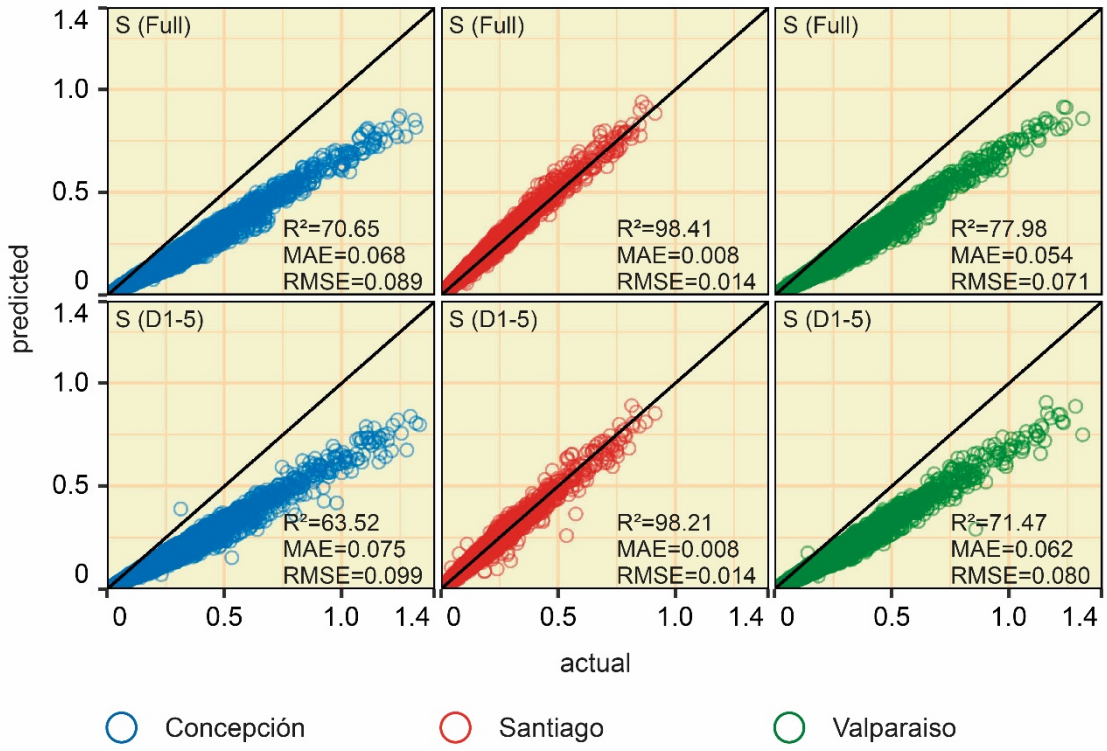

Fig. A3. Cloud points between the actual values and the predicted values for MLP_V (approach 1). The testing dataset used is indicated in brackets.

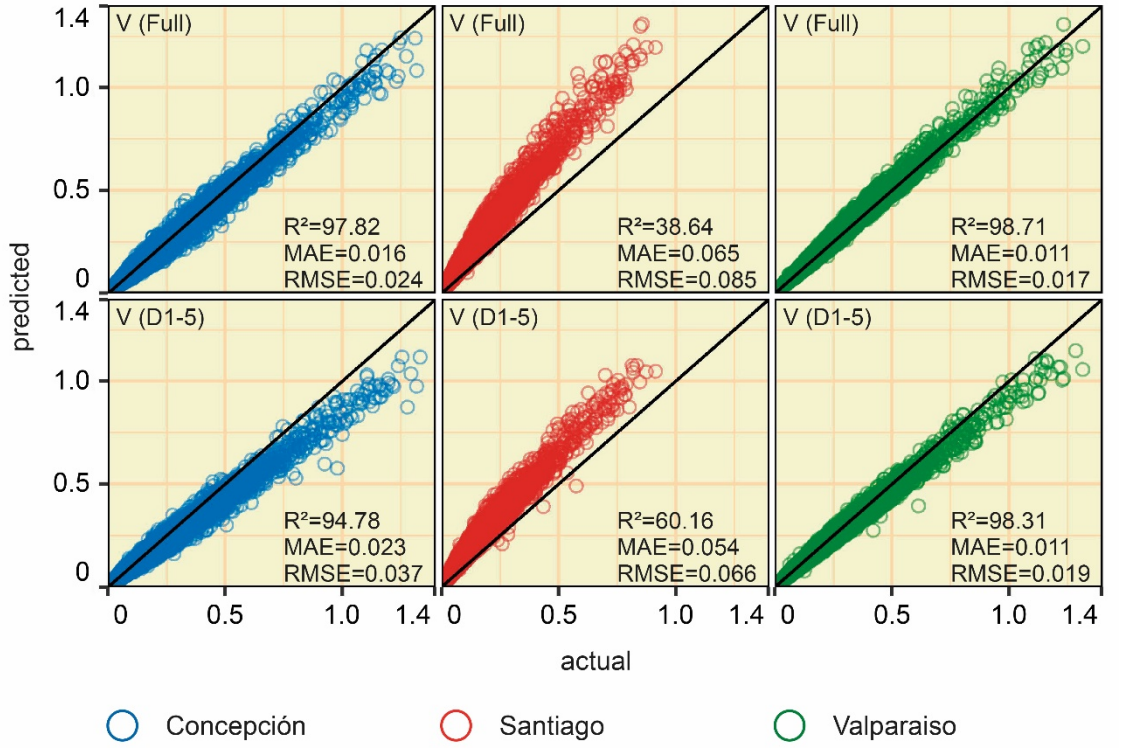

Fig. A4. Cloud points between the actual values and the predicted values for $MLP_{C\text{-}S}$ (approach 1). The testing dataset used is indicated in brackets.

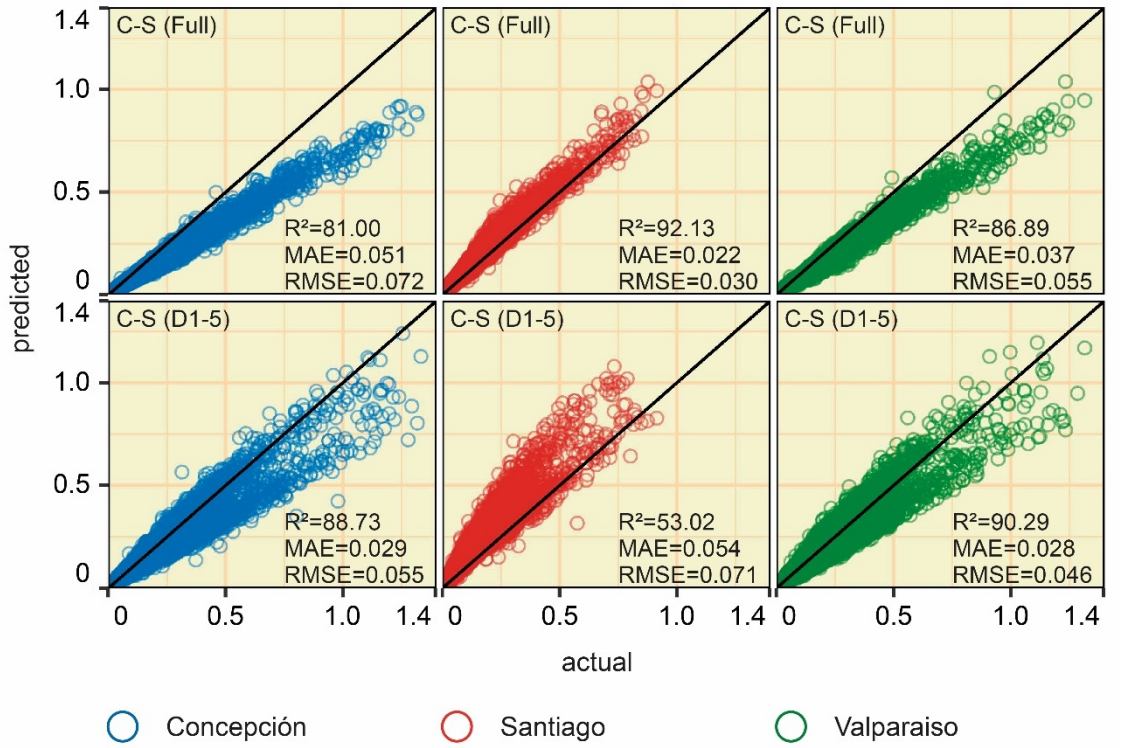

Fig. A5. Cloud points between the actual values and the predicted values for $MLP_{C\text{-}V}$ (approach 1). The testing dataset used is indicated in brackets.

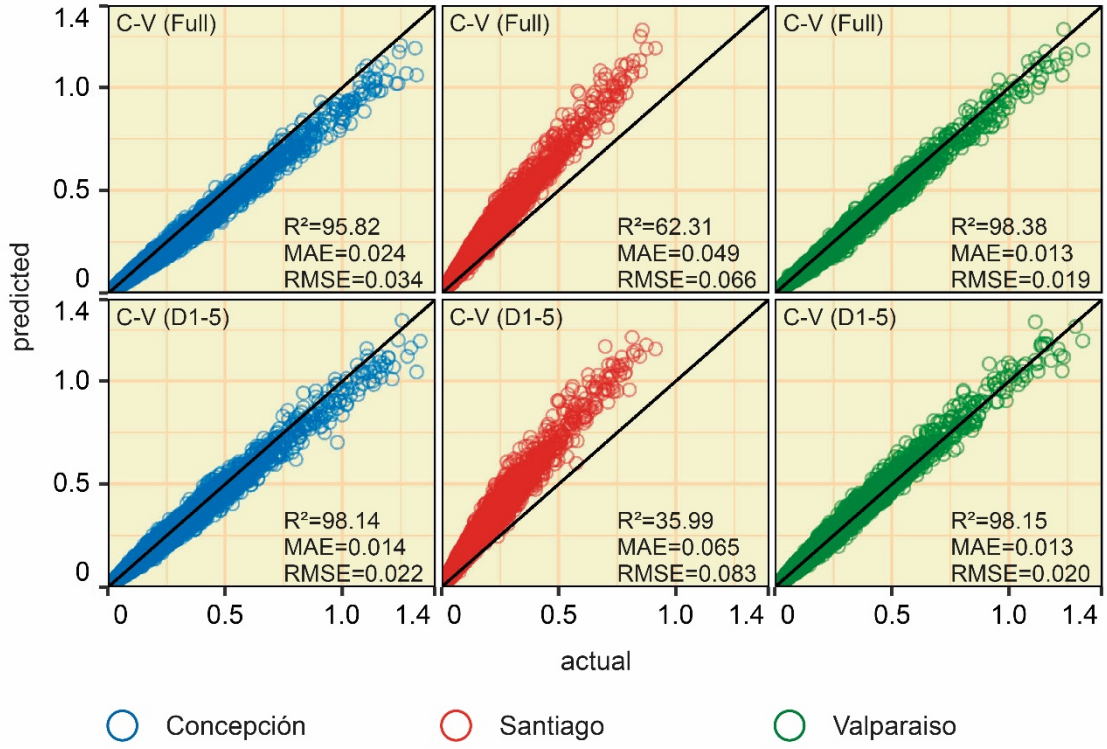

Fig. A6. Cloud points between the actual values and the predicted values for $MLP_{S\text{-}V}$ (approach 1). The testing dataset used is indicated in brackets.

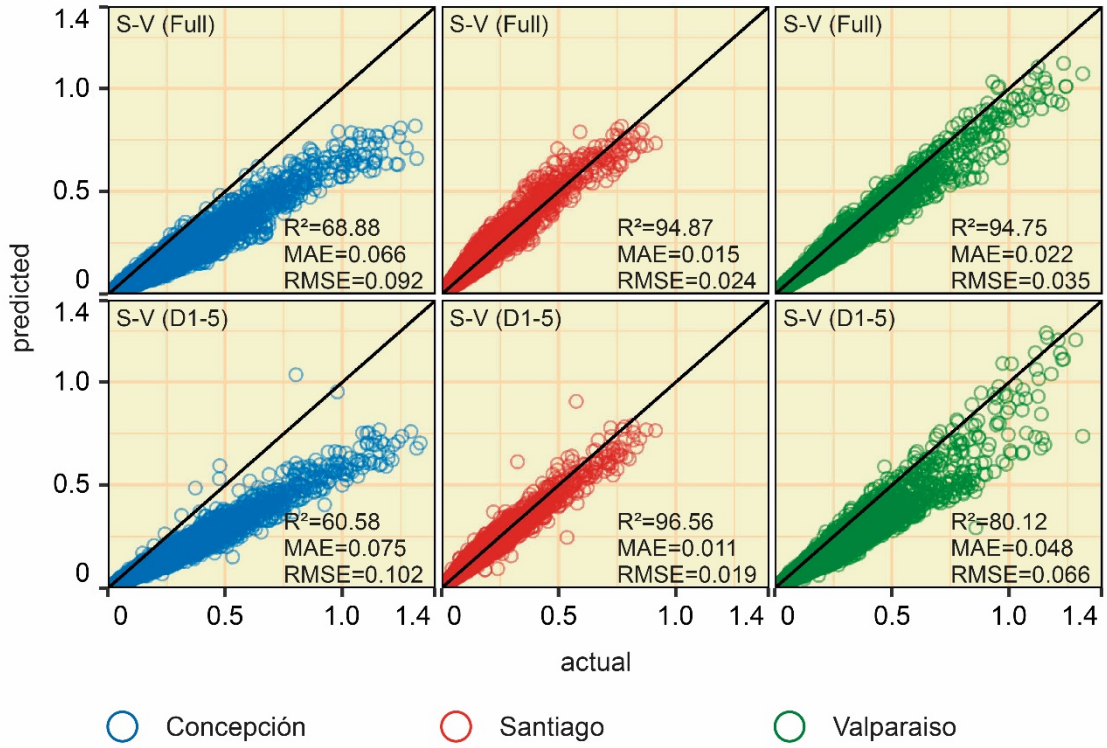

Fig. A7. Cloud points between the actual values and the predicted values for $MLP_{C\text{-}S\text{-}V}$ (approach 1). The testing dataset used is indicated in brackets.

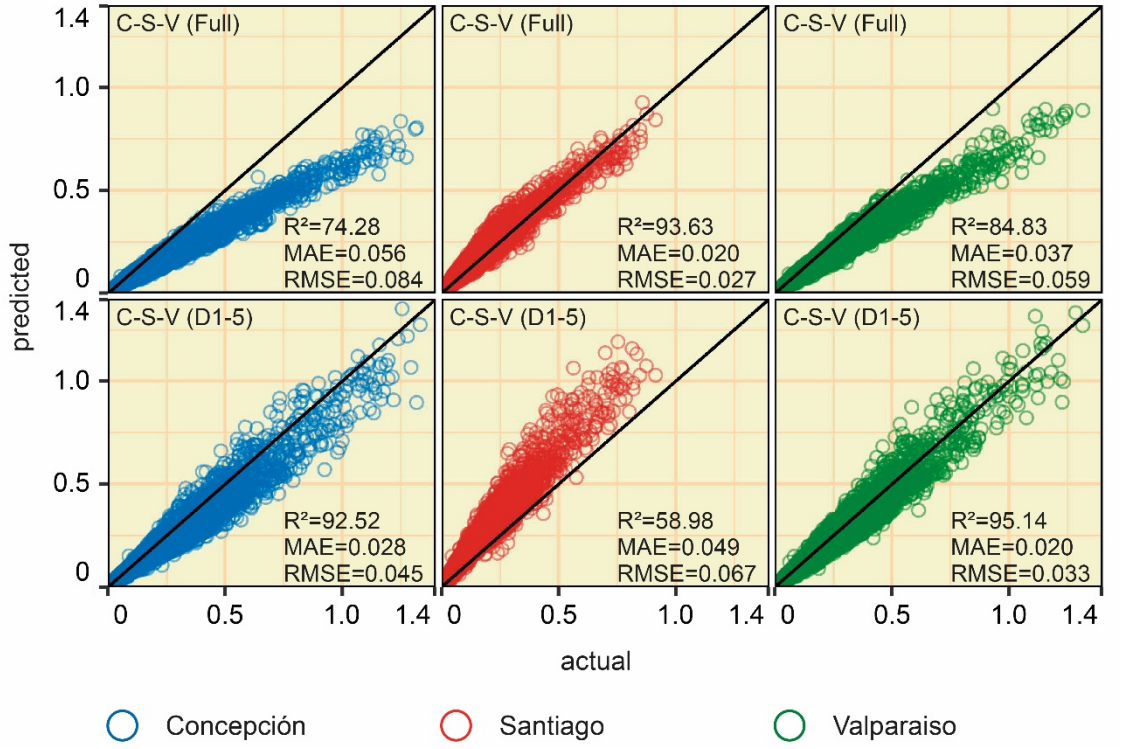

Fig. A8. Cloud points between the actual values and the predicted values for MLP_{C} (approach 2). The testing dataset used is indicated in brackets.

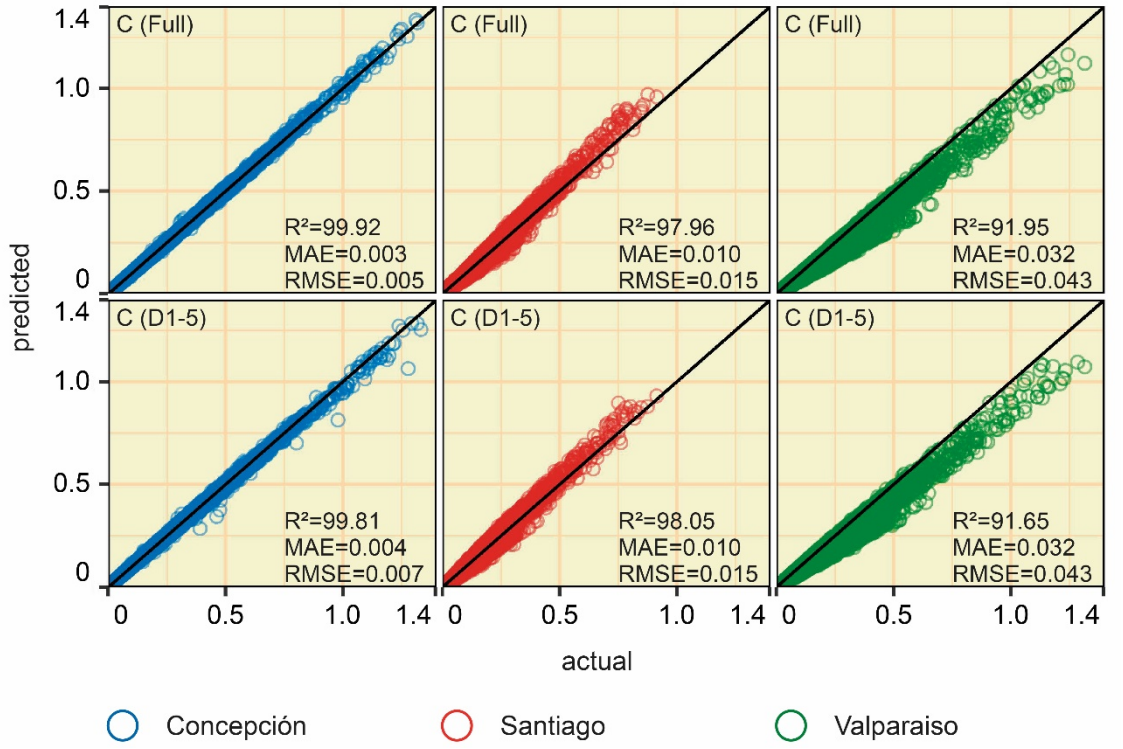

Fig. A9. Cloud points between the actual values and the predicted values for MLP_S (approach 2). The testing dataset used is indicated in brackets.

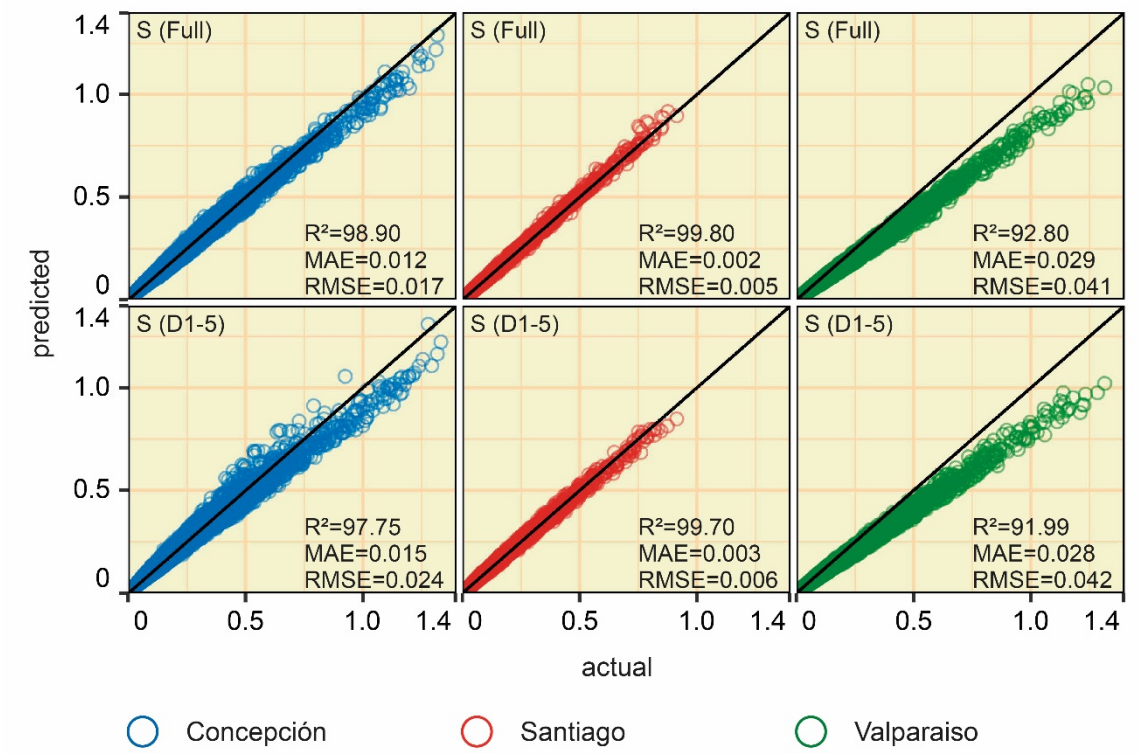

Fig. A10. Cloud points between the actual values and the predicted values for MLP_V (approach 2). The testing dataset used is indicated in brackets.

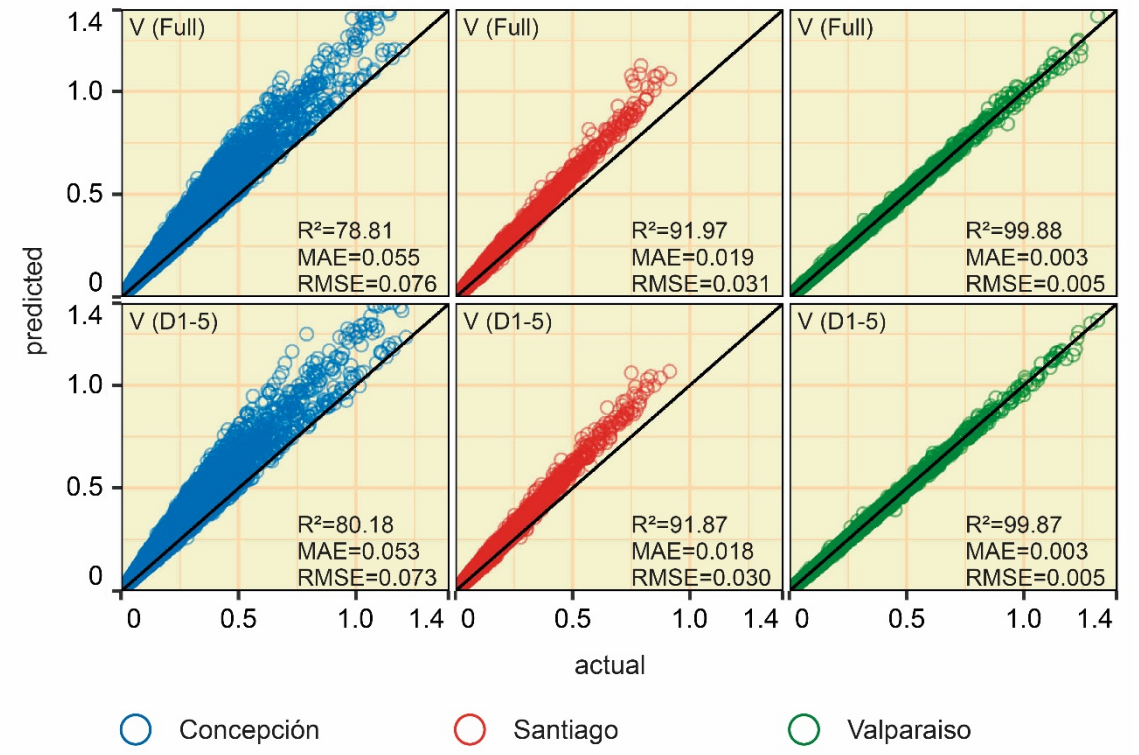

Fig. A11. Cloud points between the actual values and the predicted values for $MLP_{C\text{-}S}$ (approach 2). The testing dataset used is indicated in brackets.

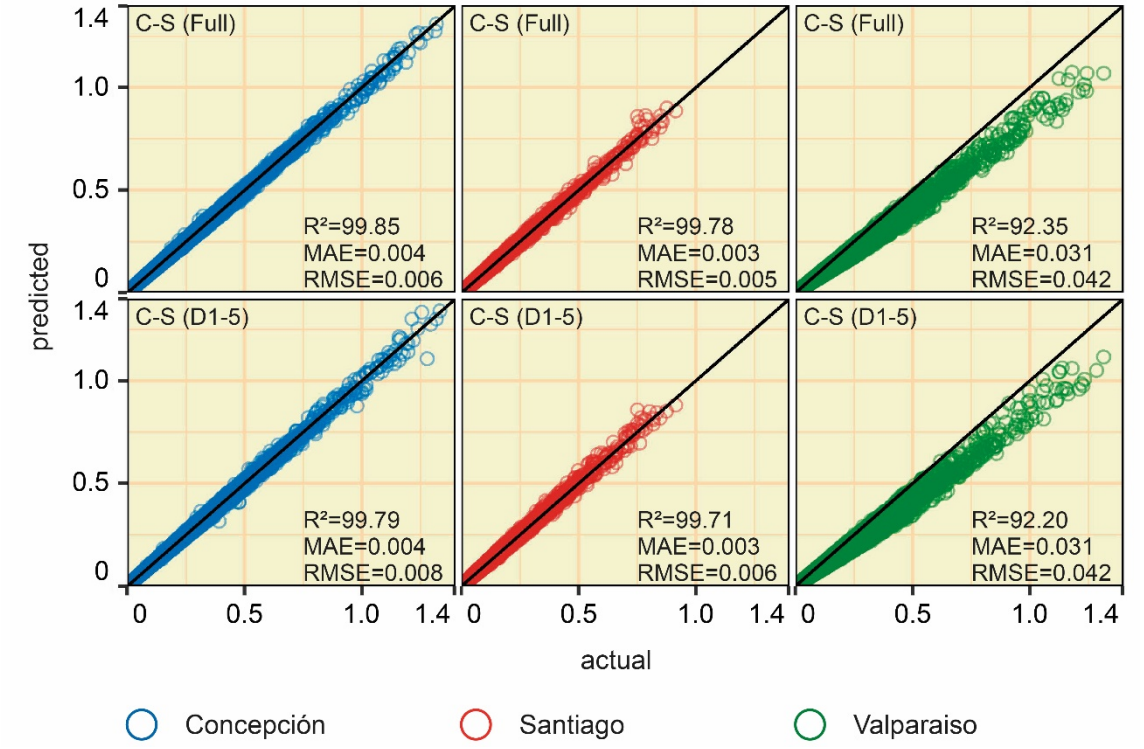

Fig. A12. Cloud points between the actual values and the predicted values for $MLP_{C\text{-}V}$ (approach 2). The testing dataset used is indicated in brackets.

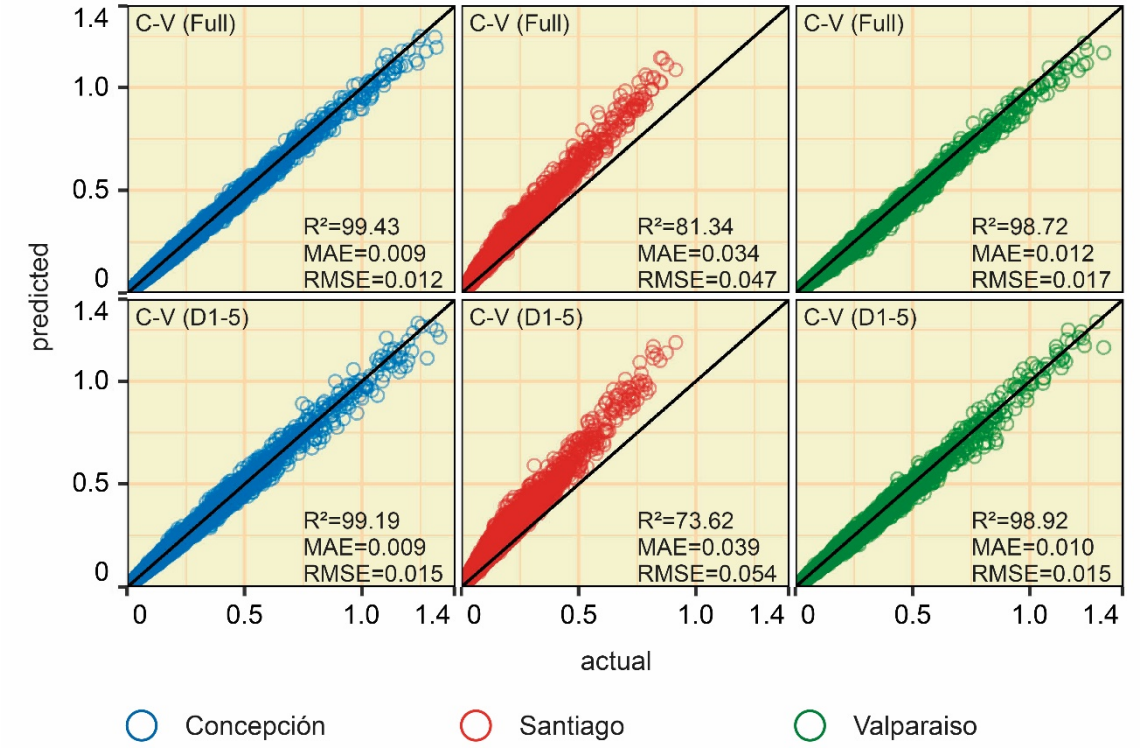

Fig. A13. Cloud points between the actual values and the predicted values for $MLP_{S\text{-}V}$ (approach 2). The testing dataset used is indicated in brackets.

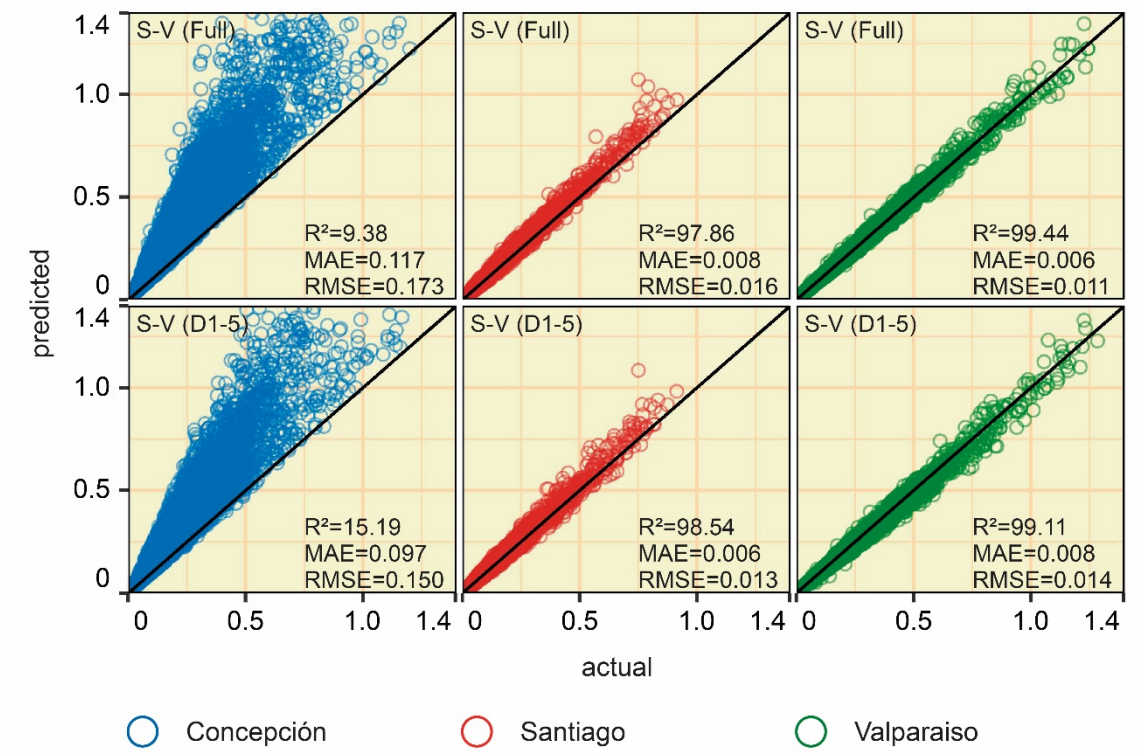

Fig. A14. Cloud points between the actual values and the predicted values for $MLP_{C\text{-}S\text{-}V}$ (approach 2). The testing dataset used is indicated in brackets.