




Article

Prediction of Fuel Poverty Potential Risk Index Using Six Regression Algorithms: A Case-Study of Chilean Social Dwellings

David Bienvenido-Huertas ^{1,*}, Jesús A. Pulido-Arcas ², Carlos Rubio-Bellido ¹ and Alexis Pérez-Fargallo ³¹ Department of Building Construction II, University of Seville, 41012 Seville, Spain; carlosrubio@us.es² Graduate School of Arts and Sciences, University of Tokyo, Tokyo 153-8902, Japan; jpulido@g.ecc.u-tokyo.ac.jp³ Department of Building Science, University of Bío-Bío, Concepción 410300, Chile; aperezf@ubiobio.cl

* Correspondence: jbienvenido@us.es

Abstract: In recent times, studies about the accuracy of algorithms to predict different aspects of energy use in the building sector have flourished, being energy poverty one of the issues that has received considerable critical attention. Previous studies in this field have characterized it using different indicators, but they have failed to develop instruments to predict the risk of low-income households falling into energy poverty. This research explores the way in which six regression algorithms can accurately forecast the risk of energy poverty by means of the fuel poverty potential risk index. Using data from the national survey of socioeconomic conditions of Chilean households and generating data for different typologies of social dwellings (e.g., form ratio or roof surface area), this study simulated 38,880 cases and compared the accuracy of six algorithms. Multilayer perceptron, M5P and support vector regression delivered the best accuracy, with correlation coefficients over 99.5%. In terms of computing time, M5P outperforms the rest. Although these results suggest that energy poverty can be accurately predicted using simulated data, it remains necessary to test the algorithms against real data. These results can be useful in devising policies to tackle energy poverty in advance.

Keywords: fuel poverty potential risk index; multilayer perceptron; *K*-nearest neighbors; tree models; support vector regression



Citation: Bienvenido-Huertas, D.; Pulido-Arcas, J.A.; Rubio-Bellido, C.; Pérez-Fargallo, A. Prediction of Fuel Poverty Potential Risk Index Using Six Regression Algorithms: A Case-Study of Chilean Social Dwellings. *Sustainability* **2021**, *13*, 2426. <https://doi.org/10.3390/su13052426>

Academic Editor: Vincenzo Costanzo

Received: 21 January 2021

Accepted: 18 February 2021

Published: 24 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

1.1. Energy Poverty: Definition and Conceptualization

Energy poverty, also called fuel poverty or energy vulnerability, takes place when households cannot keep comfortable temperatures inside their homes or cannot access energy services at a reasonable cost [1,2]. This phenomenon has received the attention of the scientific community and society in recent years due to its political and social implications [3,4]. Previous research has shown that energy poverty is a driver for the physical deterioration of residents [5,6] and is even responsible for a higher death rate in winter due to the poor thermal conditions inside buildings [7,8]. There is a wide consensus about the fact that energy poverty stems from a combination of high-energy prices, low family income, inefficient buildings, and outdated electrical household appliances [9,10]. Additionally, occupant behavior can contribute to new cases of energy poverty [11,12].

The measurement of energy poverty is a challenge because its condition is culturally sensitive and private, as well as temporal and dynamic [2]. It is a relative concept influenced by different variables, whose importance may vary depending on the context [13]. The limited availability of data and indicators, together with a lack of consensus on how energy poverty should be conceptualized and measured, makes the matter even worse. Energy poverty indicators are thus a pressing need in this research and political landscape [14]. The United Kingdom pioneered the study of fuel poverty in 1991, with the

publication of the seminal work by Boardman [15] that later spread across other European countries. Nowadays, several European countries, such as Slovakia, France and Ireland, have conceptualized fuel poverty building upon a multidimensional approach based on indicators [16]. Those can be divided into six categories: (i) indicators based on the relationship between household income and expenses [17,18]; (ii) multidimensional indicators [19,20]; (iii) indicators based on self-reported housing condition [21]; (iv) indicators of econometric analysis [22,23]; (v) indicators associated with the energy rating of dwellings [24,25]; and (vi) indicators calculated according to the thermal comfort [26]. From a wider perspective, energy poverty includes other factors, such as social exclusion [5] and the limited availability of some energy sources [19]. Outside Europe, fuel poverty has been addressed mainly from the economic perspective, that is, as a balance between household income and expenses. Nevertheless, in countries like Japan, more recent attention has focused on the multiple facets of fuel poverty [27].

A considerable amount of literature has been published on fuel poverty in developed countries, but the existing accounts fail to describe the situation in underdeveloped or developing countries. In this regard, Chile's current situation is worth to be noted for two reasons: First, the country has continuously implanted, since the mid-20th century, a strong policy on social housing, providing the vulnerable strata of society with affordable dwellings; as per the last census of 2016, around one million and a half dwellings have been completed [28]. On the contrary, the geography of the country gives a very particular energy market; Chile has one of the highest prices of electricity among the member countries of the OECD [29]. Despite being a member of this organization since 2011, the gap between the richest and poorer strata of society is still evident. Previous research has shown that a considerable number of families face difficulties in paying the energy to keep their homes comfortable [30], and many others cannot adequately ventilate or heat them up [31]. According to a recent report by the Chilean government, 20% of the homes do not have hot water, and the indoor temperature in summer and winter is not acceptable in 76% of the surveyed homes [30]. A study conducted by the Chilean Chamber of Construction has presented the breakdown of the energy use of Chilean households: 56% is allocated for keeping their homes within acceptable thermal conditions, 18% for hot water, and 26% for electrical appliances [29].

These circumstances compelled the Chilean government to take action. In 2014 the Chilean Ministry of Energy presented the Energy Agenda, whose main tasks was the design and execution of a long-term Energy Policy for 2050 [32], which builds upon four aspects: (i) security and quality in the supply chain, (ii) energy as a driver for the development of the country, (iii) compatibility with the environment, and (iv) energy efficiency and education. Within the second aspect, one of the lineaments was the conceptualization and measurement of energy poverty to establish specific policies towards it. The objective is to reduce energy poverty levels by 50% by 2035 and by 100% by 2050 [33]. As previously mentioned, owing to the fact that measuring energy poverty strongly depends on the culture, climate, and socioeconomic conditions of each country [34], the government of Chile signed an agreement with The United Nations Development Program to elaborate a conceptual and methodological framework specifically designed for Chile [35].

1.2. Energy Poverty in Chile

Over the past decade, the conceptualization of energy poverty has progressed in Chile, but much of the research has exclusively focused on identifying this issue ex-post. However, there has been little discussion on how energy poverty can be predicted in advance thus far; that is before a family moves into a new house with new heating and cooling systems, electrical appliances and energy consumption. Understanding this research gap, the authors took the lead and started a new line of research that offered some important insights into the importance of predicting energy poverty in advance [31,36,37]. Evidence for these studies suggests that two main factors should be considered: The household income and the adopted comfort model; both of them are introduced below.

According to the Chilean national socioeconomic survey, Chilean families can be divided into 10 different groups (called deciles) depending on their income. The 10th decile corresponds to the wealthiest families and the 1st decile to the poorest. Table 1 indicates the ranges of the mean monthly income corresponding to each decile in 2006, 2009, 2011, and 2013.

Table 1. Variation of ranges of Chilean deciles in recent years [38].

Year	Mean Monthly Household Income [\$]									
	Decile									
	I	II	III	IV	V	VI	VII	VIII	IX	X
2006	88,955	216,673	293,257	364,414	428,116	513,218	642,546	790,981	1,097,094	2,464,055
2009	65,151	211,744	283,602	370,674	417,867	510,817	632,655	776,025	1,093,990	2,543,369
2011	84,243	229,739	304,091	385,760	448,741	520,257	636,893	793,491	1,099,645	2,505,654
2013	100,707	261,125	353,135	428,817	510,310	630,642	730,877	901,626	1,291,669	2,857,726

Second, the energy expenditure of households can be modeled according to the gap between the indoor and outdoor temperature. In this regard, a much-debated question is whether to use a static or an adaptive thermal comfort model. The static comfort model, which is adopted by many national standards, considers static setpoint temperatures. However, it fails to explain people's ability to adapt themselves to the thermal variations inside their homes within certain limits. Current Chilean standards use this approach (Table 2).

Table 2. Comfort limits indicated in the sustainable construction standards for housing in Chile.

Comfort Limits (°C)			
Day		Night	
Lower	Upper	Lower	Upper
20	27	17	27

The adaptive comfort model builds upon the fact that inhabitants can adapt to thermal variations taking certain actions: Adapting their clothing and ventilating the building. These adaptive measures are feasible when indoor temperature, which is a function of the outdoor thermal variations [39], is within a certain range. Previous research has shown that adaptive thermal models can better represent the real energy consumption of buildings, which can vary between 10% and 18% when compared with the theoretical approach of the static model [40].

Since evidence suggests that the adaptive comfort model can better explain the real energy consumption of buildings, the authors have recently examined the influence of the application of such a model on the quantification of energy poverty and have developed a new index called the fuel poverty potential risk index (FPPRI). The theoretical basis is detailed in previous publications by the authors [41] and also has been tested against different scenarios of climate change in Chile [42], also using advanced simulation techniques, such as multiple linear regression [37].

Evidence from these previous studies suggests that these techniques can predict the risk of fuel poverty with acceptable accuracy, but it still remains unknown whether other advanced techniques may outperform the former. There is still uncertainty whether fuel poverty can be predicted with different regression algorithms. The specific objective of this study is to test the accuracy of six regression algorithms to predict the risk of fuel poverty, expressed by the FPPRI (i) multilayer perceptron (MLP); (ii) *K*-nearest neighbors (*K*-NN); (iii) classification and regression tree (CART); (iv) random forest (RF); (v) M5P; and (vi) support vector regression (SVR). These algorithms were chosen because they are commonly used in the prediction of energy consumption in the building industry, as pointed out by a

recent study [43], which concluded that interplay between the architecture of the models and the predicted phenomenon exerted a strong influence on the final accuracy. Debate continues about, which is the algorithm that best-fits each problem, and therefore this exploratory study aims to unravel the accuracy of these six advanced regression techniques in predicting the risk of falling into fuel poverty of financially deprived households in Chile by means of the FPPRI index.

It is expected that this research will contribute to a better understanding of how to predict energy poverty in advance, therefore making it possible to detect and tackle the problem in a swift and effective way. This new understanding should help designers and policymakers in building and delivering affordable and efficient social dwellings and in allocating financially deprived families in new houses whose energy expenditure would be suitable for their income level.

This paper is organized in the following way: First, an explanation of the predictive models is provided. Second, the predictor variables for the models are determined, and afterward, the data processing and results are presented. Finally, the results obtained by all models are compared and evaluated, which leads to the discussion of results and conclusions on the feasibility of such models.

2. Materials and Methods

2.1. Regression Algorithms

As previously mentioned, six regression algorithms were considered in this study. For each one of them, a brief explanation of its architecture and its main characteristics is provided. This information will be used later on to discuss the results of each one of them.

2.1.1. Multilayer Perceptron

A neural network is a statistical model simulating the neurological brain structure to solve linear and nonlinear problems [44]. This model is a computation paradigm, which allows complex problems, both classification [45,46] and regression problems [47,48], to be tackled. Neural networks have been widely used in the field of energy analysis over the last years: (i) Deb et al. [49] developed a neural network model to predict the energy saving associated with HVAC in office buildings and compared the results to multiple linear regression. The results determined a better performance for the neural network than linear regression, with a correlation coefficient higher than 37.25%; (ii) in a later study also by Deb et al. [50], a model was developed to predict the cooling load of the next day using data of the previous five days. The methodology showed an adequate accuracy, with a correlation higher than 94% in the days analyzed; (iii) Magalhães et al. [51] developed an artificial neural network model to predict the heating energy demand based on the occupants' behavior. The results obtained models with a correlation coefficient higher than 93%; (iv) Kljajic et al. [52] developed a neural network to predict the efficiency of boilers according to the operating performance as well as to predict possible improvement measures; and (v) Kialashaki and Reisel [53] conducted a large-scale application of the algorithm, designing a model to estimate the energy demand in the residential sector of the United States using historical data from the last years.

From the various architectures of artificial neural networks, MLPs are those providing the best features because they are supervised models with universal approximation capacities [54,55]. The MLPs are characterized by presenting an architecture of three or more layers (see Figure 1): an input layer, one or several intermediate layers, and an output layer. There are a series of nodes in each layer. The output value of each neuron is obtained by the sum of values of the input neurons weighted by synaptic weights and by applying an activation function (Equation (1)). These connections spread to the output layer (Equation (2)), obtaining the model's response (\hat{Y}_{MLP}).

$$y_k = \sigma \left(\sum_{j=1}^d w_{kj}^{(1)} x_j + w_{k0}^{(1)} x_0 \right) \quad (1)$$

$$\hat{Y}_{MLP} = \sigma \left(\sum_{k=1}^M w_{lk}^{(2)} \sigma \left(\sum_{j=0}^d w_{kj}^{(1)} x_j \right) + w_{l0}^{(2)} y_0 \right) \quad (2)$$

where x_j are the values of the input layer, $w_{k0}^{(1)}$ and x_0 are the weight and the input value of the bias neuron of the input layer, respectively, $w_{kj}^{(1)}$ are the weights of the hidden layer, $w_{l0}^{(2)}$ and y_0 are the weight and the input value of the bias neuron of the hidden layer, respectively, $w_{lk}^{(2)}$ are the weights of the output layer, y_k is the output of a neuron of the hidden layer, and σ is the activation function. For this study, a sigmoidal activation function was used both in the hidden and the output layer (Equation (3)). The advantage of this kind of function is the possibility of comprising an infinite input set into a finite output set.

$$\sigma = \frac{1}{1 + e^{-x}} \quad (3)$$

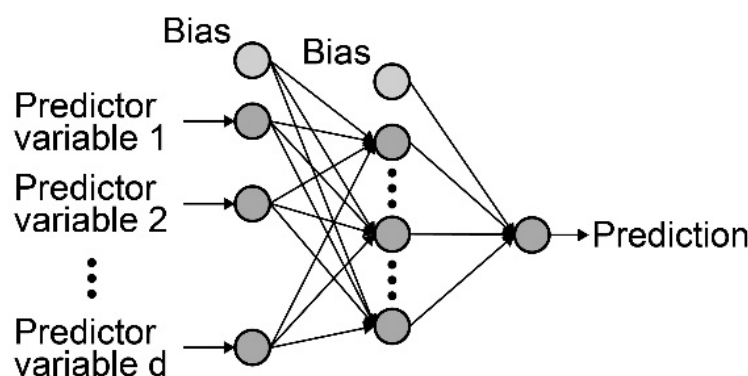


Figure 1. Scheme of a multilayer perceptron (MLP) model.

The main objective of the algorithm is the adjustment of synaptic weights guaranteeing that the predicted output value for each input vector is near to the actual output value. For this purpose, a learning algorithm is applied to a training dataset. The models were trained by backpropagation [56], using the Broyden–Fletcher–Goldfarb–Shanno (BFGS) algorithm [57], which belongs to the set of quasi-Newton methods. The number of nodes and the number of hidden layers is analyzed to determine the best model [47]. For this study, architectures of 1 hidden layer were considered because they had better performance than more complex structures [58], but the number of nodes in such layer was analyzed.

2.1.2. K-Nearest Neighbors

The K -NN algorithm (also known as instance-based learning with parameter K) classifies observations based on the majority class between the nearest observations [59]. These observations, or neighbors, are chosen from the training dataset in which the classification value is known. The K -NN algorithm is used for class label problems [60] and regression analysis [61], implying their use in the energy analysis of installations and buildings: (i) Ramakrishna Madeti and Singh [62] developed a K -NN model to detect failures in photovoltaic devices. The percentage of instances correctly classified by this model was higher than 98%, obtaining better performance than other existing models; (ii) Rodger [63] developed a K -NN model to predict the heating energy demand. The values predicted by the model presented a correlation higher than 94% with respect to the actual values; and (iii) a model to determine the number of occupants inside the building during the monitoring was used [64]. The results obtained by the model designed in this study presented a higher accuracy than those obtained by linear discriminant functions.

In the regression analysis, the K -NN algorithm determines the value of the observation analyzed as an average of the nearest neighbors' values (K) of the training dataset. To deter-

mine the closeness between observations, a metric distance is used. This distance strongly affects the performance of the model [65], so it is necessary to study which distance presents a better performance in the model. In the present study, the Minkowski distance was used (Equation (4)). This distance is characterized by the fact that the Minkowski distance is equal to other metric distances depending on the value assigned to q [66]: (i) the Manhattan distance (q equal to 1); (ii) the Euclidean distance (q equal to 2); and (iii) the Chebyshev distance (when q is near to the positive infinity). After determining the distances, the output value of the model is given by Equation (5). The output value (\hat{Y}_{kNN}) is obtained as the average of values (y_m) of the K nearest points. As can be seen, K is a key parameter in the model's response [67]. Both q and K were therefore analyzed due to their influence on the performance of the model.

$$d_q = \left[\sum_{r=1}^p (x_{ir} - x_{jr})^q \right]^{1/q} \quad (4)$$

$$\hat{Y}_{kNN} = \frac{1}{K} \sum_{m=1}^K y_m \quad (5)$$

2.1.3. Classification and Regression Tree

The CART algorithm, developed by Breiman et al. [68], constructs models in a reverse tree shape. Such models divide the input space into subregions, simplifying complex with simple problems [69]. Trees are composed of internal nodes corresponding to variables, arches to values of the root node, and leaves to the value of the dependent variable. Trees are easy to interpret because a scheme of nodes and leaves is available, thus easing the understanding of the solution adopted for the problem [70]. For this reason, trees are used in several works: (i) Tso and Yau [71] developed models to predict energy consumption by using three different algorithms: CART, MLP, and linear regression. The results revealed that both the CART and the MLP model obtained better performance than the linear regression; (ii) Mousa et al. [72] analyzed the use of a CART model to estimate the air change rate in buildings; and (iii) in another study [73], a CART model was developed to predict the monthly energy consumption in residential buildings. The results showed that the tree model obtained better performance than linear regression and multivariate adaptive regression splines.

The optimal structure is obtained by the binary recursive partitioning process during the development of the model: each node of the model has a partitioning rule. This rule is established by decreasing the total residual sum square. After carrying out the induction, the application of the pruning (that is, the removal of the inefficient nodes) allows the complexity of the model to be generalized and reduced [74]. The depth of the tree and the minimum number of instances per node is needed to be established in the configuration of the algorithm. These parameters affect the complexity and performance of the model, so they were analyzed.

2.1.4. Random Forest

As indicated above, the simplicity and interpretation of CART models make their use easier. However, many studies reflect the limitations of such models [75,76]. RF allows a solution to these problems to be adopted by creating a set of CART models generated in parallel to reduce the variance and the bias of the RF model [77,78]. RF is an ensemble learning algorithm, which obtains a better behavior than an individual model [79]. Another advantage of RF is the possibility of using large datasets, as well as not being affected by atypical values [80]. This aspect allows this model to be used to predict energy consumption. In this sense, Li et al. [81] developed an RF model to predict the daily electricity consumption in companies. The model had a more accurate prediction than artificial neural networks and support vector machines. Likewise, its monthly [82] and hourly [83] use as a predictive control system of HVAC systems has been analyzed in some studies.

For the RF training, N bootstrapped sample sets are drawn from the training dataset [78]. Each bootstrapped sample allows a forest regression tree to be generated. Moreover, each node of each tree is divided using a subset of predictors m randomly selected. This reduces the influence of the strongest predictors. After generating the set of T -trees and following the RF model structure (see Figure 2), the regression estimator (\hat{Y}_{RF}) is as follows:

$$\hat{Y}_{RF} = \frac{1}{T} \sum_{t=1}^T \hat{Y}_t \quad (6)$$

where T is the number of trees, and \hat{Y}_t is the output of the t -th tree.

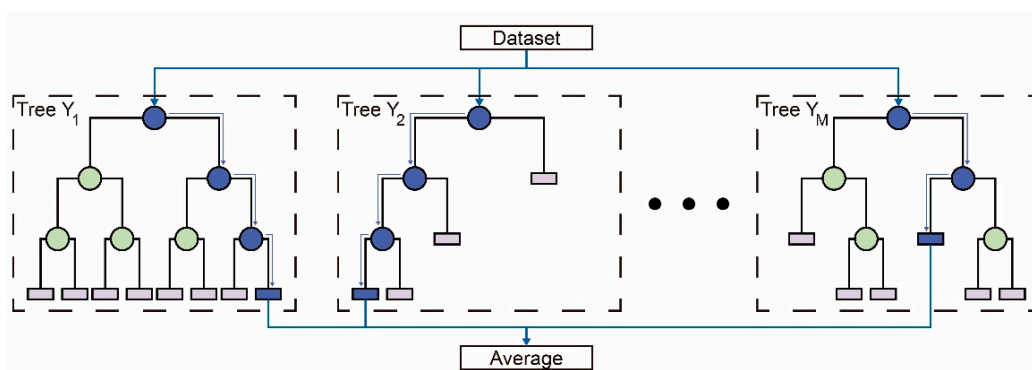


Figure 2. Scheme of a random forest (RF) model.

The number of trees used by the model influences its performance [84]. For this reason, different numbers of trees were analyzed to obtain the model with the most efficient behavior.

2.1.5. M5P

The M5P algorithm (also known as M5) is an evolution of the CART algorithm [85,86]. Unlike the CART model, M5P combines decision trees with multiple regression: a decision tree is constructed following the same structure of reverse tree from the CART model, but a multiple linear regression (MLR) model is adjusted in each leaf of the model (Equation (7)). The algorithm, therefore, develops an MLR model in each subregion. In the development of the M5P tree, the internal variation of subsets for the class values of each branch is minimized instead of maximizing the information gain. After constructing the model, the pruning allows the overfitting to be reduced. The advantages of the models generated by this algorithm are the efficient use of huge amounts of numeric variables and their robustness because of the lack of values in the instances of the dataset analyzed [87,88]. The use of this model for the energy characterization of buildings has therefore increased in recent years. In this sense, Li et al. Afsarian et al. [89] developed an M5P model to predict the total energy consumption in a reference building. The results obtained a R^2 greater than 90% in most datasets. In another study by Jeffrey Kuo et al. [90], a comparative study of regression models was conducted to estimate the energy consumption in convenience stores of Taiwan. The use of four different models (M5P, Gaussian processes, linear regression, and support vector regression) was compared. The results obtained by M5P obtained a more accurate estimation than the other models. Likewise, M5P has a good performance in determining oscillations in energy prices. In a study by Azofra et al. [91], a model was developed to determine the variations of the cost of electricity tariffs because of the generation of photovoltaic and wind power in Spain. The model was used to analyze various scenarios, thus determining the possible saving in the cost of electricity tariff.

$$\hat{Y}_{MLR} = \beta_0 + \sum_{i=1}^v (\beta_i x_i) + \varepsilon \quad (7)$$

where β_0 is the independent term, β_i are the regression coefficients, x_i are the predictor variables, and ε is the error.

2.1.6. Support Vector Regression

SVR is an application for regression analysis of the support vector machines [92]. SVR consists of transforming the root input variable in the space of higher dimension where two classes are separated by an optimal hyperplane [93,94]. Some advantages of SVR are the facts that they are less subject to overfitting [95] and their good behavior when tackling nonlinear problems [96]. It is an algorithm widely developed in the energy control of installations, such as lighting [97] and HVAC systems [98]. Likewise, SVR has been used in some studies to estimate the energy consumption in residential buildings [99,100] and offices [101,102].

For their mathematical approach, support vector machines approach the relationship between the output and input parameters by using the following equation:

$$f(x) = w^T \cdot \varphi(x) + b \quad (8)$$

where $\varphi(x)$ is representative of a nonlinear mapping function, w is a weight vector, and b is the bias term. w and b can be predicted by minimizing the regularized risk function (Equation (9)), guaranteeing the restrictions of Equation (10).

$$\min \left\{ C \sum_{i=1}^N (\zeta_i + \zeta_i^*) + \frac{\|w\|^2}{2} \right\} \quad (9)$$

$$\begin{aligned} y_i - w^T \cdot \varphi(x_i) - b &\leq \psi + \zeta_i \quad i = 1, 2, \dots, N \\ w^T \cdot \varphi(x_i) + b - y_i &\leq \psi + \zeta_i^* \quad i = 1, 2, \dots, N \\ \zeta_i, \zeta_i^* &\geq 0 \\ C &> 0 \end{aligned} \quad (10)$$

where N is the sampling number, C is penalty parameter to control the compensation between the regularization term ($\|w\|^2$) and the training error, ψ is the maximum error allowed, ζ_i and ζ_i^* represent the distance of the actual values from the upper and lower limit of the error allowed, respectively.

Finally, the output of the SVR (Equation (11)) is obtained by introducing Lagrange multipliers (δ_i and δ_i^*) and a Kernel function ($K(x, x_i)$). In this regard, it is essential to determine which kernel function is used. There are different kernel functions: linear, polynomial, sigmoidal, and radial basis function (RBF). For this study, the RBF kernel is used (Equation (12)) because models with the best performance are obtained, and a lower number of parameters should be considered to improve the behavior of the model [74,103].

$$f(x) = \sum_{i=1}^N (\delta_i + \delta_i^*) \cdot K(x, x_i) + b \quad (11)$$

$$K(x, x_i) = \exp\left(-\gamma \|x - x_i\|^2\right) \quad (12)$$

The parameters that should be considered to improve the performance of the model are C and γ . To determine the model with the best performance, different combinations of C and γ were used.

2.2. Dataset and Model Construction

2.2.1. Description of the Dataset

Based on the calculation methodology by Pérez-Fargallo et al. [31,36], the procedure to determine the dataset was described in a previous study [37]. The case studies were configured using different parameters of the A1 and B1 typology of the social dwelling of the Chilean Ministry of Housing and Urbanism (MINVU in Spanish) [104]. A total

of 7776 case studies were configured. Likewise, each case study is considered to belong to each of the 5 poorest deciles [105]. Thus, 5 different datasets per decile, composed of 7776 instances, were obtained. In total, 38,880 cases were considered.

The predictor variables were as follows: (i) orientation (OR), (ii) form ratio (FR), (iii) volume (V), (iv) surface in contact with the ground (S_G), (v) horizontal surface area in contact with another dwelling (S_H), (vi) roof surface area (S_R), (vii) vertical surface area in contact with another dwelling (S_V), (viii) shadow distance (D), (ix) shadow height (H), (x) energy price (EP), and (xi) income (IN). Figure 3 shows Pearson's correlation coefficients and the p -values. For nearly all of them, the null hypothesis can be rejected, and the Pearson's coefficient shows no evident correlation. A more surprising correlation is found between the group of variables related to the design of the dwelling (S_G , S_R , S_H , S_V and V), for which the null hypothesis cannot be rejected, and moderate correlations are found in some cases. This is because those are interweaved in the design of the dwelling, but this does not compromise the accuracy of the model. The dependent variable was the FPPRI, which was logarithmically transformed to improve the accuracy of the final model.

	OR	FR	V	S_G	S_H	S_R	S_V	D	H	EP	IN
OR											
FR	(1) 0.0001 (2) 0.9828										
V	(1) -0.0001 (2) 0.9830	(1) -0.0049 (2) 0.3331									
S_G	(1) 0.0007 (2) 0.8939	(1) -0.0016 (2) 0.7545	(1) 0.3214 (2) 0.0000								
S_H	(1) -0.0003 (2) 0.9564	(1) -0.0023 (2) 0.6542	(1) 0.4589 (2) 0.0000	(1) -0.3321 (2) 0.0000							
S_R	(1) -0.0004 (2) 0.9296	(1) -0.0005 (2) 0.9184	(1) 0.0715 (2) 0.0000	(1) -0.4627 (2) 0.0000	(1) -0.3318 (2) 0.0000						
S_V	(1) 0.0001 (2) 0.9835	(1) 0.1754 (2) 0.0000	(1) 0.3412 (2) 0.0000	(1) 0.1549 (2) 0.0000	(1) 0.1321 (2) 0.0000	(1) 0.0003 (2) 0.9540					
D	(1) -0.0001 (2) 0.9838	(1) 0.0002 (2) 0.9678	(1) -0.0002 (2) 0.9682	(1) 0.0003 (2) 0.9520	(1) 0.0019 (2) 0.7051	(1) 0.0018 (2) 0.7283	(1) 0.0001 (2) 0.9897				
H	(1) -0.0001 (2) 0.9838	(1) 0.0002 (2) 0.9678	(1) -0.0002 (2) 0.9682	(1) -0.0001 (2) 0.9831	(1) 0.0002 (2) 0.9715	(1) -0.0003 (2) 0.9550	(1) 0.0001 (2) 0.9918	(1) -0.0002 (2) 0.9696			
EP	(1) 0.0039 (2) 0.4461	(1) 0.0017 (2) 0.7408	(1) 0.0165 (2) 0.0011	(1) 0.0000 (2) 0.9953	(1) 0.0084 (2) 0.0965	(1) 0.0067 (2) 0.1859	(1) 0.0031 (2) 0.5374	(1) -0.0203 (2) 0.0001	(1) -0.0048 (2) 0.3489		
IN	(1) -0.0046 (2) 0.3631	(1) -0.0004 (2) 0.4288	(1) -0.0027 (2) 0.5938	(1) 0.0033 (2) 0.5198	(1) -0.0055 (2) 0.2788	(1) 0.0000 (2) 0.9944	(1) 0.0007 (2) 0.8884	(1) -0.0005 (2) 0.9221	(1) -0.0015 (2) 0.7726	(1) 0.0007 (2) 0.8898	

(1) Pearson correlation value
(2) P-value

Figure 3. Pearson's correlation values and p -values between predictor variables. The significance level of the p -values was 0.05.

2.2.2. Training and Validation Procedure

As indicated above, five datasets containing 7776 instances were obtained per decile. Likewise, a full dataset with the case studies of the five deciles was generated (38,880 observations). These datasets were used to generate 6 different models for each regression algorithm: one per each decile (D1, D2, D3, D4, and D5) and another per all deciles (T). The datasets were randomly divided into two sets: 75% of the observations corresponded to the training dataset, and the remaining 25% to the testing dataset.

For the models' training, a 10-fold cross-validation was carried out. Cross-validation allows the bias and the variance of the model to be reduced [106]. All training datasets were randomly divided into 10 subsets. For each fold, 9 subsets were used for the training, and the remaining for the testing (Figure 4). The performance of the model is obtained by the average value of the 10-fold.

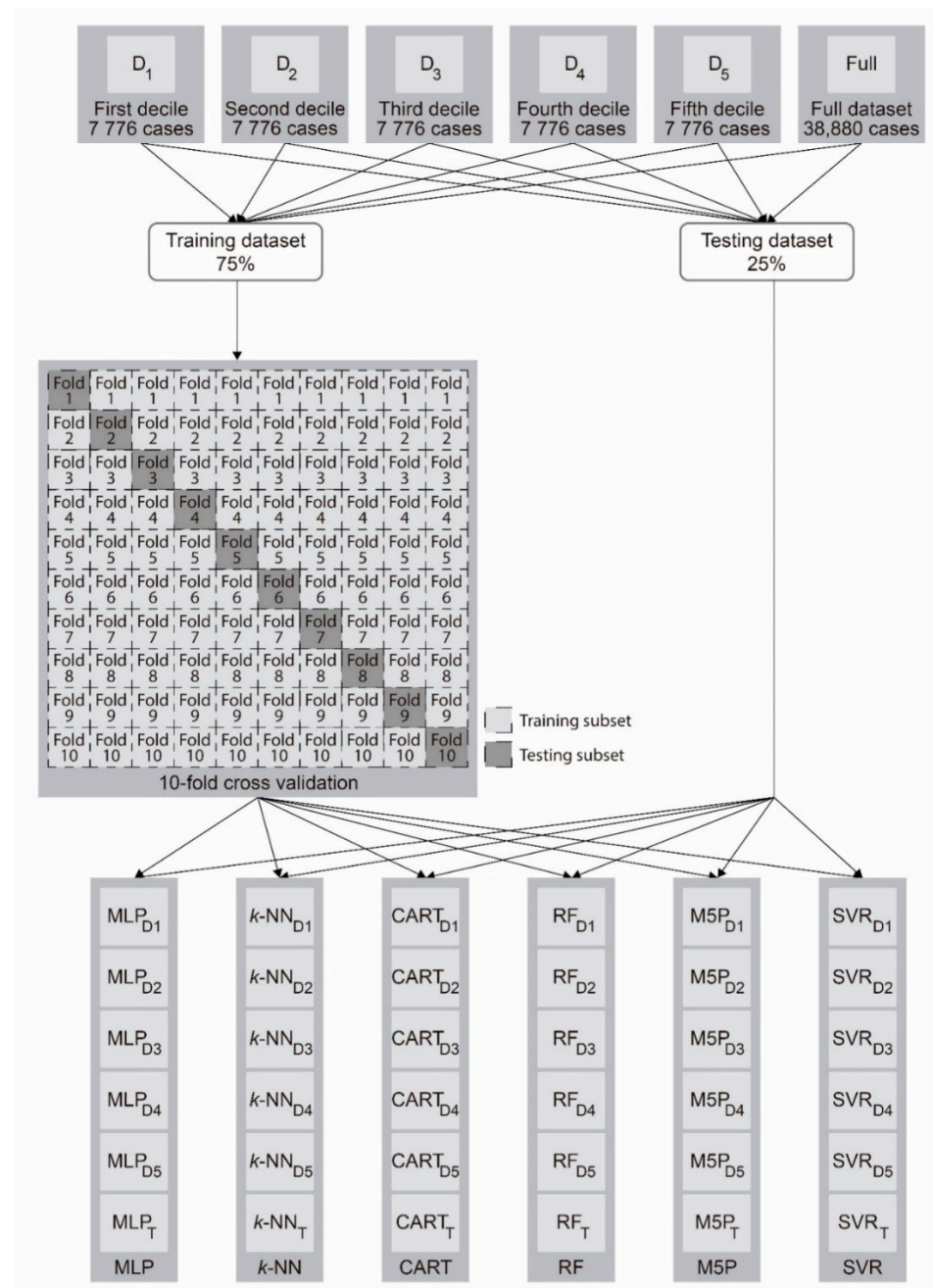


Figure 4. Flowchart of the training and testing.

Three statistical parameters were used to evaluate the performance of the models: (i) the correlation coefficient (R^2) (Equation (13)), the root mean square error ($RMSE$) (Equation (14)), and the mean absolute error (MAE) (Equation (15)). The use of these parameters allows the performance of the models to be correctly defined [107] and also are amongst the most used when comparing the performance of algorithms with different

architecture [43]. The lower the value of $RMSE$ and MAE , the greater the accuracy. The threshold value of R^2 was set at 0.95 to consider the predictions sufficiently accurate [37,48].

$$R^2 = \left(1 - \frac{\sum_{i=1}^n (t_i - m_i)^2}{\sum_{i=1}^n (t_i - \bar{t}_i)^2} \right) \quad (13)$$

$$RMSE = \left(\frac{\sum_{i=1}^n (t_i - m_i)^2}{n} \right)^{1/2} \quad (14)$$

$$MAE = \frac{\sum_{i=1}^n |t_i - m_i|}{n} \quad (15)$$

where m_i is the model's prediction, t_i is the actual value, and n is the number of instances in the dataset.

3. Results and Discussion

This section is organized as follows: First, the results from each algorithm are presented, along with a discussion of the particular features of their architecture. Finally, all algorithms are compared on the basis of the three parameters: R^2 , MAE , and $RMSE$.

3.1. Performance of the Models

3.1.1. MLP Models

The performance of the MLPs model was analyzed on the basis of the number of neurons in the hidden layer. As mentioned in Section 2.1, only one hidden layer was considered because it usually delivers better performance than more complex architectures [58]. Table 3 indicates the optimal number of neurons obtained for each model and the performance obtained in both the training and testing phases. What stands out in this table is that the optimal number of neurons in the hidden layer was very similar between the different models (between 12 and 14 nodes). Regarding the accuracy, R^2 was greater than 99.8% for all models in the training and testing phases, MAE was lower than 0.010, and $RMSE$ lower than 0.012. The accuracy did not vary when considering the full dataset or each decile separately.

Table 3. Results of training and testing of the MLP models.

Model and Dataset	Number of Nodes	Training			Testing		
		R^2	MAE	$RMSE$	R^2	MAE	$RMSE$
MLP _{D1} (D ₁)	13	0.998	0.009	0.012	0.998	0.010	0.012
MLP _{D2} (D ₂)	13	0.999	0.009	0.012	0.999	0.009	0.011
MLP _{D3} (D ₃)	12	0.998	0.010	0.013	0.999	0.009	0.012
MLP _{D4} (D ₄)	12	0.998	0.011	0.014	0.999	0.008	0.011
MLP _{D5} (D ₅)	13	0.998	0.009	0.012	0.999	0.009	0.011
MLP _T (Full _{dataset})	14	0.999	0.009	0.011	0.999	0.008	0.010

3.1.2. K-NN Models

For K -NN models, two parameters, q and K , have an influence on its accuracy. In this case, values of K ranged between 1 and 25, and values of q between 1 and 4 (Figure 5). A different behavior was observed depending on the type of dataset used: for models per deciles, a value of q of 2 (corresponding to the Euclidean distance) obtained the best performance, whereas, for K -NN_T, a value of q of 4 obtained the best statistical parameters. In this regard, the number of close instances used by the model varied: for models per deciles, values of K between 9 and 11 obtained acceptable results (in the testing, MAE was between 0.059 and 0.072, and $RMSE$ between 0.076 and 0.094), whereas, for K -NN_T, the number of K was 2. The performance of K -NN_T was lower than models per deciles because a low number of instances was used to obtain a representative result (in the testing, MAE

was 0.082 and *RMSE* was 0.110). In this sense, both in the training and testing phases, R^2 was always below 0.95, except for one case, *MAE* and *RMSE* were very similar for all sets of data (Table 4).

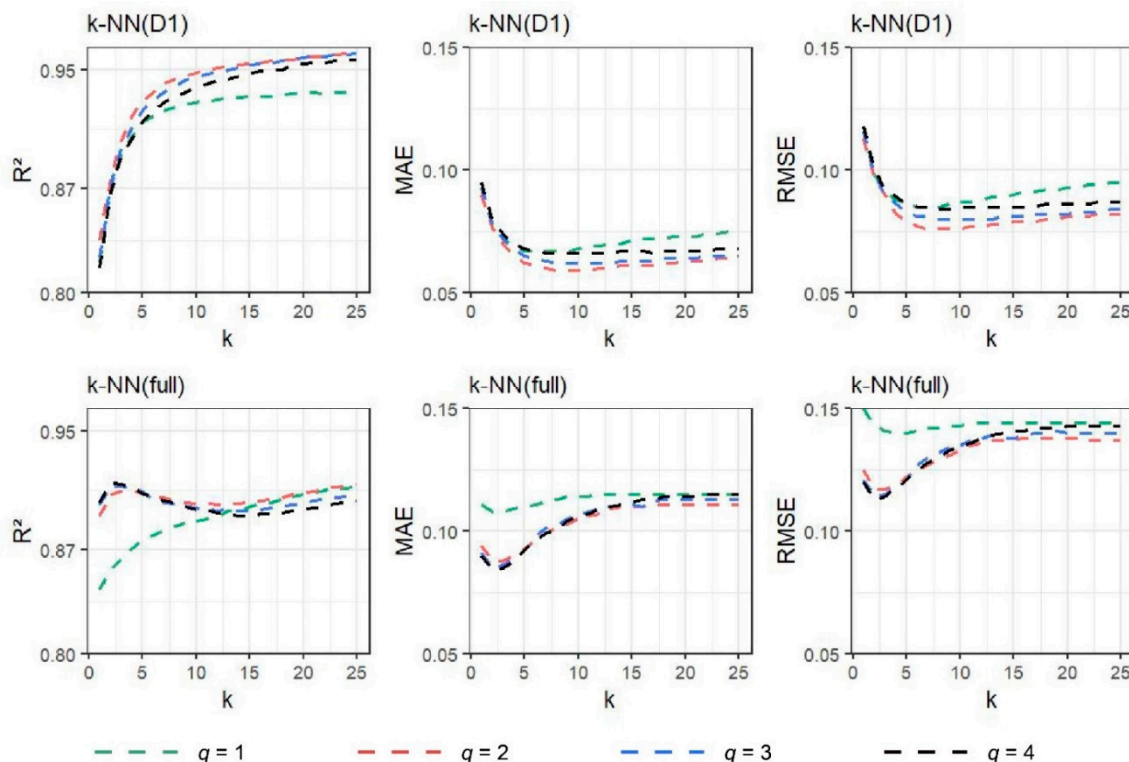


Figure 5. Values of R^2 , *MAE*, and *RMSE* obtained for K and q .

Table 4. Results of training and testing of the K -nearest neighbors (K -NN) models.

Model and Dataset	Training					Testing		
	q	K	R^2	<i>MAE</i>	<i>RMSE</i>	R^2	<i>MAE</i>	<i>RMSE</i>
K -NN _{D1} (D ₁)	2	10	0.948	0.059	0.076	0.952	0.059	0.076
K -NN _{D2} (D ₂)	2	9	0.943	0.071	0.092	0.942	0.072	0.094
K -NN _{D3} (D ₃)	2	11	0.946	0.069	0.089	0.947	0.068	0.088
K -NN _{D4} (D ₄)	2	10	0.946	0.066	0.085	0.944	0.066	0.086
K -NN _{D5} (D ₅)	2	10	0.945	0.065	0.085	0.948	0.063	0.082
K -NN _T (Full _{dataset})	4	2	0.915	0.084	0.113	0.917	0.082	0.110

3.1.3. CART Models

In this model, the parameters of depth and instances per node were analyzed: the depth of the tree oscillated between 2 and 12 (the maximum value was established according to the number of variables of the dataset), and the number of observations per node ranged between 2 and 30. Figure 6 is a sample of the analysis carried out. The performance of the model is strongly influenced by the depth of the tree: the higher the depth, the greater the correlation (values of R^2 greater than 97.80%) and the lower the *MAE* and *RMSE*. The most optimal depth for all models was, therefore, the maximum suggested (12 levels). Although this could mean overfitting to training data, the performance obtained in the testing phase was similar to that obtained in training (Table 5); therefore, it was not considered necessary to generalize the tree (i.e., to simplify it) to obtain a better performance in new observations. When comparing the accuracy for the full dataset and for each decile, the full model ($CART_T$) delivered better results than each decile separately in the testing phase. R^2 increased between 1.0 and 1.3%, and *MAE* and *RMSE* had values

lower than 0.008 and 0.010, respectively. This is because the structure obtained is better adjusted to the full dataset than to the models per deciles.

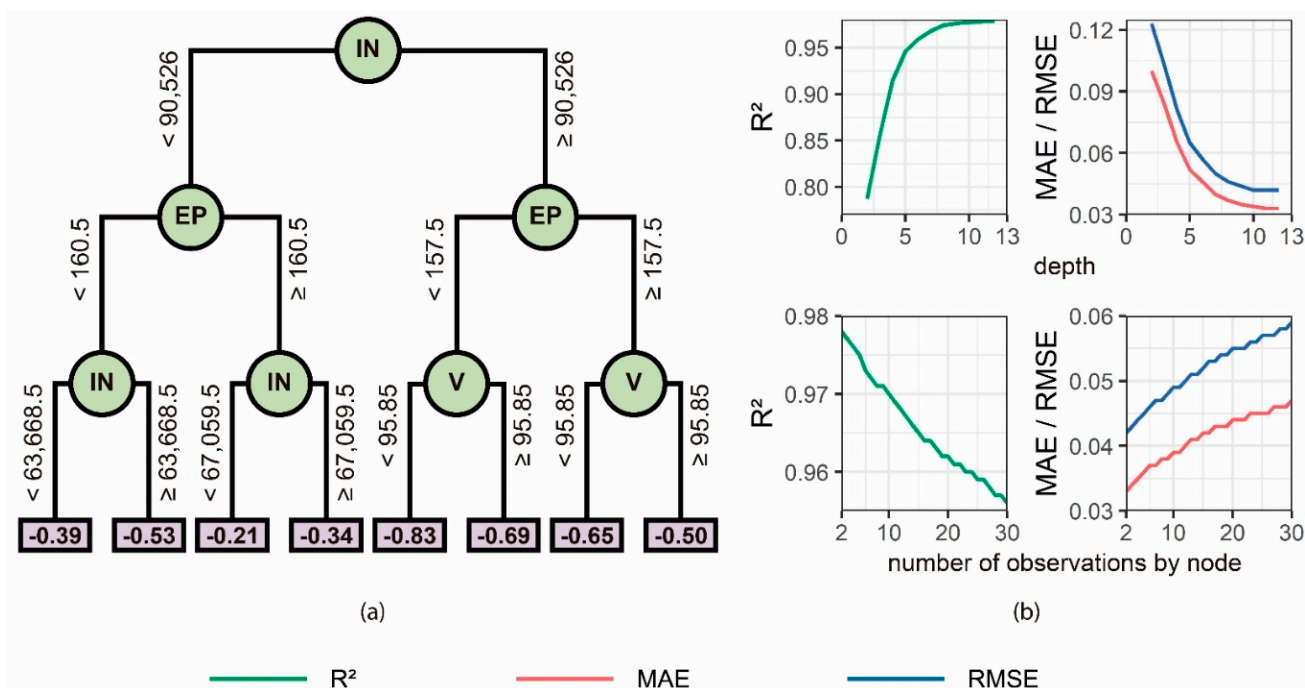


Figure 6. Training of the classification and regression tree (CART)_{D1} model: (a) scheme of the CART model for a depth of 3; and (b) statistical parameters both of different values of depth and the number of observations by node.

Table 5. Results of training and testing of the CART models.

Model and Dataset	Training			Testing		
	R^2	MAE	RMSE	R^2	MAE	RMSE
CART _{D1} (D ₁)	0.978	0.033	0.042	0.980	0.033	0.041
CART _{D2} (D ₂)	0.982	0.035	0.045	0.983	0.035	0.044
CART _{D3} (D ₃)	0.981	0.035	0.044	0.981	0.034	0.043
CART _{D4} (D ₄)	0.980	0.034	0.043	0.981	0.034	0.042
CART _{D5} (D ₅)	0.980	0.035	0.043	0.981	0.034	0.042
CART _T (Full _{dataset})	0.992	0.027	0.034	0.993	0.027	0.034

3.1.4. RF Models

In this model, the number of trees influences the performance of the model. For this reason, the behavior of the model with a number of trees between 1 and 200 was analyzed by using both the same depth and number of instances used in the CART models. Figure 7 represents the performance obtained for the RF_{D1} model, whereas Table 6 indicates the values obtained for all models. As can be seen in Figure 7, the performance of the model improved as the number of trees increased. However, when the optimal value was reached (149 in the case of D1), the performance of the model did not improve: the number of trees only increased the time required to train the model. Likewise, the optimal number obtained for each model was very similar, varying between 143 and 153. With respect to the models obtained, the performance was very similar for models per deciles and the full model. In this sense, R^2 obtained values between 99.3 and 99.6%, whereas MAE and RMSE oscillated between 0.021 and 0.023 and between 0.027 and 0.030, respectively.

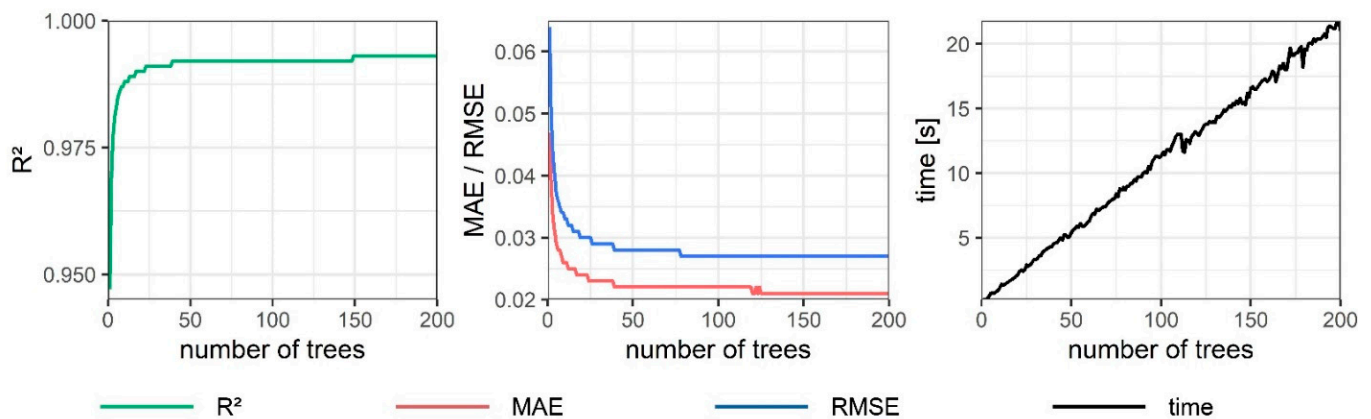


Figure 7. Influence of the number of trees on the behavior of RF_{D1} .

Table 6. Results of training and testing of the RF models.

Model and Dataset	Number of Trees	Training			Testing		
		R^2	MAE	RMSE	R^2	MAE	RMSE
RF_{D1} (D_1)	149	0.993	0.021	0.027	0.993	0.022	0.027
RF_{D2} (D_2)	148	0.994	0.023	0.030	0.994	0.023	0.029
RF_{D3} (D_3)	148	0.993	0.023	0.029	0.994	0.022	0.028
RF_{D4} (D_4)	143	0.993	0.022	0.028	0.993	0.022	0.028
RF_{D5} (D_5)	144	0.993	0.022	0.028	0.994	0.021	0.027
RF_T (Full _{dataset})	153	0.996	0.022	0.028	0.996	0.022	0.028

3.1.5. M5P Models

In M5P models, no parameter was analyzed to optimize their performance. As can be seen in Figure 8, M5P determines the optimal depth of the model. Therefore, it is not necessary to analyze this aspect in such performances. Regarding the values of the statistical parameters, the performance was very similar (Table 7): R^2 was between 0.995 and 0.998, MAE between 0.014 and 0.016, and RMSE between 0.018 and 0.020.

3.1.6. SVR Models

As indicated in Section 2.1.6, both C and γ affect the performance of the SVR models. Different combinations of these two parameters were therefore analyzed. In this sense, C adopted values between 0.1 and 4, whereas γ was analyzed using values between 0.05 and 0.2. Figure 9 represents the results of the various combinations in SVR_{D1} , and Table 8 indicates the optimal configuration and the results obtained by each model. Several aspects can be appreciated from the analysis: (i) optimal values of the models were obtained from values of γ of 0.15. Although best values in the statistical parameters were obtained as the value of γ increased, the value of C obtained the same values for lower values of γ . Values γ lower than 0.15 did not obtain the same optimal performance; (ii) high values of γ and C influenced the training time. In this regard, the time difference between different values of γ was greater than 1000 s, slowing down the training with high γ . Thus, the optimal value of γ obtained for all models was 0.15, whereas the value of C varied between 1.6 and 2.1 (Table 8). Likewise, the performance obtained both in the training and testing phases was similar. SVR_T presents R^2 quite similar to models per deciles in the testing, MAE is greater than 0.002, and RMSE ranges between 0.004 and 0.006.

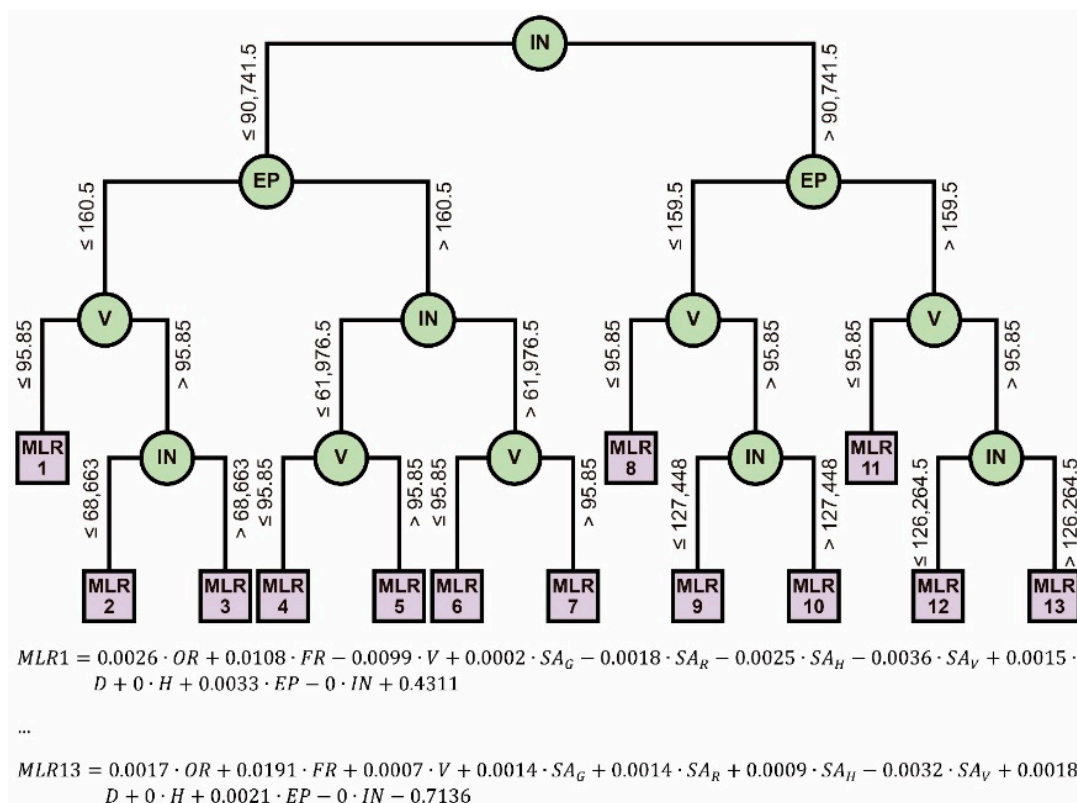


Figure 8. Structure of the M5P_{D1} model.

Table 7. Results of training and testing of the M5P models.

Model and Dataset	Training			Testing		
	R ²	MAE	RMSE	R ²	MAE	RMSE
M5P _{D1} (D ₁)	0.995	0.016	0.020	0.996	0.015	0.019
M5P _{D2} (D ₂)	0.997	0.016	0.020	0.996	0.016	0.020
M5P _{D3} (D ₃)	0.996	0.016	0.020	0.996	0.016	0.019
M5P _{D4} (D ₄)	0.996	0.015	0.019	0.996	0.016	0.019
M5P _{D5} (D ₅)	0.996	0.016	0.019	0.996	0.016	0.020
M5P _T (Full _{dataset})	0.998	0.014	0.018	0.998	0.014	0.018

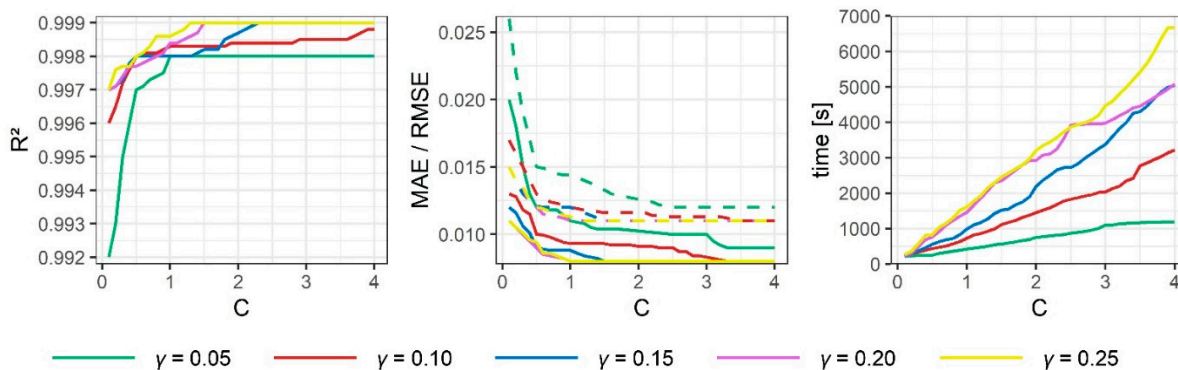


Figure 9. Influence of C and γ on the behavior of SVR_{D1}. Values of RMSE are represented by the discontinuous line.

Table 8. Results of training and testing of the support vector regression (SVR) models.

Model and Dataset	C	γ	Training			Testing		
			R^2	MAE	RMSE	R^2	MAE	RMSE
SVR _{D1} (D ₁)	2.0	0.15	0.999	0.008	0.011	0.998	0.009	0.012
SVR _{D2} (D ₂)	1.6	0.15	0.999	0.010	0.013	0.999	0.009	0.013
SVR _{D3} (D ₃)	1.9	0.15	0.999	0.009	0.012	0.999	0.009	0.011
SVR _{D4} (D ₄)	2.0	0.15	0.999	0.009	0.012	0.999	0.009	0.012
SVR _{D5} (D ₅)	2.0	0.15	0.999	0.009	0.012	0.999	0.009	0.012
SVR _T (Full _{dataset})	2.1	0.15	0.999	0.012	0.018	0.998	0.011	0.017

3.2. Comparison of Models

The algorithms were compared on the basis of two parameters: (I) the individual performance obtained by each decile and (ii) the performance obtained in the estimation of the 5 deciles together. Figure 10 shows the scatter plots comparing the actual and predicted values of the total dataset (as well as the values of the statistical parameters), and Table 9 is a compilation of the statistical parameters obtained per deciles separately. Likewise, the point obtained by each model per decile is represented in Appendix A (Figures A2–A6). Several aspects regarding the comparative analysis are worth noting. First, the estimation conducted by each model was merged in a combined estimation (D1–5) with the corresponding values of R^2 , MAE, and RMSE. Then, the full model of each algorithm was used to evaluate the estimation conducted in each decile.

Regarding the performance obtained by each decile, there were different results. The algorithm that delivered with the worst performance was K-NN: R^2 was lower than 95% for almost all decile models (only D1 obtained R^2 of 95%), whereas K-NN_T had a different behavior depending on the decile analyzed, obtaining the best behavior in deciles 3, 4, and 5, and the worst in deciles 1 and 2 (0.880 and 0.939). Therefore, the performance obtained using the K-NN algorithm was not considered adequate (see Figure 10).

The use of the other algorithms gave optimal correlation coefficients (higher than 98% in all cases). The analysis of the error parameters revealed that CART and RF models obtained error values higher than MLP, M5P, and SVR: MAE and RMSE increased regarding the highest value of M5P of 21.43% and 23.53%, respectively. Therefore, the performance obtained by these two models, although being acceptable, did not present the adjustment degree of MLP, M5P, and SVR. Regarding these three algorithms, the results were quite similar for all the cases analyzed: R^2 oscillated between 0.996 and 0.999, MAE between 0.008 and 0.018, and RMSE between 0.010 and 0.026. Despite the performance obtained by the full models in each decile was optimal, SVR_T had the worst performance in deciles 1 and 2. This aspect influenced the estimation conducted by the full model, which was worse than that conducted by the set of decile models (MAE decreased by 0.002, and RMSE by 0.005). On the other hand, the scatter plots obtained by MLP and M5P were similar, both with an adequate adjustment.

Thus, these two algorithms obtained the best performance, particularly if the time required to train the SVR models is considered (Table 10). The time needed to carry out the training of these models was more than 1990s, and SVR_T needed more than 18 h for the training. The remaining algorithms needed quite short times: MLP and K-NN consumed more time (391.2 and 653.2 s, respectively).

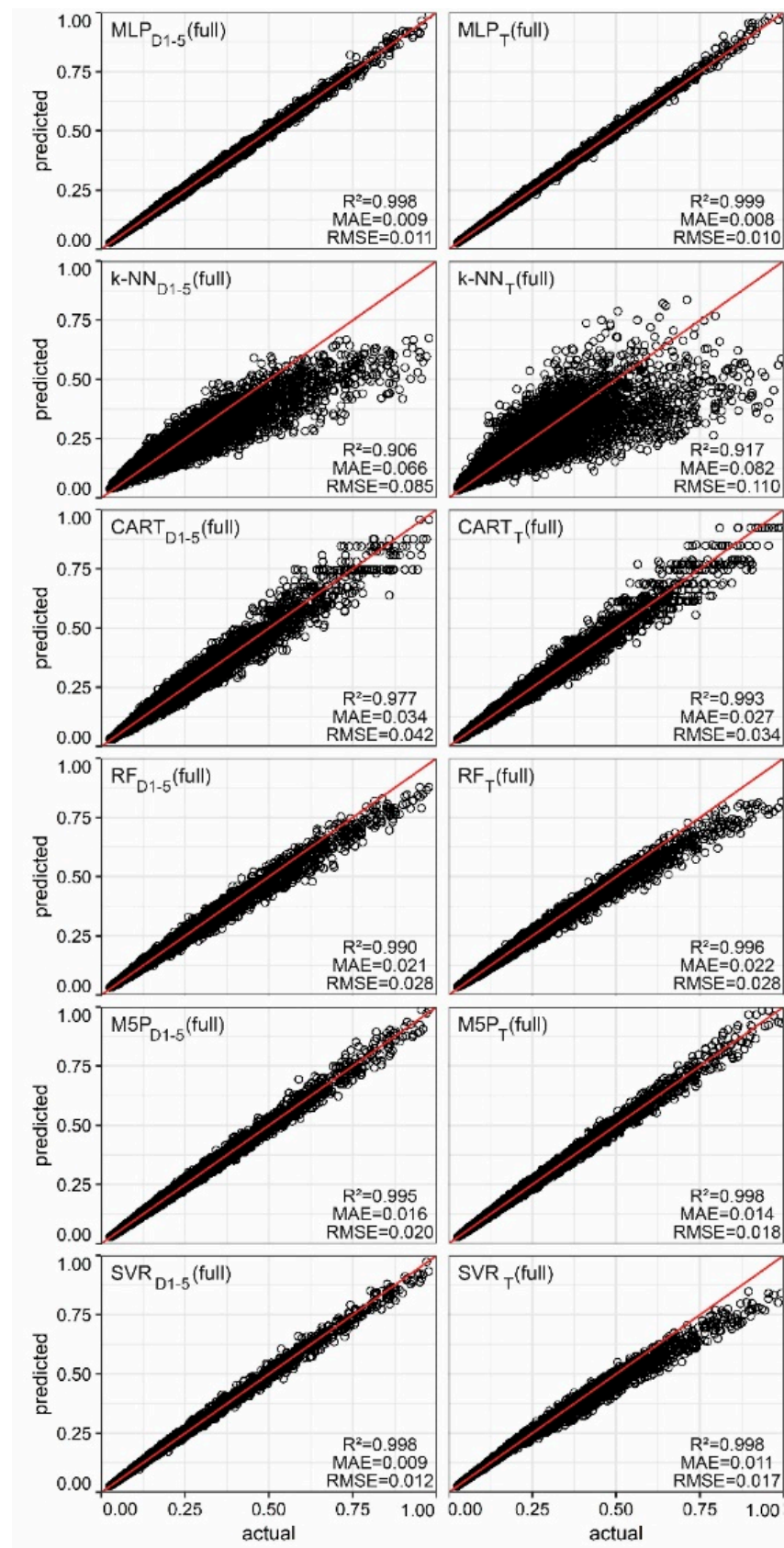


Figure 10. Comparison of actual and predicted values obtained by each model. The testing dataset used is indicated in brackets.

Table 9. Statistical parameters were obtained by each regression algorithm in the testing phase per each decile.

Test Per Each Decile Model				Test Per Total Model			
Model	R ²	MAE	RMSE	Model	R ²	MAE	RMSE
MLP				MLP			
MLP _{D1} (D ₁)	0.998	0.010	0.012	MLP _T (D ₁)	0.999	0.008	0.011
MLP _{D2} (D ₂)	0.999	0.009	0.011	MLP _T (D ₂)	0.999	0.008	0.010
MLP _{D3} (D ₃)	0.999	0.009	0.012	MLP _T (D ₃)	0.999	0.008	0.010
MLP _{D4} (D ₄)	0.999	0.008	0.011	MLP _T (D ₄)	0.999	0.008	0.010
MLP _{D5} (D ₅)	0.999	0.009	0.011	MLP _T (D ₅)	0.999	0.008	0.010
K-NN				K-NN			
K-NN _{D1} (D ₁)	0.952	0.059	0.076	K-NN _T (D ₁)	0.880	0.077	0.106
K-NN _{D2} (D ₂)	0.942	0.072	0.094	K-NN _T (D ₂)	0.939	0.058	0.082
K-NN _{D3} (D ₃)	0.947	0.068	0.088	K-NN _T (D ₃)	0.963	0.045	0.062
K-NN _{D4} (D ₄)	0.944	0.066	0.086	K-NN _T (D ₄)	0.968	0.043	0.059
K-NN _{D5} (D ₅)	0.948	0.063	0.082	K-NN _T (D ₅)	0.967	0.045	0.062
CART				CART			
CART _{D1} (D ₁)	0.980	0.033	0.041	CART _T (D ₁)	0.988	0.023	0.031
CART _{D2} (D ₂)	0.983	0.035	0.044	CART _T (D ₂)	0.992	0.023	0.030
CART _{D3} (D ₃)	0.981	0.034	0.043	CART _T (D ₃)	0.992	0.023	0.029
CART _{D4} (D ₄)	0.981	0.034	0.042	CART _T (D ₄)	0.992	0.022	0.028
CART _{D5} (D ₅)	0.981	0.034	0.042	CART _T (D ₅)	0.991	0.023	0.029
RF				RF			
RF _{D1} (D ₁)	0.993	0.022	0.027	CART _T (D ₁)	0.994	0.020	0.027
RF _{D2} (D ₂)	0.994	0.023	0.029	CART _T (D ₂)	0.996	0.018	0.024
RF _{D3} (D ₃)	0.994	0.022	0.028	CART _T (D ₃)	0.996	0.017	0.021
RF _{D4} (D ₄)	0.993	0.022	0.028	CART _T (D ₄)	0.996	0.017	0.021
RF _{D5} (D ₅)	0.994	0.021	0.027	CART _T (D ₅)	0.995	0.021	0.028
M5P				M5P			
M5P _{D1} (D ₁)	0.996	0.015	0.019	M5P _T (D ₁)	0.996	0.014	0.018
M5P _{D2} (D ₂)	0.996	0.016	0.020	M5P _T (D ₂)	0.997	0.014	0.017
M5P _{D3} (D ₃)	0.996	0.016	0.019	M5P _T (D ₃)	0.997	0.014	0.017
M5P _{D4} (D ₄)	0.996	0.016	0.019	M5P _T (D ₄)	0.997	0.014	0.017
M5P _{D5} (D ₅)	0.996	0.016	0.020	M5P _T (D ₅)	0.997	0.015	0.018
SVR				SVR			
SVR _{D1} (D ₁)	0.998	0.009	0.012	SVR _T (D ₁)	0.995	0.018	0.026
SVR _{D2} (D ₂)	0.999	0.009	0.013	SVR _T (D ₂)	0.997	0.013	0.021
SVR _{D3} (D ₃)	0.999	0.009	0.011	SVR _T (D ₃)	0.999	0.009	0.011
SVR _{D4} (D ₄)	0.999	0.009	0.012	SVR _T (D ₄)	0.999	0.008	0.011
SVR _{D5} (D ₅)	0.999	0.009	0.012	SVR _T (D ₅)	0.999	0.008	0.011

Table 10. Time required in the training of models.

Algorithm	Time (s)					
	Model					
	D1	D2	D3	D4	D5	T
MLP	77.3	71.6	66.0	66.1	71.2	391.2
K-NN	5.3	5.6	5.7	5.8	5.6	653.2
CART	2.4	2.3	2.1	2.0	2.1	11.3
RF	20.3	20.6	21.3	20.2	21.1	95.1
M5P	5.1	4.9	5.1	5.4	5.2	26.2
SVR	2188.2	1991.3	2195.3	2130.3	2142.0	67,771.3

4. Conclusions

The present exploratory study set out with the aim of clarifying the best algorithm to predict FPPRI of financially deprived households in Chile. Considering their performance, it can be concluded that multilayer perceptron, support vector regression and M5P are, in this order, the best option. Correlation coefficients were higher than 99.5%, and the values

of MAE and RMSE were lower than 0.018 and 0.026, respectively. Including computing time as a variable, SVR should be discarded in favor of M5P and MLP. Finally, since M5P delivers the shortest computing time for all of them, it could be labeled as the best option in terms of speed and performance.

Comparison of the findings with those of a previous study by the authors that assessed the accuracy of multiple linear regression (MLR) models also confirms that M5P outperforms the former [37]. The MLR delivered values of R² between 0.807 and 0.963 and values of RMSE between 0.013 and 0.091. A possible explanation for this might be the architecture of the M5P model itself, which combines a CART algorithm with an MLR model in each subregion; that is, each leaf of the tree has one MLR model, which is the regression algorithm used in the previous study. As pointed out by other authors, the particular architecture of this model seems to be suited for the prediction of the variations of energy prices [91] and the energy consumption [89], as in our case. In terms of computing time, the CART algorithm was the fastest, and adding an MLR model to each leaf doubles computing time but still places this algorithm as the second-fastest. Additionally, this study has clarified that the design of the M5P model should follow some recommendations. Even though the model automatically determines the depth of the tree, caution should be exercised so as not to consider more than 150 trees, as the results from the original CART model suggest.

These results also have two main methodological implications. First, the M5P models can be easily programmed in different languages because it is composed by if-then rules, and also because it divides the input space into a linear regression model, delivering an equation whose mathematical meaning is easier to grasp. Second, the FPPRI was tested against a nationally representative sample size in the Chilean context, but the methodology can be extrapolated to other countries by making the necessary adjustments in the predictor variables.

The most important limitation of the present study lies in the fact that the tested algorithms are data-driven or black-box models; that is, all data were artificially generated. Further studies need to be carried out using real data from the energy expenditure of Chilean households in order to validate the results from the present study. The biggest challenge would be compiling a statistically representative pool of data about the energy expenditure of households. The CASEN survey still does not cover this aspect, which should be considered in the future as part of the so-called “multidimensional poverty”.

Overall, and considering its scientific contribution and also its limitation, this study has provided a deeper insight into the prediction of fuel poverty by using different data-driven algorithms. These findings may have a number of important implications for the design of future policies tackling energy poverty before reallocating families into a new house, thus providing the basis towards a sustainable development of social housing that considers the economic situation of financially deprived households.

Author Contributions: Conceptualization, methodology, investigation, writing—original draft preparation, and editing and visualization, D.B.-H., J.A.P.-A., C.R.-B. and A.P.-F. All authors have read and agreed to the published version of the manuscript.

Funding: This research has been funded by the “Consejería de Economía, Conocimiento, Empresas y Universidad de la Junta de Andalucía” through a postdoctoral research programme at University of Sevilla.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would also like to acknowledge that this paper is part of the project “Conicyt Fondecyt Regular 1200551-Energy poverty prediction based on social housing architectural design in the central and central-southern zones of Chile: an innovative index to analyze and reduce the risk of energy poverty” funded by the National Agency for Research and Development (ANID).

In addition, we would like to acknowledge to the research group “Confort ambiental y pobreza energética (+CO-PE)” of the University of the Bío-Bío for supporting this research.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

<i>Symbols</i>	
b	Bias term (support vector regression)
C	Penalty parameter (support vector regression)
d_q	Minkowski distance (K -nearest neighbors)
EC	Energy consumption
EC_R	Real energy consumption
EC_S	Simulated energy consumption
EC_{SC}	Average monthly cooling energy consumption starting from the simulation
EC_{SE+L}	Average monthly energy consumption of equipment and lighting from the simulation
EC_{SH}	Average monthly heating energy consumption from the simulation
FPI	Fuel poverty index
$FPI_{adaptive}$	Fuel poverty index with the approach by Pérez-Fargallo et al.
H_c	Occupied hours in thermal comfort applying category III of EN 15,251
H_d	Unoccupied hours of the analyzed period
H_t	Total hours of the analyzed period
I	Household income
K	Nearest neighbors' number (K -nearest neighbors)
$K(x, x_i)$	Kernel function (support vector regression)
m_i	Model's prediction
MAE	Mean absolute error
n	Number of instances in the dataset
N	Sampling number (support vector regression)
P	Energy price used
P_C	Electricity price for cooling
P_{E+L}	Electricity price for equipment and lighting
P_H	Energy of fuel price for heating
R^2	Correlation coefficient
$RMSE$	Root mean square error
t_i	Actual value
T	Number of trees (random forest)
TI	Threshold income
w	Weight vector (support vector regression)
$w_{k0}^{(1)}$	Weight of the bias neuron of the input layer (multilayer perceptron)
$w_{kj}^{(1)}$	Weights of the hidden layer (multilayer perceptron)
$w_{i0}^{(2)}$	Weight of the bias neuron of the hidden layer (multilayer perceptron)
$w_{ik}^{(2)}$	Weights of the output layer (multilayer perceptron)
x_0	Input value of the bias neuron of the input layer (multilayer perceptron)
x_i	Predictor variables of multiple linear regression
x_j	Values of the input layer (multilayer perceptron)
y_0	Input value of the bias neuron of the hidden layer (multilayer perceptron)
y_k	Output of a neuron of the hidden layer (multilayer perceptron)
y_m	Values of the K nearest points (K -nearest neighbors)
\hat{Y}_{kNN}	Output value of K -nearest neighbors
\hat{Y}_{RF}	Output value of random forest
\hat{Y}_t	Output of the t -th tree (random forest)
<i>Greek letters</i>	
β_0	Independent term of multiple linear regression
β_i	Regression coefficients of multiple linear regression

δ_i	Lagrange multiplier (support vector regression)
δ_i^*	Lagrange multiplier (support vector regression)
ε	Error of multiple linear regression
ξ_i	Distance of the actual values from the upper limit of the error allowed (support vector regression)
ξ_i^*	Distance of the actual values from the lower limit of the error allowed (support vector regression)
μ_{D_n}	Average of income distribution of decile n
σ	Activation function (multilayer perceptron)
$\sigma_{D_n}^2$	Variance of income distribution of decile n
ψ	Maximum error allowed (support vector regression)

Abbreviations

BFGS	Broyden–Fletcher–Goldfarb–Shanno algorithm
CART	Classification and regression tree
D	Distance to the closest building
EP	Energy price
FPPRI	Fuel poverty potential risk index
FR	Form ratio
H	Shadow height
IN	Income
K-NN	K-nearest neighbors
MINVU	Chilean Ministry of Housing and Urbanism
MLP	Multilayer perceptron
MLR	Multiple linear regression
OR	Orientation
RF	Random forest
S_G	Surface in contact with the ground
S_H	Horizontal surface area in contact with other dwellings
S_R	Roof surface area
S_V	Vertical surface area in contact with other dwellings
SVR	Support vector regression
V	Volume

Appendix A

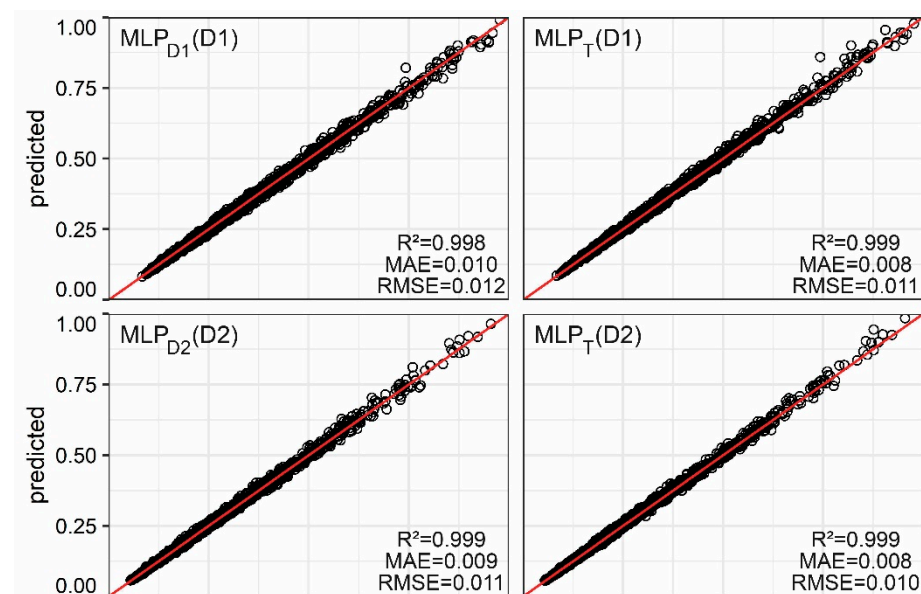


Figure A1. Cont.

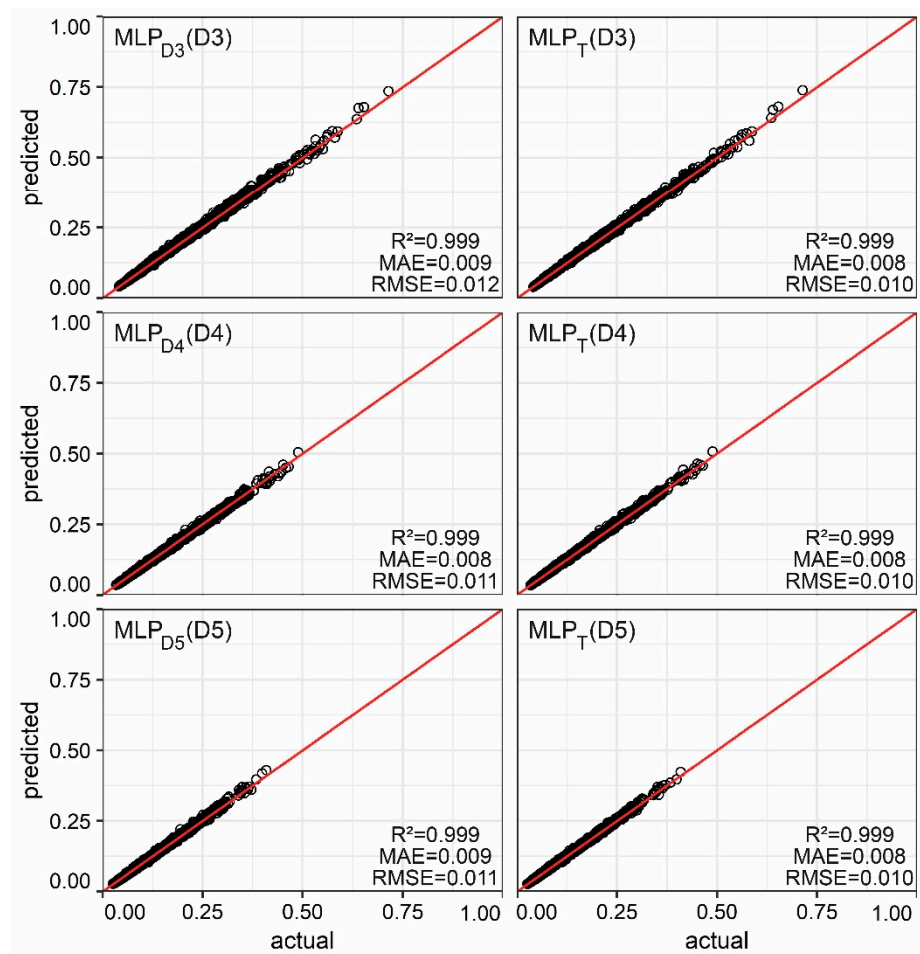


Figure A1. Point cloud between the actual and predicted values per deciles (D1, D2, D3, D4, and D5) obtained by the MLP models in the testing. The testing dataset used is indicated in brackets.

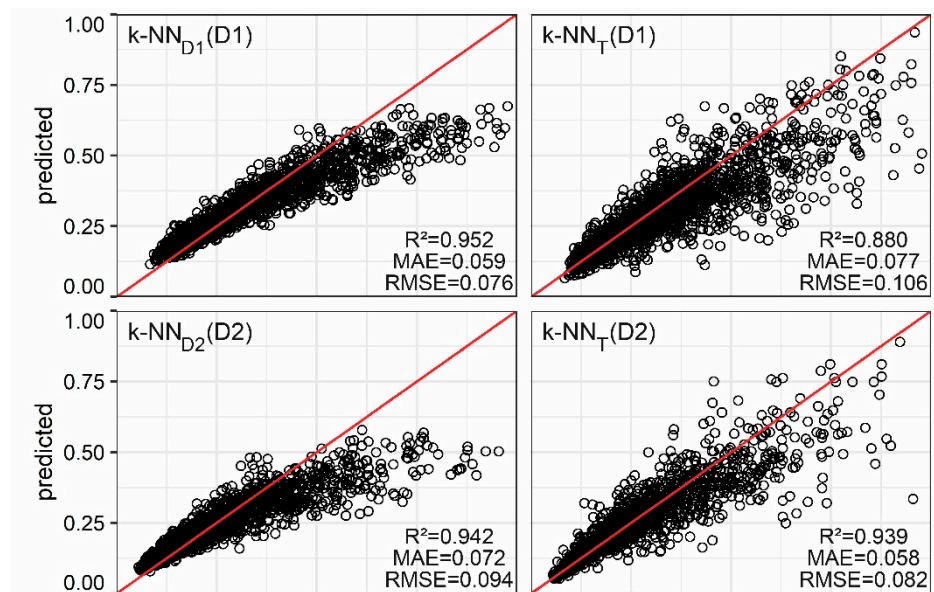


Figure A2. Cont.

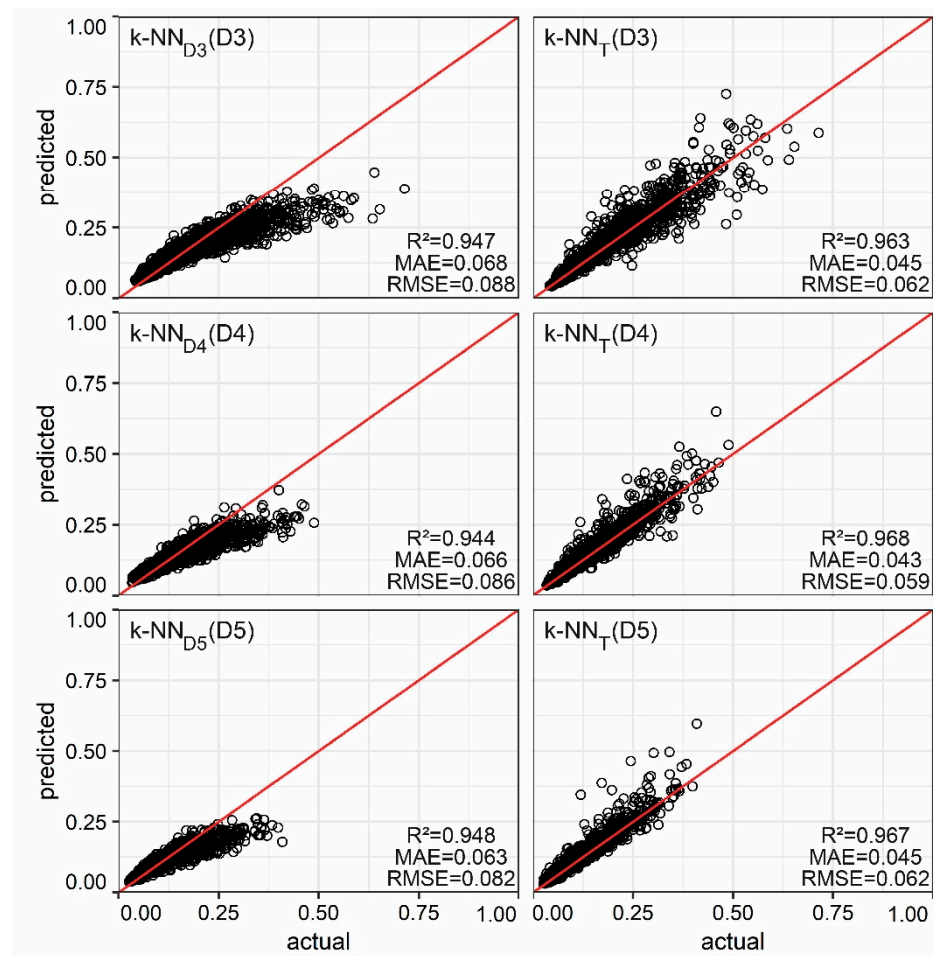


Figure A2. Point cloud between the actual and predicted values per deciles (D1, D2, D3, D4, and D5) obtained by the K-NN models in the testing. The testing dataset used is indicated in brackets.

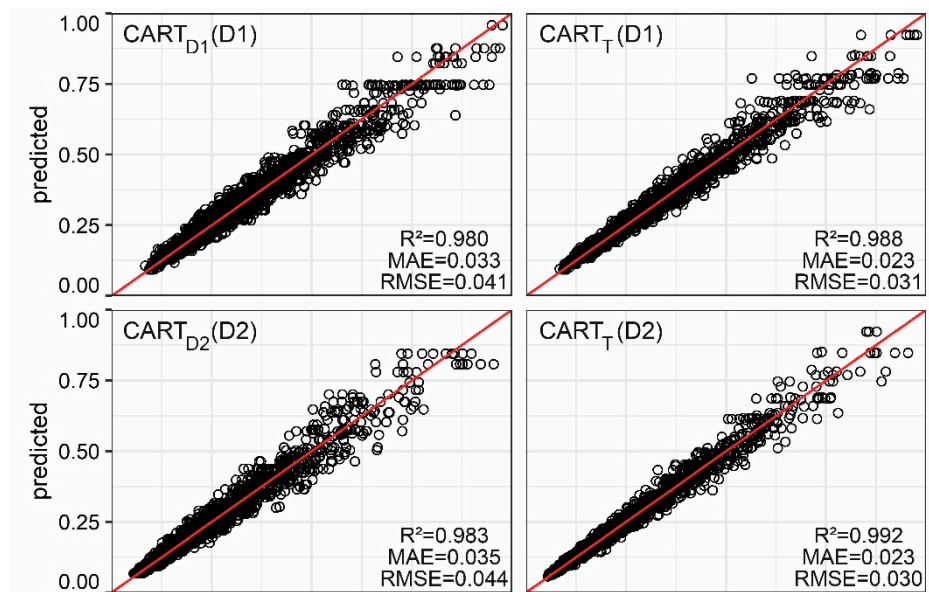


Figure A3. Cont.

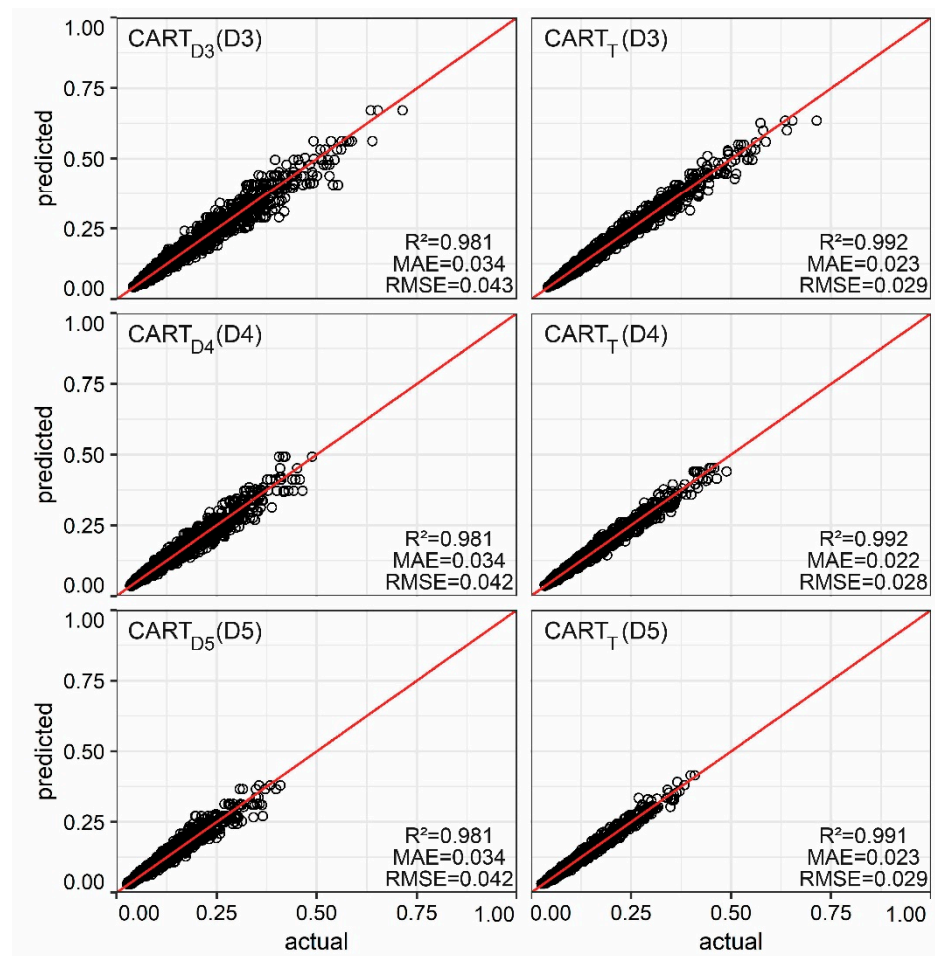


Figure A3. Point cloud between the actual and predicted values per deciles (D1, D2, D3, D4, and D5) obtained by the CART models in the testing. The testing dataset used is indicated in brackets.

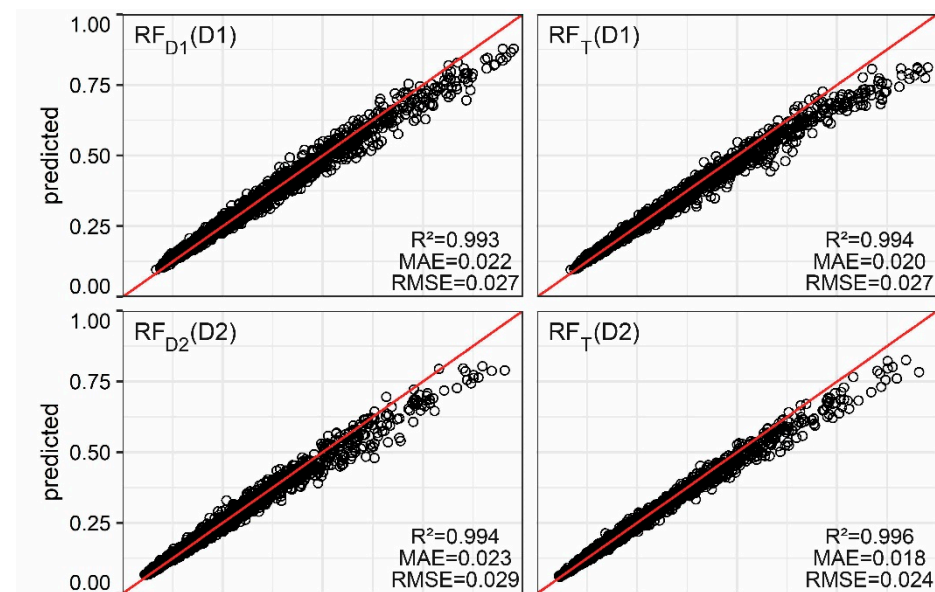


Figure A4. Cont.

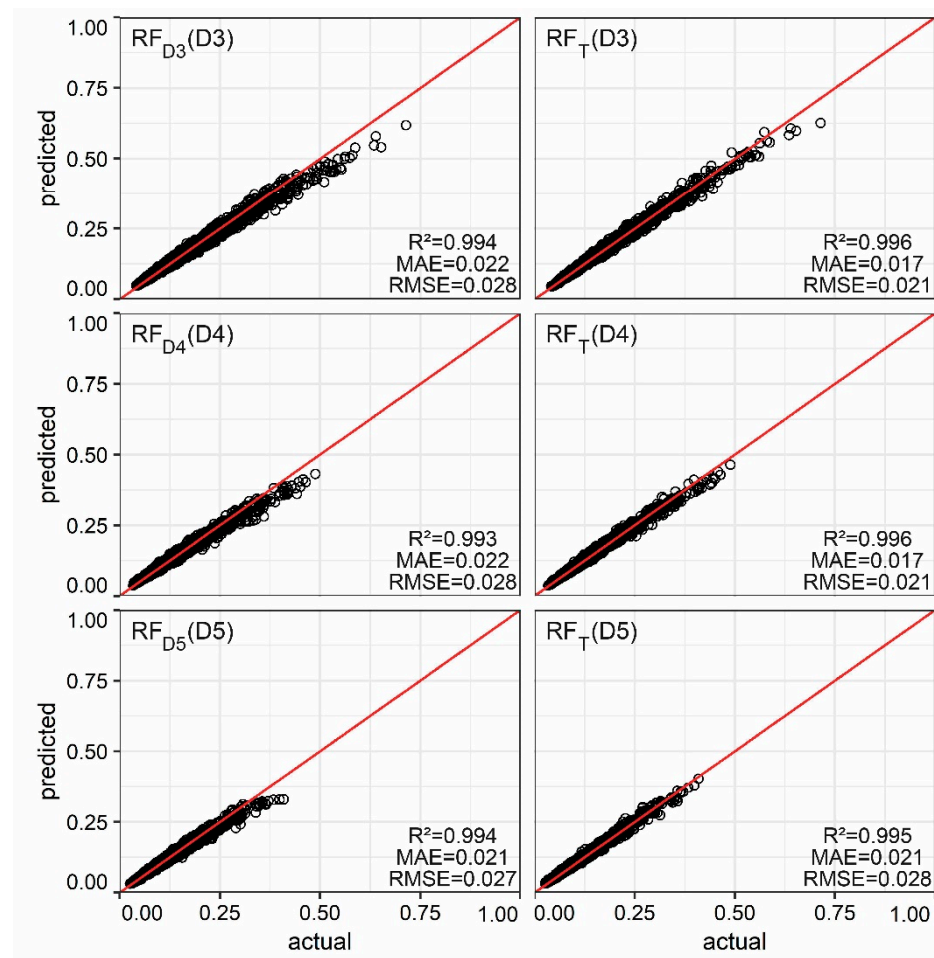


Figure A4. Point cloud between the actual and predicted values per deciles (D1, D2, D3, D4, and D5) obtained by the RF models in the testing. The testing dataset used is indicated in brackets.

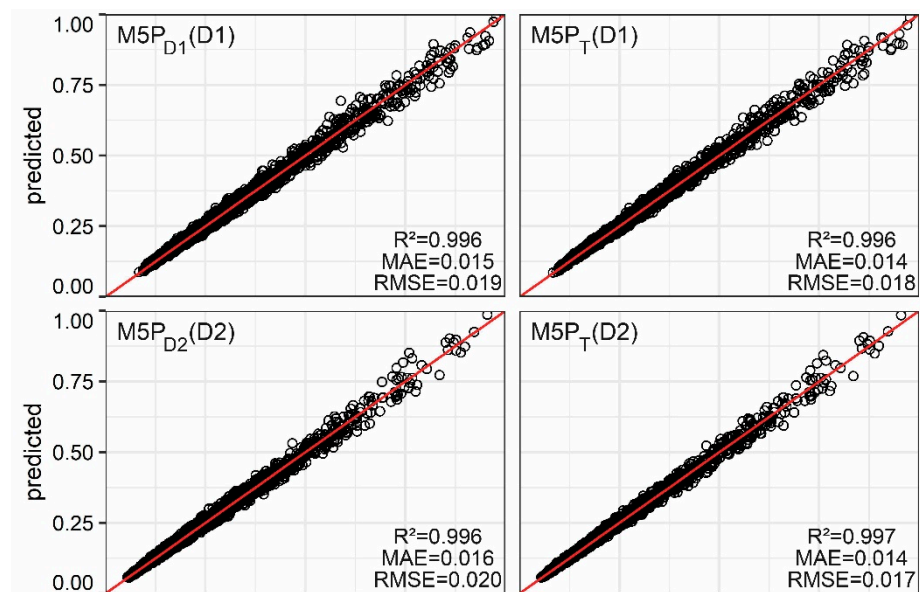


Figure A5. Cont.

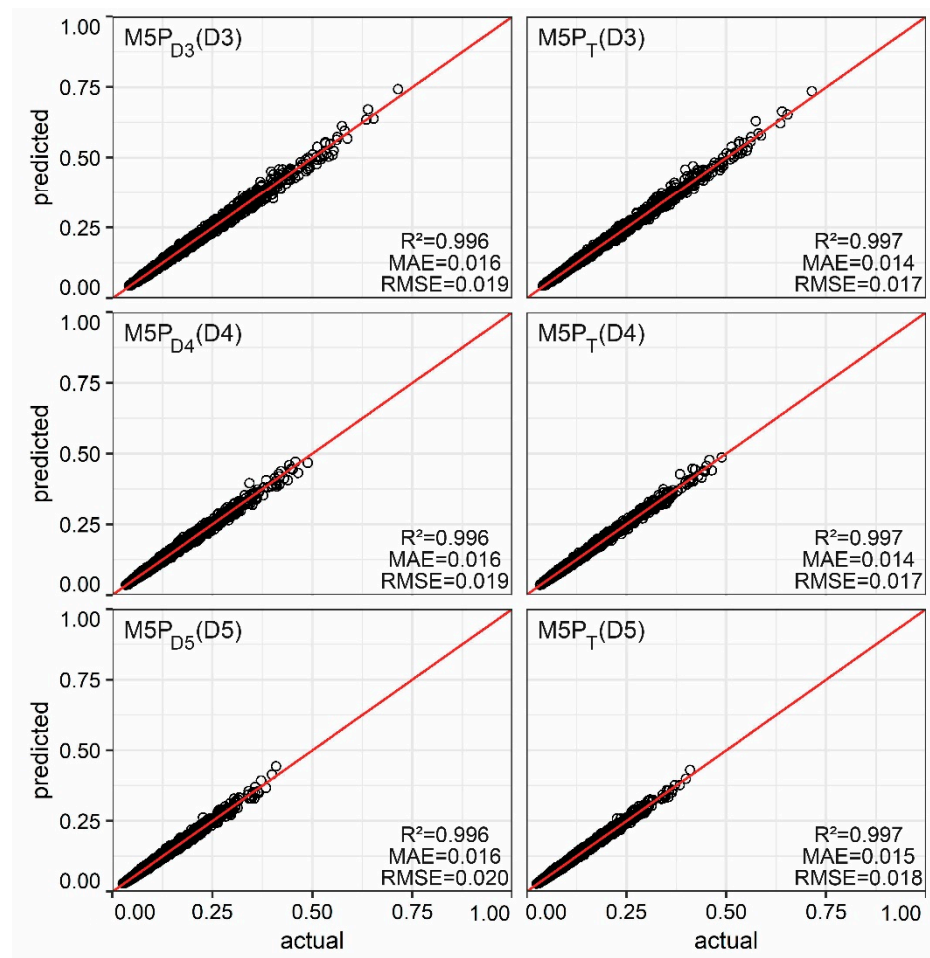


Figure A5. Point cloud between the actual and predicted values per deciles (D1, D2, D3, D4, and D5) obtained by the M5P models in the testing. The testing dataset used is indicated in brackets.

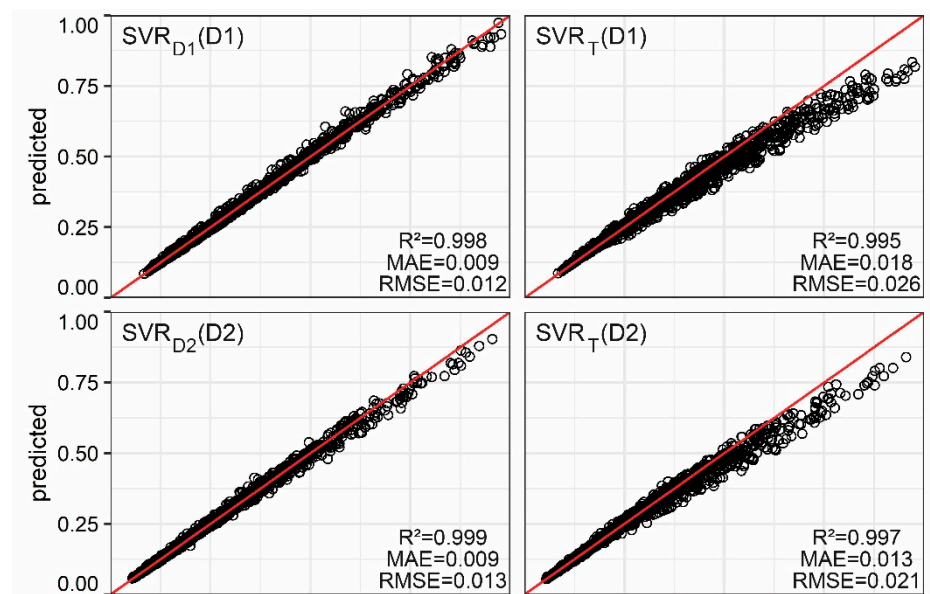


Figure A6. Cont.

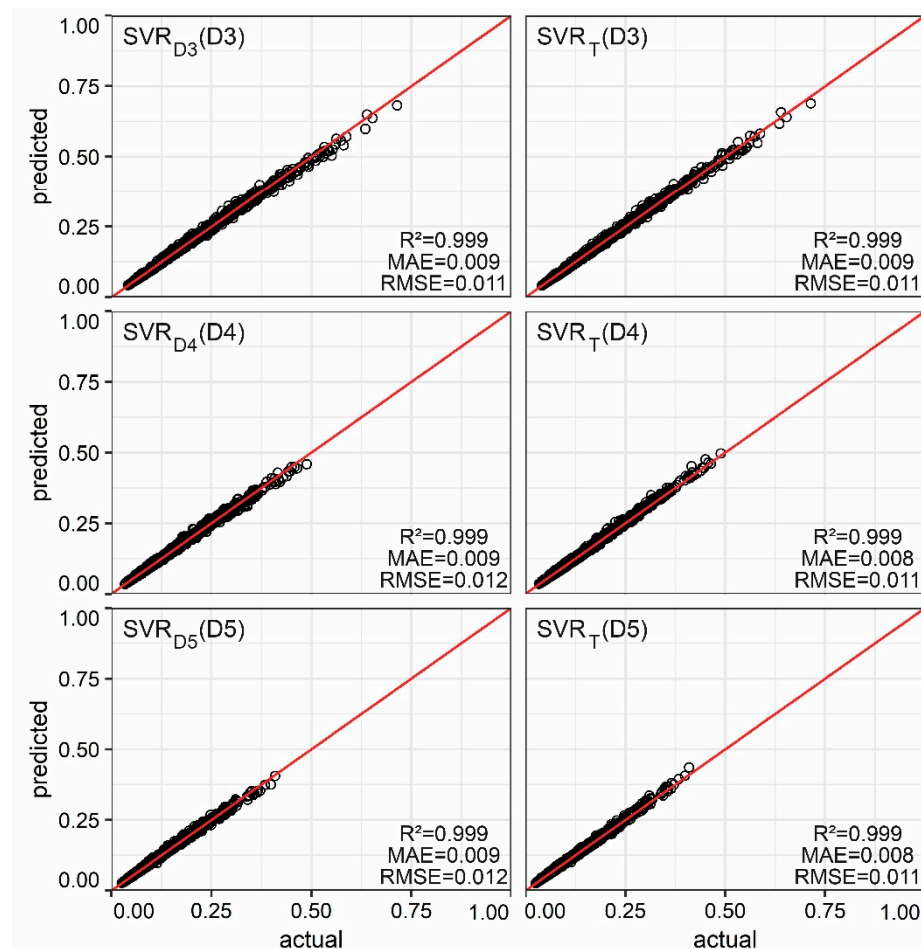


Figure A6. Point cloud between the actual and predicted values per deciles (D1, D2, D3, D4, and D5) obtained by the SVR models in the testing. The testing dataset used is indicated in brackets.

References

- Bouzarovski, S.; Petrova, S. A global perspective on domestic energy deprivation: Overcoming the energy poverty–fuel poverty binary. *Energy Res. Soc. Sci.* **2015**, *10*, 31–40. [[CrossRef](#)]
- Thomson, H.; Bouzarovski, S.; Snell, C. Rethinking the measurement of energy poverty in Europe: A critical analysis of indicators and data. *Indoor Built Environ.* **2017**, *26*, 879–901. [[CrossRef](#)] [[PubMed](#)]
- Pachauri, D.S. Energy Use and Energy Access in Relation to Poverty. *Econ. Polit. Wkly.* **2004**, *39*, 271–278.
- González-Eguino, M. Energy poverty: An overview. *Renew. Sustain. Energy Rev.* **2015**, *47*, 377–385. [[CrossRef](#)]
- Thomson, H.; Snell, C. Quantifying the prevalence of fuel poverty across the European Union. *Energy Policy* **2013**, *52*, 563–572. [[CrossRef](#)]
- Middlemiss, L.; Gillard, R. Fuel poverty from the bottom-up: Characterising household energy vulnerability through the lived experience of the fuel poor. *Energy Res. Soc. Sci.* **2015**, *6*, 146–154. [[CrossRef](#)]
- Liddell, C.; Morris, C.; Thomson, H.; Guiney, C. Excess winter deaths in 30 European countries 1980–2013: A critical review of methods. *J. Public Health* **2015**, *38*, 806–814. [[CrossRef](#)] [[PubMed](#)]
- Braubach, M.; Ferrand, A. Energy efficiency, housing, equity and health. *Int. J. Public Health* **2013**, *58*, 331–332. [[CrossRef](#)] [[PubMed](#)]
- Rosenow, J.; Platt, R.; Flanagan, B. Fuel poverty and energy efficiency obligations—A critical assessment of the supplier obligation in the UK. *Energy Policy* **2013**, *62*, 1194–1203. [[CrossRef](#)]
- Ambrose, A.R. Improving energy efficiency in private rented housing: Why don't landlords act? *Indoor Built. Environ.* **2015**, *24*, 913–924. [[CrossRef](#)]
- Love, J.; Cooper, A.C. From social and technical to socio-technical: Designing integrated research on domestic energy use. *Indoor Built. Environ.* **2015**, *24*, 986–998. [[CrossRef](#)]
- Snell, C.; Bevan, M.; Thomson, H. Justice, fuel poverty and disabled people in England. *Energy Res. Soc. Sci.* **2015**, *10*, 123–132. [[CrossRef](#)]
- Scarpellini, S.; Rivera-Torres, P.; Suárez-Perales, I.; Aranda-Usón, A. Analysis of energy poverty intensity from the perspective of the regional administration: Empirical evidence from households in southern Europe. *Energy Policy* **2015**, *86*, 729–738. [[CrossRef](#)]
- Programa de las Naciones Unidas para el Desarrollo. *Pobreza Energética: Análisis de Experiencias Internacionales y Aprendizajes para Chile*; Programa de las Naciones Unidas para el Desarrollo: Santiago, Chile, 2018.

15. Boardman, B. *Fuel Poverty: From Cold Homes to Affordable Warmth*; John Wiley & Sons Ltd.: London, UK, 1991.
16. Thomson, H.; Snell, C.; Liddell, C. Fuel poverty in the European Union: A concept in need of definition? *People Place Policy Online* **2016**, *10*, 5–24. [[CrossRef](#)]
17. Schuessler, R. Energy Poverty Indicators: Conceptual Issues—Part I: The Ten-Percent-Rule and Double Median/Mean Indicators. *SSRN Electron. J.* **2014**, *14*. [[CrossRef](#)]
18. Rademaekers, K.; Yearwood, J.; Ferreira, A.; Pye, S.; Hamilton, I.; Agnolucci, P.; Grover, D.; Karásek, J.; Anisimova, N. *Selecting Indicators to Measure Energy Poverty*; Trinomics: Rotterdam, The Netherlands, 2014.
19. Nussbaumer, P.; Bazilian, M.; Modi, V. Measuring energy poverty: Focusing on what matters. *Renew. Sustain. Energy Rev.* **2012**, *16*, 231–243. [[CrossRef](#)]
20. Narula, R.; Kodiyat, T.P. The Growth of Outward FDI and the Competitiveness of the Underlying Economy: The Case of India. In *UNU-MERIT Working Papers 2013*; UNU-MERIT: Maastricht, The Netherlands, 2013; pp. 1–26. [[CrossRef](#)]
21. European Commission. *European Union Statistics on Income and Living Conditions (EU-SILC)*; Eurostat: Luxembourg, 2014.
22. Miniaci, R.; Scarpa, C.; Valbonesi, P. Fuel Poverty and the Energy Benefits System: The Italian Case. *SSRN Electron. J.* **2014**. [[CrossRef](#)]
23. Legendre, B.; Ricci, O. Measuring fuel poverty in France: Which households are the most fuel vulnerable? *Energy Econ.* **2015**, *49*, 620–628. [[CrossRef](#)]
24. Florio, P.; Teissier, O. Estimation of the Energy Performance Certificate of a housing stock characterised via qualitative variables through a typology-based approach model: A fuel poverty evaluation tool. *Energy Build.* **2015**, *89*, 39–48. [[CrossRef](#)]
25. Fabbri, K. Building and fuel poverty, an index to measure fuel poverty: An Italian case study. *Energy* **2015**, *89*, 244–258. [[CrossRef](#)]
26. Sánchez-Guevara Sánchez, C.; Neila Gonzalez, F.J.; Hernández Aja, A. Towards a fuel poverty definition for Spain. In *Proceedings of the World Sustainable Building Conference 2014*, Barcelona, Spain, 28–30 October 2014; pp. 11–17.
27. Okushima, S. Gauging energy poverty: A multidimensional approach. *Energy* **2017**, *137*, 1159–1166. [[CrossRef](#)]
28. MINVU. Estadísticas históricas. Minist. Vivienda. 2016. Available online: <https://www.minvu.cl/elementos-tecnicos/estadisticas/estadisticas-de-edificacion/> (accessed on 20 January 2001).
29. Corporación de Desarrollo Tecnológico de la Cámara Chilena de la Construcción. *Estudio de Usos Finales y Curva de Oferta de la Conservación de la Energía en el Sector Residencial*; Corporación de Desarrollo Tecnológico de la Cámara Chilena de la Construcción: Las Condes, Chile, 2010.
30. Ministerio de Energía. *Desarrollo de un Marco Conceptual y Metodológico para Abordar la Pobreza Energética en Chile*; Ministerio de Energía: Santiago, Chile, 2017.
31. Pérez-Fargallo, A.; Rubio-Bellido, C.; Pulido-Arcas, J.A.; Trebilcock, M. Development policy in social housing allocation: Fuel poverty potential risk index. *Indoor Built Environ.* **2017**, *26*, 980–998. [[CrossRef](#)]
32. Ministerio de Energía. *Agenda de Energía. Un Desafío País, Progreso para Todos*; Ministerio de Energía: Santiago, Chile, 2014.
33. Ministerio de Energía. *Energía 2050. Política Energética de Chile*; Ministerio de Energía: Santiago, Chile, 2017.
34. García-Ochoa, R. *Pobreza Energética en América Latina*; CEPAL: Santiago, Chile, 2014.
35. Programa de las Naciones Unidas para el Desarrollo. *Documento de Proyecto. Desarrollo de un Marco Conceptual y Metodológico para Abordar la Pobreza Energética en Chile*; Programa de las Naciones Unidas para el Desarrollo: Santiago, Chile, 2016.
36. Pérez-Fargallo, A.; Bienvenido-Huertas, D.; Rubio-Bellido, C.; Trebilcock, M. Energy Poverty Risk Mapping Methodology Considering the User's Thermal Adaptability: The Case of Chile. *Energy Sustain. Dev.* **2020**, *58*, 63–77. [[CrossRef](#)]
37. Pino-Mejías, R.; Pérez-Fargallo, A.; Rubio-Bellido, C.; Pulido-Arcas, J.A. Artificial neural networks and linear regression prediction models for social housing allocation: Fuel Poverty Potential Risk Index. *Energy* **2018**, *164*, 627–641. [[CrossRef](#)]
38. Ministerio de Desarrollo Social de Chile. *CASEN 2013. Evolución y Distribución del Ingreso de los Hogares (2006–2013)*; Ministerio de Desarrollo Social de Chile: Santiago, Chile, 2015.
39. 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, Switzerland, 2007.
40. Attia, S.; Carlucci, S. Impact of different thermal comfort models on zero energy residential buildings in hot climate. *Energy Build.* **2015**, *102*, 117–128. [[CrossRef](#)]
41. Bienvenido-Huertas, D.; Pérez-Fargallo, A.; Alvarado-Amador, R.; Rubio-Bellido, C. Influence of Climate on the Creation of Multilayer Perceptrons to Analyse the Risk of Fuel Poverty. *Energy Build.* **2019**, *198*, 38–60. [[CrossRef](#)]
42. Pérez-Fargallo, A.; Rubio-Bellido, C.; Pulido-Arcas, J.A.; Guevara-García, F.J. Fuel Poverty Potential Risk Index in the context of climate change in Chile. *Energy Policy* **2018**, *113*, 157–170. [[CrossRef](#)]
43. Bourdeau, M.; Zhai, X.Q.; Nefzaoui, E.; Guo, X.; Chatellier, P. Modeling and forecasting building energy consumption: A review of data-driven techniques. *Sustain. Cities Soc.* **2019**, *48*. [[CrossRef](#)]
44. Haykin, S.S. *Neural Networks and Learning Machines*; Pearson: Upper Saddle River, NJ, USA, 2009; Volume 3.
45. Raghu, S.; Sriraam, N. Optimal configuration of multilayer perceptron neural network classifier for recognition of intracranial epileptic seizures. *Expert Syst. Appl.* **2017**, *89*, 205–221. [[CrossRef](#)]
46. Zhou, W.; Jia, J. A learning framework for shape retrieval based on multilayer perceptrons. *Pattern Recognit. Lett.* **2019**, *117*, 119–130. [[CrossRef](#)]
47. Bienvenido-Huertas, D.; Moyano, J.; Rodríguez-Jiménez, C.E.; Marín, D. Applying an artificial neural network to assess thermal transmittance in walls by means of the thermometric method. *Appl. Energy* **2019**, *234*, 1–14. [[CrossRef](#)]

48. Pino-Mejías, R.; Pérez-Fargallo, A.; Rubio-Bellido, C.; Pulido-Arcas, J.A. Comparison of linear regression and artificial neural networks models to predict heating and cooling energy demand, energy consumption and CO₂ emissions. *Energy* **2017**, *118*, 24–36. [[CrossRef](#)]
49. Deb, C.; Lee, S.E.; Santamouris, M. Using artificial neural networks to assess HVAC related energy saving in retrofitted office buildings. *Sol. Energy* **2018**, *163*, 32–44. [[CrossRef](#)]
50. Deb, C.; Eang, L.S.; Yang, J.; Santamouris, M. Forecasting diurnal cooling energy load for institutional buildings using Artificial Neural Networks. *Energy Build.* **2016**, *121*, 284–297. [[CrossRef](#)]
51. Magalhães, S.M.; Leal, V.M.; Horta, I.M. Modelling the relationship between heating energy use and indoor temperatures in residential buildings through Artificial Neural Networks considering occupant behavior. *Energy Build.* **2017**, *151*, 332–343. [[CrossRef](#)]
52. Kljajić, M.; Gvozdenac, D.; Vukmirović, S. Use of Neural Networks for modeling and predicting boiler's operating performance. *Energy* **2012**, *45*, 304–311. [[CrossRef](#)]
53. Kialashaki, A.; Reisel, J.R. Modeling of the energy demand of the residential sector in the United States using regression models and artificial neural networks. *Appl. Energy* **2013**, *108*, 271–280. [[CrossRef](#)]
54. Barron, A.R. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Trans. Inf. Theory* **1993**, *39*, 930–945. [[CrossRef](#)]
55. Cybenko, G. Approximation by superpositions of a sigmoidal function. *Math. Control. Signals Syst.* **1989**, *2*, 303–314. [[CrossRef](#)]
56. Rumelhart, D.E.; Hinton, G.E.; Williams, R.J. Learning representations by back-propagating errors. *Nat. Cell Biol.* **1986**, *323*, 533–536. [[CrossRef](#)]
57. Chambers, L.G.; Fletcher, R. *Practical Methods of Optimization*; John Wiley & Sons: Hoboken, NJ, USA, 1980.
58. Kumar, R.; Aggarwal, R.; Sharma, J. Energy analysis of a building using artificial neural network: A review. *Energy Build.* **2013**, *65*, 352–358. [[CrossRef](#)]
59. Aha, D.W.; Kibler, D.; Albert, M.K. Instance-based learning algorithms. *Mach. Learn.* **1991**, *6*, 37–66. [[CrossRef](#)]
60. Beckel, C.; Sadamori, L.; Staake, T.; Santini, S. Revealing household characteristics from smart meter data. *Energy* **2014**, *78*, 397–410. [[CrossRef](#)]
61. Hu, C.; Jain, G.; Zhang, P.; Schmidt, C.; Gomadam, P.; Gorka, T. Data-driven method based on particle swarm optimization and k-nearest neighbor regression for estimating capacity of lithium-ion battery. *Appl. Energy* **2014**, *129*, 49–55. [[CrossRef](#)]
62. Madeti, S.R.; Singh, S. Modeling of PV system based on experimental data for fault detection using kNN method. *Sol. Energy* **2018**, *173*, 139–151. [[CrossRef](#)]
63. Rodger, J.A. A fuzzy nearest neighbor neural network statistical model for predicting demand for natural gas and energy cost savings in public buildings. *Expert Syst. Appl.* **2014**, *41*, 1813–1829. [[CrossRef](#)]
64. Szczurek, A.; Maciejewska, M.; Pietrucha, T. Occupancy determination based on time series of CO₂ concentration, temperature and relative humidity. *Energy Build.* **2017**, *147*, 142–154. [[CrossRef](#)]
65. Hartigan, J.A.; Wong, M.A. Algorithm AS 136: A K-Means Clustering Algorithm. *J. R. Stat. Soc. Ser. C* **1979**, *28*, 100–108. [[CrossRef](#)]
66. Black, P.E. *DADS: The On-Line Dictionary of Algorithms and Data Structures*; NIST: Gaithersburg, MD, USA, 2020.
67. Ahmed, N.K.; Atiya, A.F.; El Gayar, N.; El-Shishiny, H. An Empirical Comparison of Machine Learning Models for Time Series Forecasting. *Econ. Rev.* **2010**, *29*, 594–621. [[CrossRef](#)]
68. Breiman, L.; Friedman, J.H.; Olshen, R.A.; Stone, C.J. *Classification and Regression Trees*; CRC Press: Boca Raton, FL, USA, 2017.
69. Sun, W. River ice breakup timing prediction through stacking multi-type model trees. *Sci. Total. Environ.* **2018**, *644*, 1190–1200. [[CrossRef](#)]
70. Xu, M.; Watanachaturaporn, P.; Varshney, P.K.; Arora, M.K. Decision tree regression for soft classification of remote sensing data. *Remote. Sens. Environ.* **2005**, *97*, 322–336. [[CrossRef](#)]
71. Tso, G.K.; Yau, K.K. Predicting electricity energy consumption: A comparison of regression analysis, decision tree and neural networks. *Energy* **2007**, *32*, 1761–1768. [[CrossRef](#)]
72. Mousa, W.A.; Lang, W.; Auer, T.; Yousef, W.A. A pattern recognition approach for modeling the air change rates in naturally ventilated buildings from limited steady-state CFD simulations. *Energy Build.* **2017**, *155*, 54–65. [[CrossRef](#)]
73. Williams, K.T.; Gomez, J.D. Predicting future monthly residential energy consumption using building characteristics and climate data: A statistical learning approach. *Energy Build.* **2016**, *128*, 1–11. [[CrossRef](#)]
74. Rodriguez-Galiano, V.; Sanchez-Castillo, M.; Chica-Olmo, M.; Chica-Rivas, M. Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. *Ore Geol. Rev.* **2015**, *71*, 804–818. [[CrossRef](#)]
75. Dudoit, S.; Fridlyand, J.; Speed, T.P. Comparison of discrimination methods for the classification of tumors using gene expression data. *J. Am. Stat. Assoc.* **2002**, *97*, 77–87. [[CrossRef](#)]
76. Lariviere, B.; Vandenpoel, D. Predicting customer retention and profitability by using random forests and regression forests techniques. *Expert Syst. Appl.* **2005**, *29*, 472–484. [[CrossRef](#)]
77. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
78. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [[CrossRef](#)]
79. Dietterich, T.G. An Experimental Comparison of Three Methods for Constructing Ensembles of Decision Trees: Bagging, Boosting, and Randomization. *Mach. Learn.* **2000**, *40*, 139–157. [[CrossRef](#)]
80. Assouline, D.; Mohajeri, N.; Scartezzini, J.-L. Large-scale rooftop solar photovoltaic technical potential estimation using Random Forests. *Appl. Energy* **2018**, *217*, 189–211. [[CrossRef](#)]

81. Li, C.; Tao, Y.; Ao, W.; Yang, S.; Bai, Y. Improving forecasting accuracy of daily enterprise electricity consumption using a random forest based on ensemble empirical mode decomposition. *Energy* **2018**, *165*, 1220–1227. [[CrossRef](#)]
82. Manjarres, D.; Mera, A.; Perea, E.; Lejarazu, A.; Gil-Lopez, S. An energy-efficient predictive control for HVAC systems applied to tertiary buildings based on regression techniques. *Energy Build.* **2017**, *152*, 409–417. [[CrossRef](#)]
83. Wang, Z.; Wang, Y.; Zeng, R.; Srinivasan, R.S.; Ahrentzen, S. Random Forest based hourly building energy prediction. *Energy Build.* **2018**, *171*, 11–25. [[CrossRef](#)]
84. Zhou, Y.; Qiu, G. Random forest for label ranking. *Expert Syst. Appl.* **2018**, *112*, 99–109. [[CrossRef](#)]
85. Quinlan, J.R. Learning with Continuous Classes. In Proceedings of the Australian Joint Conference on Artificial Intelligence, Hobart, Australia, 16–18 November 1992; Volume 92, pp. 343–348.
86. Wang, Y.; Witten, I.H. *Induction of Model Trees for Predicting Continuous Classes*; University of Waikato: Hamilton, New Zealand, 1996.
87. Behnood, A.; Behnood, V.; Gharehveran, M.M.; Alyamac, K.E. Prediction of the compressive strength of normal and high-performance concretes using M5P model tree algorithm. *Constr. Build. Mater.* **2017**, *142*, 199–207. [[CrossRef](#)]
88. Lin, L.; Wang, Q.; Sadek, A.W. A combined M5P tree and hazard-based duration model for predicting urban freeway traffic accident durations. *Accid. Anal. Prev.* **2016**, *91*, 114–126. [[CrossRef](#)]
89. Afsarian, F.; Saber, A.; Pourzangbar, A.; Olabi, A.G.; Khanmohammadi, M.A. Analysis of recycled aggregates effect on energy conservation using M5' model tree algorithm. *Energy* **2018**, *156*, 264–277. [[CrossRef](#)]
90. Kuo, C.-F.J.; Lin, C.-H.; Lee, M.-H. Analyze the energy consumption characteristics and affecting factors of Taiwan's convenience stores-using the big data mining approach. *Energy Build.* **2018**, *168*, 120–136. [[CrossRef](#)]
91. Azofra, D.; Martínez, E.; Jiménez, E.; Blanco, J.; Azofra, F.; Saenz-Díez, J. Comparison of the influence of photovoltaic and wind power on the Spanish electricity prices by means of artificial intelligence techniques. *Renew. Sustain. Energy Rev.* **2015**, *42*, 532–542. [[CrossRef](#)]
92. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [[CrossRef](#)]
93. Scholkopf, B.; Smola, A.J. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*; MIT Press: Cambridge, MA, USA, 2001.
94. Smola, A.J.; Schölkopf, B. A tutorial on support vector regression. *Stat. Comput.* **2004**, *14*, 199–222. [[CrossRef](#)]
95. Crawford, E.B.; Carruthers, J.M. Connection and Coherence between and among European Instruments in the Private International Law of Obligations. *Int. Comp. Law Q.* **2014**, *63*, 1–29. [[CrossRef](#)]
96. Zendeheboudi, A.; Baseer, M.; Saidur, R. Application of support vector machine models for forecasting solar and wind energy resources: A review. *J. Clean. Prod.* **2018**, *199*, 272–285. [[CrossRef](#)]
97. Caicedo, D.; Pandharipande, A. Sensor Data-Driven Lighting Energy Performance Prediction. *IEEE Sens. J.* **2016**, *16*, 6397–6405. [[CrossRef](#)]
98. Liu, J.; Liu, J.; Chen, H.; Yuan, Y.; Li, Z.; Huang, R. Energy diagnosis of variable refrigerant flow (VRF) systems: Data mining technique and statistical quality control approach. *Energy Build.* **2018**, *175*, 148–162. [[CrossRef](#)]
99. Jain, R.K.; Smith, K.M.; Culligan, P.J.; Taylor, J.E. Forecasting energy consumption of multi-family residential buildings using support vector regression: Investigating the impact of temporal and spatial monitoring granularity on performance accuracy. *Appl. Energy* **2014**, *123*, 168–178. [[CrossRef](#)]
100. Niu, D.; Wang, Y.; Wu, D.D. Power load forecasting using support vector machine and ant colony optimization. *Expert Syst. Appl.* **2010**, *37*, 2531–2539. [[CrossRef](#)]
101. Ding, Y.; Zhang, Q.; Yuan, T.; Yang, K. Model input selection for building heating load prediction: A case study for an office building in Tianjin. *Energy Build.* **2018**, *159*, 254–270. [[CrossRef](#)]
102. Zhao, J.; Liu, X. A hybrid method of dynamic cooling and heating load forecasting for office buildings based on artificial intelligence and regression analysis. *Energy Build.* **2018**, *174*, 293–308. [[CrossRef](#)]
103. Zuo, R.; Carranza, E.J.M. Support vector machine: A tool for mapping mineral prospectivity. *Comput. Geosci.* **2011**, *37*, 1967–1975. [[CrossRef](#)]
104. Ministerio de Vivienda y Urbanismo. *DS 47—Ordenanza General de la Ley General de Urbanismo y Construcciones*; Ministerio de Vivienda y Urbanismo: Santiago, Chile, 1992.
105. Ministerio de Desarrollo Social. *Resultados Encuesta Casen 2013*; Ministerio de Desarrollo Social: Santiago, Chile, 2013.
106. Kirschen, R.H.; O'Higgins, E.A.; Lee, R.T. The Royal London Space Planning: An integration of space analysis and treatment planning. *Am. J. Orthod. Dentofac. Orthop.* **2000**, *118*, 448–455. [[CrossRef](#)] [[PubMed](#)]
107. Alimissis, A.; Philippopoulos, K.; Tzani, C.; Deligiorgi, D. Spatial estimation of urban air pollution with the use of artificial neural network models. *Atmospheric Environ.* **2018**, *191*, 205–213. [[CrossRef](#)]