PREDICTION OF THE MAINTENANCE PERFORMANCE COST IN DWELLINGS AND BUILDING SITES LOCATED IN SPAIN USING MULTILAYER PERCEPTRONS

David Bienvenido-Huertas[1], David Marín[1], Daniel Sánchez-García[2], Pedro Fernández-Valderrama[1], Juan Moyano[1]

Universidad de Sevilla,

[1]Departamento de Expresión Gráfica e Ingeniería en la Edificación

[2]Departamento de Construcciones Arquitectónicas II

Received: 3/Jan/2019 - Reviewed: 14/Feb/2019 - Accepted: 18/Feb/2019 - DOI: http://dx.doi.org/10.6036/9061

PREDICCIÓN DEL IMPORTE DE ACTUACIONES DE MANTENIMIENTO EN VIVIENDAS Y SOLARES UBICADOS EN ESPAÑA USANDO PERCEPTRONES MULTICAPA

RESUMEN:

La eficiente gestión de los activos inmobiliarios es una de las actividades con mayor demanda en el sector de la ingeniería de la edificación. La determinación y programación de los trabajos de mantenimiento resulta fundamental para que las entidades financieras establezcan el orden de inversión en función del importe presupuestado. Dado que es un proceso de trabajo lento, resulta necesario su optimización. En este trabajo se desarrollan dos perceptrones multicapa (PM) para la determinación del importe económico en los trabajos de mantenimiento en las dos tipologías de activos inmobiliarios de mayor interés para el sector de la edificación: solares y viviendas. En base al entrenamiento con 76 casos de estudio para solares y 317 para viviendas, se obtuvieron que las configuraciones de PM óptimas fueron las de 6 y 12 nodos, respectivamente, y se determinaron las variables de entrada que más influyen en su comportamiento. Además, los PM presentaron un comportamiento más óptimo con respecto a los de regresión lineal múltiple. Finalmente, se testearon los PM ante 15 casos de estudio nuevos para cada modelo, prediciendo los importes presupuestados para los trabajos de mantenmiento con una desviación inferior al 11% con respecto al valor real en la mayoría de ellos.

Palabras clave: perceptrón multicapa, importe presupuestado, solar, vivienda, activo inmobiliario.

ABSTRACT:

The effective asset management of real estate is an area of great interest to the building engineering sector as a whole. The determination and programming of maintenance tasks is essential to allow finance entities to establish market order according to a given budget. The process works slowly, and some optimization is generally required. In this paper, two multilayer perceptrons (MLPs) are developed to determine the economic cost of maintenance works in the two types of real estate asset of more interest to building sector: building sites and dwellings. After training using 76 case studies for building sites and 317 for dwellings, the optimal MLP configurations are shown to have 6 and 12 nodes respectively, and the input variables that most influenced their behavior are also determined. Furthermore, the MLPs showed more optimal behavior than models using multiple linear regression. Finally, the MLPs were tested for 15 new case studies for each model, predicting the budgeted costs of the associated maintenance works with deviations of less than 11% compared with the actual value in most cases.

Keywords: multilayer perceptron, budgeted cost, building site, dwelling, real estate asset.

1.- INTRODUCTION

The outbreak of the financial crisis (following the bankruptcy of the Lehman Brothers in September 2008) was accompanied by significant changes in the economies of international markets [1], reflected in part in the particular characteristics in each country [2]. In Spain, one of the main factors was the failure of the real estate market due to rapid urban development between 1990 and 2006, together with speculation on the ultimate price of individual dwellings, with annual increases of up to 18% [3], although there are other possible causes, such as the buying and selling policy developed by the national government [4], or the tendency of Spanish citizens to purchase properties, to the detriment of rental values [5]. Consequently, finance entities played an important role during the peak of the real estate boom by financing, on the one hand, most of the land purchases, and on the other hand, much of the construction work carried out by property developers along with the purchasing of homes by citizens. This tendency was different between the autonomous regions of the country. In some of them, such as Andalusia, the building sector grew in importance in its productive process [6].

However, since the beginning of the crisis, the vast amount of activity generated by these entities has increased the number of mortgage defaults and foreclosures due to a lack of liquidity, both of commercial entities and individuals [7,8]. In this regard, there were more than 600 000 foreclosures in Spain [9] between 2008 and 2015. One of the key consequences was an increase in the number of real estate assets owned by banking institutions. Among the wide variety of types of asset are [10]: (i) a seemingly endless run of housing developments, (ii) dwellings, both single-family homes and flats, (iii) business premises, and (iv) building sites. The great variety of asset typologies makes their management something of a challenge.

Nowadays, financial institutions are the most important real estate businesses [10]. In recent years, banking institutions have created specialist companies in real estate asset management, causing nationalized savings account managers to move their assets to the Management Company for Assets Arising from Bank Reorganisation (SAREB) in Spain [3,11], although some external investor groups have in fact acquired most of the existing stock [12]. The aim of these companies is to maximize the value of these assets by acquiring commercial advantage by means of their market placement [13,14], thereby reducing the risk associated with the assets as well as increasing the creditworthiness of the financial entities [15]. In effect, these companies provide the banks with warranties for adequate economic recovery, meaning that asset management is now one of the main activities in the real estate sector. However, for adequate management guaranteeing effective exploitation and promotion in the market, both human and time resources are needed. These resources are used to determine the state of the asset, as well as to establish the activities required to guarantee the adequacy of the conditions for habitability, although the process can in fact take a long time. Technical personnel from specialist companies must visit each asset, inspect its pathologies and record any imperfections, and issue a detailed report on the economic valuation of any maintenance works. One of the most significant aspects of this process is the economic valuation of the maintenance works required, which normally takes a considerable time and influences decision-making by the financial entity on the suitability of the process as a whole. Moreover, the huge variety of asset typologies (e.g. dwellings, building sites, commercial properties, garages, offices, or industrial units) can make their management quite challenging. Thus, techniques that speed up this working process could allow the technicians to work more effectively as well as affording the more rapid exploitation of the existing real estate stock.

In this regard, the use of artificial intelligence has been widely applied in various industries as a system for optimizing the work order and reducing working times [16]. From the various existing algorithms, artificial neural networks (ANNs) is a technique that has shown high reliability and viability [17]. ANNs, which have been the focus of much of the research in this area over the last few years along with expert systems, are computational algorithms that imitate the behavior of brain neurons with the aim of building adaptive information-processing systems, which can be used to predict an efficient response [18]. Their main characteristics are their learning, adaptation and generalization ability, allowing us to estimate unknown situations with respect to a set of training data. In the building engineering sector, ANN has been applied more and more over the last few years through different methods, such as: (i) estimation of the economic value of business premises [19]. The estimations carried out by ANNs obtained an increase of 10% in the correlation between the actual and the estimated values with respect to the traditional hedonic modelling approaches; (ii) prediction of costs during the project phase. In the study by Lesniak and Juszczyk [20], the authors analysed the feasibility of using ANNs to estimate the site overhead costs in the early stages of a project. The estimations of the model obtained a coefficient of determination of 84.13%; (iii) Hola and Schabowicz [21] used ANNs to estimate the execution time cost of earthworks. The model works by using input variables of the machines to be used so that the best machined is selected; (iv) Luo et al. [22] developed an ANN to optimize the selection process of suppliers. The selection model had advantages with respect to other models, such as cluster models; and (v) cost valuation of maintenance of construction equipment [23] to ease the resource planning and the decision-makings. However, there is no research on the feasibility of using ANNs to optimize the performance of real estate assets, or predicting the associated economic cost.

1.1.- AIM OF THE PAPER

This paper aims to develop prediction models using ANN, which could allow us to estimate the cost associated with the maintenance performance of real estate assets, the results of which could be useful for establishing a priority ranking when instigated by finance entities. The suggested ANN models are therefore focused on the results, with the objective of minimizing the management time of the maintenance activities of real estate assets and of determining precisely the cost of the units of work to be carried out.

Because most existing real estate assets are building sites and dwellings, two different models are suggested for each. In total, 393 case studies were compiled (76 building sites and 317 dwellings) for the training and the testing of the suggested models to determine the most adequate configuration. Furthermore, the results were compared with multiple linear regression models, given that these models are a technique widely used in other types of economic prediction, such as the sale price of real estate assets [24].

2.- METHODOLOGY

2.1.- CHARACTERISTICS OF THE ARTIFICIAL NEURAL NETWORKS DEVELOPED

In this study, the multilayer perceptron (MLP) is used as the architecture of an ANN. This specific class of ANN is characterized by dividing its structure into three or more layers: (i) an input layer, which obtains the information from the different variables; (ii) one or several hidden layers, which extract the information required; and (iii) an output layer, which indicates the value estimated by the MLP. In each layer, there are several nodes connected to those of the following layer by means of weighted connections with synaptic weights. The value estimated by the output layer is thus given by the following expression:

$$z_k = \sigma\left(\sum_{j=1}^{M} w_{kj}^{(2)} \sigma\left(\sum_{i=0}^{d} w_{ji}^{(1)} x_i\right) + w_{10}^{(2)} y_0\right) \quad (1)$$

Where z_k is the output of the final layer, σ is the activation function, $w_{kj}^{(2)}$ are the weights of the output layer, $w_{ji}^{(1)}$ are the weights of the hidden layer, x_i are the values of the input layer, $w_{10}^{(2)}$ and y_0 are the weight and the input value of the bias neuron of the hidden layer respectively.

The adjustment of synaptic weights is achieved by MLP training, which consists of applying a learning algorithm to a training dataset. In the present case, the MLP was trained by means of back propagation [25–27], using the Broyden-Fletcher-Goldfarb-Shanno algorithm (BFGS) [28], which belongs to the set of Quasi-Newton methods. The training was undertaken using 10-fold cross validation. For each of the MLP layers, a sigmoidal activation function was used (Eq. (2)). Other parameters that configured the MLP were the learning rate and the momentum, with values of 0.3 and 0.2, respectively.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2)$$

Where x is the weighted sum of the inputs to the neuron.

For this study, we only considered a single architecture of MLP with a hidden layer. The number of neurons in the hidden layer and the training time varied in the design phase until the most adequate MLP model was determined according to the aim of this study. The most appropriate number of neurons, as obtained for the hidden layer, is discussed in Section 3 of this paper. To obtain the best model, the following three statistical parameters were used: (i) the linear correlation coefficient (R^2) (see Eq. (3)); (ii) the mean absolute error (MAE) (see Eq. (4)); and (iii) the root mean square error ($RMSE$) (see Eq. (5)). As a quality indicator, it was established that R^2 should have a value greater than 0.95 and that both *MAE* and *RMSE* should be as low as possible. The determination of the number of nodes in the hidden layer to achieve an optimal yield is also discussed in Section 3.

$$R^2 = 100 \cdot \left(1 - \frac{\sum_{i=1}^{n}(x_i - y_i)^2}{\sum_{i=1}^{n}(x_i - \bar{x})^2}\right) \quad [\%] \tag{3}$$

$$MAE = \frac{\sum_{i=1}^{n}|x_i - y_i|}{n} \tag{4}$$

$$RMSE = \left(\frac{\sum_{i=1}^{n}(x_i - y_i)^2}{n}\right)^{1/2} \tag{5}$$

Where x_i is the actual value of the i-th instance, y_i is the value estimated by the ANN of the i-th instance, $\bar{x}$ is the average of the actual values, and n is the number of instances in the dataset.

2.2.- MODELS SUGGESTED

For each MLP, the main factors that influenced the economic valuation of the maintenance tasks were considered. These factors constituted the input variables from the two models developed: one for building sites (MLP-BS) and another for dwellings (MLP-D) (see Fig. 1). The input variables were selected according to the criteria used by the main companies of real estate asset management to analyse the maintenance works to be carried out. The output response was the budgeted cost of the associated maintenance works.

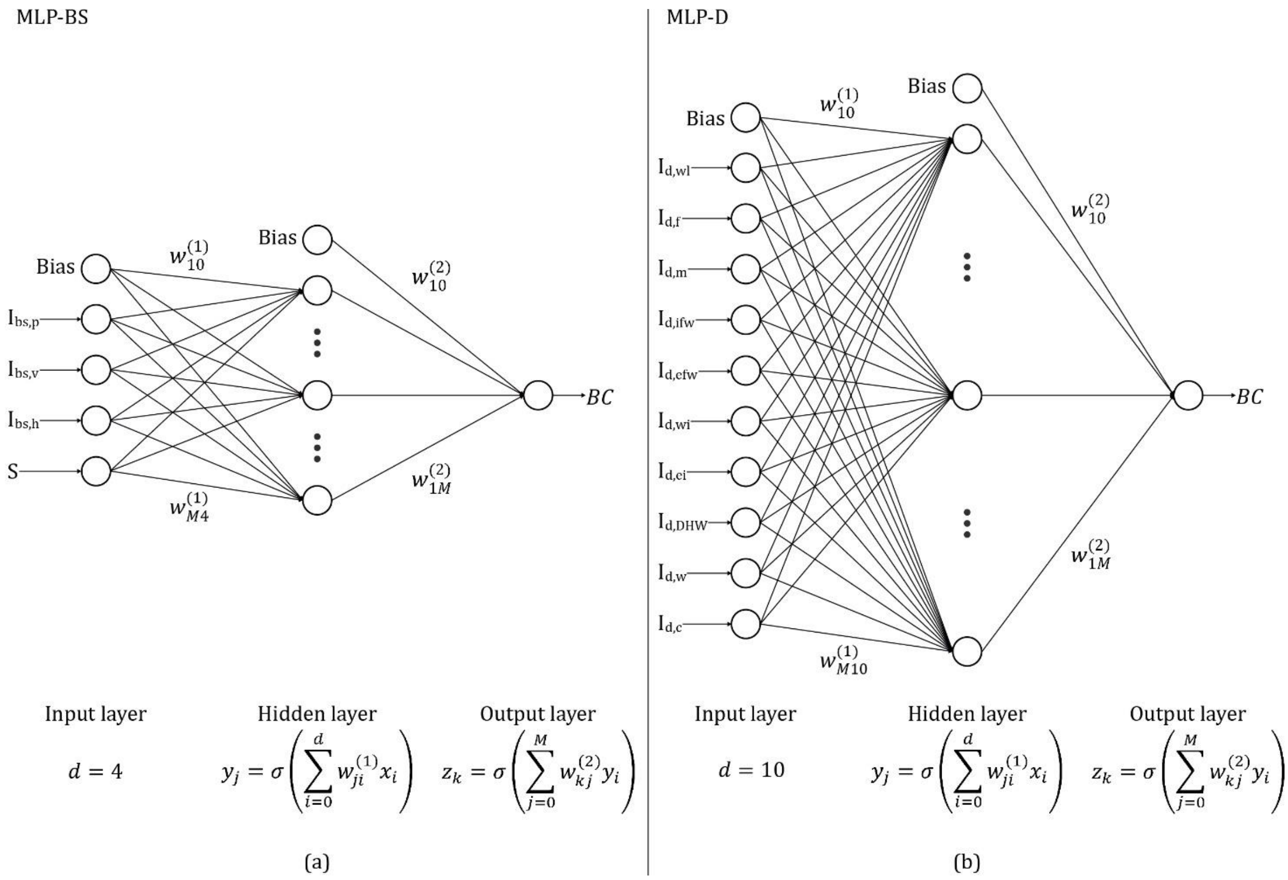


Fig. 1. Structure of the MLPs suggested: (a) building sites, (b) dwellings.

It is important to stress that the discrete values generally used in these tasks to define the input variables were normalized on a scale from 0 to 1. This normalization both made the training process of the MLP easier and hastened the process of field data acquisition by establishing a value on a scale. Values were established according to the level of severity, and the extremes of the distribution of the values corresponded to when the asset was in good condition and when the maintenance was urgent.

In the case of building sites, factors considered when carrying out the economic valuation of performances are as follows: (i) status of the perimeter fence index; (ii) status of existing vegetation index; (iii) level of health index; and (iv) surface area. These factors constituted the input variables (see Fig. 1). Table 1 (a) describes each variable and the associated values.

Inputs	Description
Status of the perimeter fence index ($I_{bs,p}$)	State of the perimeter fence of the building site, in which its safety level and adequate state of maintenance are determined. Aspects such as demolishing a fence in a very bad state, installing a new one and the removal of materials to a waste plant are included in the values associated with this index. The new fence to be installed on a building site is made up of galvanized steel poles, which are fixed by first laying foundations, and it has a simple twisted galvanized steel mesh. Urgent maintenance is considered when the state of the fence implies imminent risk for the safety of third parties, or when the fence does not exist.
Status of existing vegetation index ($I_{bs,v}$)	State of the existing vegetation on the land, in which the safety and health of the existing plants are determined. Works such as weeding and land surface cleaning by means of mechanical resources, cutting down and removing bushes and trees, pulling out stumps and the application of herbicides, and removing the material to a waste plant are included in the values associated with this index.
Level of health index ($I_{bs,h}$)	Cleanliness of the site according to the presence of waste and other factors, including wild animals or parasites such as *Siphonaptera*. Aspects such as the removal of existing waste and specific treatment needed for the site (e.g., exclusion of rats) as well as the removal of the material to a waste plant are included in the values associated with this index.
Surface (S)	Surface area of the site. Cadastral data of the building site can be used.

(a)

Inputs	Description
Index of the state of the walls ($I_{d,wl}$)	State of walls, including aspects such as verticality, the existence of damage or the quality of the finish. Damage related to moisture is not included in this index. Aspects such as the filling of holes, a complete change of finish, and even the demolition of damaged sections of walls and their rebuilding are included in the values associated with this index.
Index of the state of the floors ($I_{d,f}$)	State of floors depending on the presence of loose pieces, which can be dangerous when walking on them, as well as the existence of missing pieces. Re-laying of loose pieces and laying them on surfaces from which they are missing are included in the values associated with this index.
Index of moisture pathologies ($I_{d,m}$)	Existence of pathologies originating in moistures of different kinds (condensation, capillarity or filtration). The removal of the cause of moisture, and the upgrade of the damaged element are included in the values associated with this index.
Index of the inner finishing work ($I_{d,ifw}$)	State of the inner finishing work of the dwelling, in which the correct state of the doors, hardware and other elements are analysed. The valuation of the entrance door is also included. The repair of damage to the inner finishing work and hardware, and even the removal and installation of new elements are included in the values associated with this index.
Index of the exterior finishing work ($I_{d,efw}$)	State of the external finishing work of the dwelling, in which the correct state of windows and balcony doors is analysed. Valuation of the state of any protective elements, such as rails, is also included. The repair of damage to the external finishing work, and even the removal and the installation of new elements, are included in the values associated with this index.
Index of the state of the water installation ($I_{d,wi}$)	State of the installation of the water supply and sewage in the dwelling, analysing its operation. The installation of missing elements or those in a poor state, the repair of any breakdowns, and even the partial or total renovation of the water installation of the dwelling are included in the values associated with this index.
Index of the state of the electrical installation ($I_{d,ei}$)	State of the electrical installation of the dwelling, in which its operation is analysed. The installation of missing elements or the presence of those in a poor state, the repair of any breakdowns, and even the partial or total renovation of the electrical installation of the dwelling are included in the values associated with this index.
Index of the state of the DHW system ($I_{d,DHW}$)	State of the domestic heating water (DHW) system of the dwelling, in which its operation is analysed. The installation of missing elements or presence of those in a poor state, the repair of breakdowns, and even the partial or total renovation of the DHW system of the dwelling are included in the values associated with this index.
Index of the existence of waste ($I_{d,w}$)	Existence of waste in the dwelling, implying a risk to safety and health, such as furniture in a bad conditions. Work related to the management and removal of existing material to a waste plant are included in the values associated with this index.
Index of cleanliness level ($I_{d,c}$)	State of cleanliness and health of the dwelling, including the existence of parasites implying a risk to health, are analysed. Works such as cleaning the dwelling and fumigation are included in the values associated with this index.

(b)

Table 1. Input variables: (a) input variables in the MLP-BS model, and (b) input variables in the MLP-D model.

For dwellings, the main factors considered when determining the state of the real estate asset as well as the maintenance needed are as follows: (i) index of the state of the walls; (ii) index of the state of the floors; (iii) index of moisture pathologies; (iv) index of the state of the interior finishing work; (v) index of the state of the exterior finishing work; (vi) index of the state of the water installation; (vii) index of the state of the electrical installation; (viii) index of the state of the domestic heating water (DHW) system; (ix) index of the existence of waste; and (x) index of the level of cleanliness. These indexes are included as input variables to the MLP-D model (see Fig. 1). As with the building site variables, each of the variables is described more fully in Table 1 (b). Works related to the management and removal of material to a waste plant are included in the associated values. Unlike the model for building sites, the surface area of the dwelling is not considered because all the analysed dwellings have surface areas between 60 and 90 m^2.

The output values predicted by the model determine the cost associated with the maintenance works. In each of the case studies included in the training dataset, the cost associated with the maintenance was determined. For that purpose, a budgeting model of the construction work units was used [29]. In this model, the maintenance to be carried out is divided into a number of work units, and a cost is obtained related to the maintenance, by applying a measure or quantity to an assigned price (Eq. (6)) [29]. The total cost of the maintenance work is obtained from the sum of the direct costs, which are in turn obtained from the sum of the cost of all units (Eq. (7)) [29], and of the indirect costs (Eq. (8)) [29]. Prices were obtained using the database called “Precio Centro de la Construcción en España” [30].

$$U_i = M \cdot P \quad [€] \tag{6}$$

$$DC = \sum_{i=1}^{n} U_i \quad [€] \tag{7}$$

$$BC = DC + IC \quad [€] \tag{8}$$

Where M and P are the measure and price of the work unit U_i, DC is the amount of direct costs, BC is the budgeted cost of maintenance and IC is the amount of indirect costs.

Direct costs are expenses directly attributable to construction work units, and indirect costs are items such as indirect labor, auxiliary means, or technical and administrative personnel. Because the exact determination of indirect costs is complex due to the large number of factors that usually affect these, common practice is to use a percentage of the total of the direct costs (Eq. (9)) [29].

$$\%IC = \frac{IC}{DC} \cdot 100 \quad [\%] \tag{9}$$

Where $\%IC$ is the percentage of indirect costs related to the work.

For this study, a survey of 12 previous cases was undertaken, with the aim of determining the percentage of indirect costs that is usually assigned to this kind of maintenance (see Table 2). The results show that the use of a figure of 3% is adequate to assess the maintenance of building sites, and 20% can be used for maintenance of dwellings. These percentages were used to determine the BC for the dataset obtained.

Performance	DC	IC	$\%IC$
Building site 1	14,917.50	447.53	3.00
Building site 2	23,457.40	912.30	3.89
Building site 3	12,514.83	382.46	3.06
Building site 4	3,794.47	113.83	3.00
Building site 5	4,841.47	128.80	2.66
Building site 6	209,588.58	6,287.66	3.00
Dwelling 1	1,675.03	307.13	18.34
Dwelling 2	1,724.00	316.46	18.36
Dwelling 3	1,583.19	311.42	19.67
Dwelling 4	1,215.37	257.13	21.16
Dwelling 5	1,431.03	307.13	21.46
Dwelling 6	1,241.69	273.92	22.06

Table 2. Percentage of indirect costs for several maintenance works for real estate assets.

2.3.- DATA COLLECTION

To develop the ANN, a robust dataset, which can be used for training and validating the system, is required. For this purpose, 393 different case studies were compiled (76 from building sites and 317 from dwellings), all located in the province of Seville (see Fig. 2). These case studies were selected because of their different characteristics, with the aim of obtaining a greater heterogeneity in the dataset. For instance, dwellings with different ages and state of conservation were selected, and building sites were included in the dataset after assessing the surface and state of maintenance. In each case, the input variables were analysed, the required maintenance was determined, and an economic valuation of these works was obtained.

(a) (b)

Fig. 2. Case studies: (a) photograph of one of the analysed building sites, and (b) photographs of several aspects analysed inside a dwelling.

Despite the case studies are from Seville, the results obtained in this research can be extrapolated to other regions of Spain by calculating the costs with the "Precio Centro de la Construcción en España" database. If it is necessary to use another database as a reference point, the budgeted costs of the training dataset should be recalculated.

Regarding the variety of possible typologies of building sites, building sites were included with some generic conditions in fulfilment of the following requirements: (i) no elements of environmental pollution hazard needing specialist treatment; (ii) no constructive elements at risk of collapse; (iii) no archaeological remains; and (iv) no works in progress, machinery, or earthworks in progress or recently completed.

For the training and validation of the MLPs, 61 case studies were used for building sites and 302 for dwellings. These case studies were radomly selected. The remaining case studies were used for testing the models generated as well as for looking at their behavior in new real estate assets not included in the training process.

3.- RESULTS AND DISCUSSION

As mentioned above, the validity of the MLPs was determined using the statistical parameters R^2, *MAE* and *RMSE*. Table 3 indicates the results obtained for both MLPs, according to the number of neurons in the hidden layer for a training time of 500 (the most optimal parameters were obtained using this training time). In all the configurations, R^2 shows high values, with 97.40% being the lowest value obtained for the configuration of 4 nodes of MLP-D, and 98% or more for most of the architectures suggested. However, great variations occurred in the parameters *MAE* and RMSE, which were useful to determine the optimal architecture of both MLPs suggested. In this sense, the optimal configuration for MLP-BS was the one with 6 nodes, because it had a *MAE* of 270.05 and a *RMSE* of 357.18, with deviations regarding the next nearest value of 6.46% and 4.97%, respectively. On the other hand, the best configuration for MLP-D was the one with 12 neurons with values of *MAE* and *RMSE* of 170.97 and 290.14, with deviations of 2.22% and 7.04%, respectively.

Number of nodes in the hidden layer	R^2	*MAE*	*RMSE*
2	98.16	342.80	432.31
3	98.34	314.05	406.52
4	98.20	318.35	421.94
5	98.60	287.51	374.94
6	98.71	270.05	357.18
7	98.47	303.50	394.67
8	98.47	293.46	392.04
9	98.49	299.17	391.96
10	98.38	317.66	409.69

(a)

Number of nodes in the hidden layer	R^2	*MAE*	*RMSE*
4	97.40	223.55	421.77
5	97.44	222.77	401.91
6	97.87	195.95	362.98
7	98.35	189.68	320.75
8	98.30	183.99	326.04
9	97.82	187.74	367.85
10	98.32	182.60	321.98
11	98.32	190.16	323.17
12	98.64	170.97	290.14
13	98.39	174.74	316.28
14	98.46	175.56	310.58
15	98.34	180.68	320.17
16	98.27	185.21	328.31

(b)

Model	Input variable removed	*RMSE*	Deviation of *RMSE* [%]	Classification
MLP-BS	$I_{bs,p}$	2,128.80	496.00	1
	S	2,098.85	487.62	2
	$I_{bs,v}$	460.26	28.86	3
	$I_{bs,h}$	392.12	9.78	4
MLP-D	$I_{d,ei}$	887.35	205.84	1
	$I_{d,m}$	753.08	159.56	2
	$I_{d,wl}$	573.99	97.83	3
	$I_{d,ifw}$	529.66	82.55	4
	$I_{d,f}$	508.23	75.17	5
	$I_{d,wi}$	443.65	52.91	6
	$I_{d,w}$	414.26	42.78	7
	$I_{d,efw}$	387.41	33.53	8
	$I_{d,DHW}$	342.77	18.14	9
	$I_{d,c}$	306.94	5.79	10

(c)

Table 3. Results obtained in the training phase: (a) behavior of the MLP-BS models for a training time of 500, (b) behavior of the MLP-D models for a training time of 500, and (c) influence of the input variables on quality indicators.

These configurations produced acceptable behavior, although it was essential to determine the importance of the input variables used in both models. To gauge the impact of the input variables considered in the network behavior, the influence of removing each variable was assessed. This training and testing was carried out using the optimal configurations for each MLP (6 nodes for MLP-BS and 12 nodes for MLP-D), and all training runs ended at the same training time as that of the previous phase. To determine the impact of each variable, the variations in the statistical parameter *RMSE*, which reflects the modified network with respect to the original, were studied. In general terms, the different input variables had a considerable influence on the estimated values (see Table 3 (c)). For MLP-BS, variables with the greatest influence were the status of the perimeter fence index ($I_{bs,p}$) and the surface area (S). The first outcome relates to the high cost associated with work to adapt the safeguarding perimeter of the building site, thus it is necessary to remove the existing fence in bad conditions, to manage the waste, and to install the new fence; the surface area also decisively influences the associated direct costs of the three indexes of MLP-BS ($I_{bs,p}$, $I_{bs,v}$ and $I_{bs,h}$). From the four variables considered, the level of health index ($I_{bs,h}$) was the one with the lowest impact on the behavior of the model, with a deviation of 9.78%, because the work units for this variable were of low value in most of the building sites analysed.

For MLP-D, the variable with the highest influence was the index of the state of the electrical installation ($I_{d,ei}$), with a deviation of *RMSE* of 205.84% with respect to that obtained in the model without modifications. This was due to the high cost associated with the maintenance of the electrical installation. Other significant variables were the index of moisture pathologies ($I_{d,m}$) and the index of the state of the walls ($I_{d,wl}$), with percentage deviations of *RMSE* of 159.56% and 97.83%, respectively. The variable with the lowest influence on the network behavior was the index of cleanliness ($I_{d,c}$) due to the low cost associated with the maintenance included in this index.

With the aim of determining the efficiency of the results obtained through the optimal architecture of MLP-BS and MLP-D, the same training dataset was used in order to predict the economic valuation of maintenance works by means of multiple linear regression (MLR) (Eq. (10)), given its widespread use as an analysis technique in other fields of economic estimation, such as the price valuation of real estate asset acquisition [24]. In Eq. (10) Y is the dependent variable, β_0 is the independent term, β_i are the partial regression coefficients, x_i are the predictor variables, and ε is the error. The predictor variables were those used in the MLPs suggested, and the dependent variable was the budgeted cost for maintenance performances (*BC*).

$$Y = \beta_0 + \sum_{i=1}^{n} (\beta_i x_i) + \varepsilon \tag{10}$$

The results indicate that the quality indicators associated with MLPs were better than those of MLR (see Fig. 3). This superior behavior was due to the capacity of MLPs for detecting nonlinear relationships between the variables. In Fig. 3 it is notable how the cloud points for the estimated and predicted values showed a higher correlation in MLP than in MLR, with lower associated errors in each value. In this sense, the values for R^2 were higher in MLP than in MLR, and both *MAE* and *RMSE* had lower values. For the building sites, R^2 of the MLP showed a better adjustment than for MLR, being 5.38% higher, whereas *MAE* and *RMSE* were 54.65% and 55.35% lower, respectively. For dwellings, the R^2 values of the two systems were very similar, with MLP being slightly higher (98.32%) than MLR (97.34%), although the parameters of the error were more optimal in MLP, because MLR increased MAE by 28.99% and RMSE by 25.77%. Thus, the MLPs showed better behavior than MLR for both types of real estate assets.

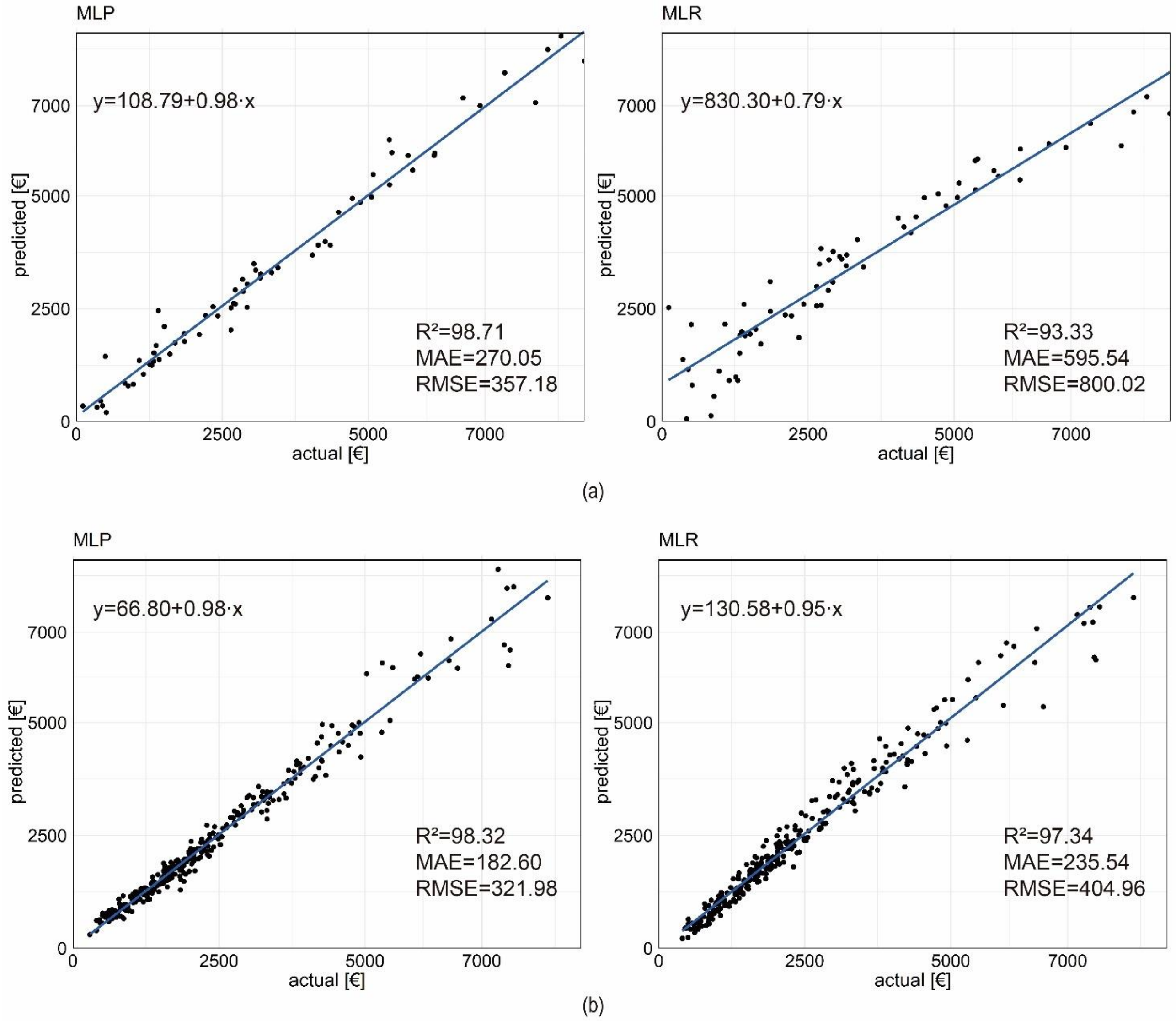
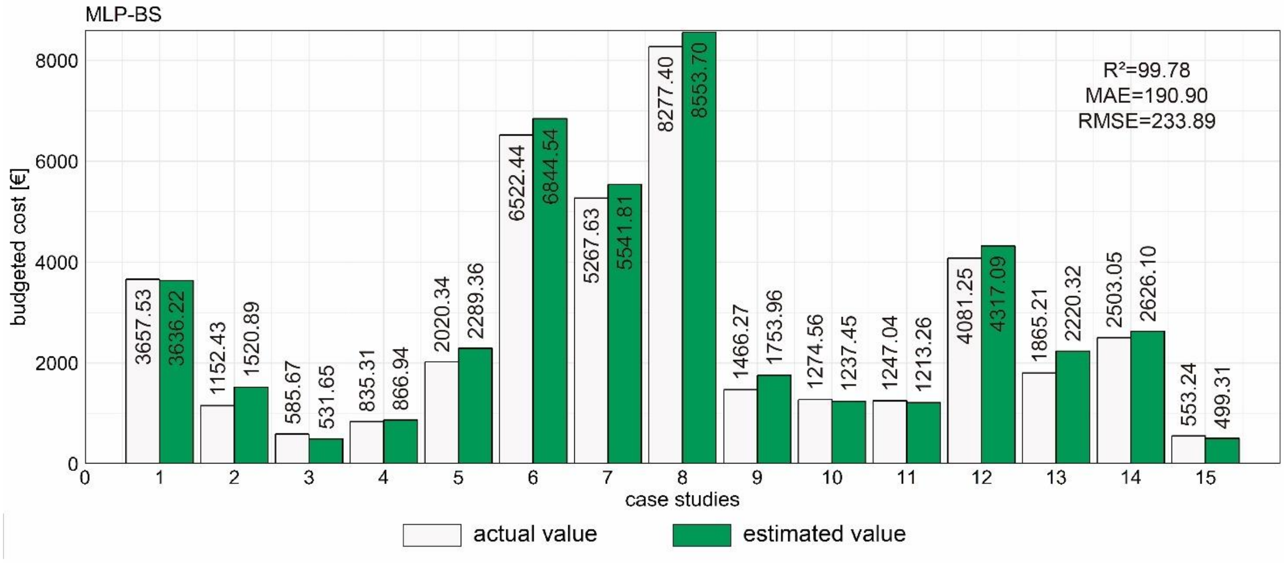
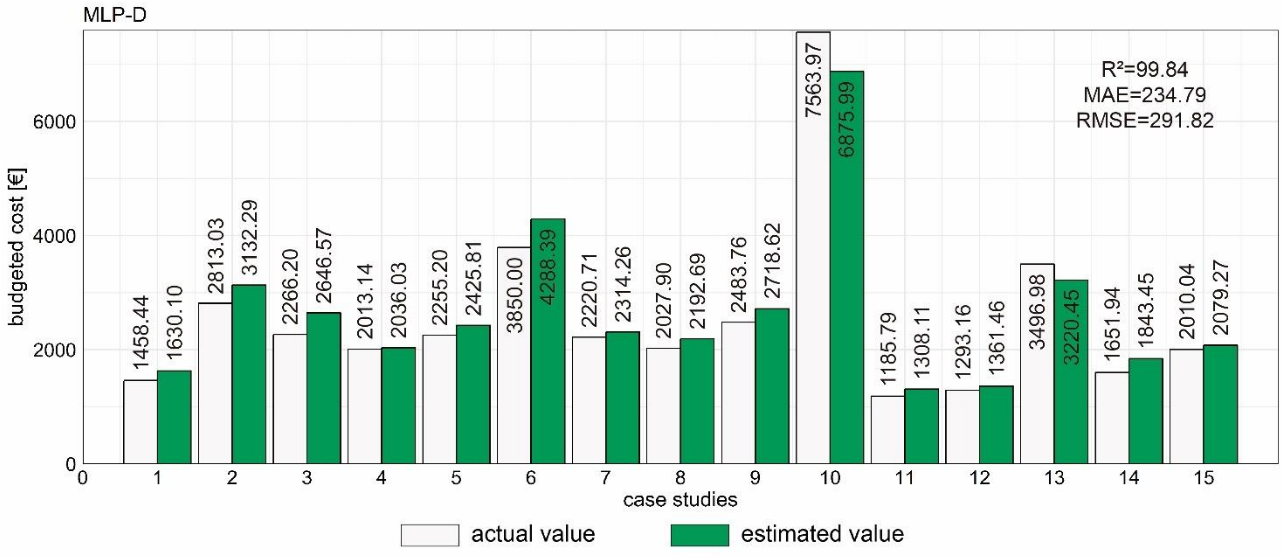


Fig. 3. Comparison between MLP and MLR: (a) comparison of the existing correlation between the values obtained by MLP and MLR for building sites, and (b) comparison of the existing correlation between the values obtained by MLP and MLR for dwellings.

Finally, 15 new case studies were used for building sites and dwellings to test the degree of adjustment and the error of the economic predictions assessed by the MLPs with respect to the actual budgeted cost. The computing time required to carry out the estimations was less than 2 s. Fig. 4 shows graphically the existing differences between the actual and the estimated values for each case study. In both MLP-BS and MLP-D, the results obtained show an adjustment degree greater than 99%, although it is important to highlight that in some case studies, differences were obtained between the actual and the estimated values. Both models showed deviations of less than 11% in most new case studies, although a greater difference was obtained in one case for each MLP: (i) in MLP-BS, a maximum percentage of 24.23% was obtained (see case study 2 in Fig. 4 (a)); and (ii) in MLP-D, a maximum percentage of 14.37% was obtained (see case study 3 in Fig. 4 (b)).



(a)



(b)

Fig. 4. Actual and estimated values for the new data used during the MLP testing of building sites (a) and dwellings (b).

Despite these deviations, the results predicted by the MLPs were close to the actual values, and the robustness of the models could be guaranteed for the new case studies. This was reflected in the reduction of values of *MAE* and *RMSE* by more than 9% in both MLPs with respect to the results of the training and validation phase. Hence, both MLPs could be used to estimate, with high adjustment, the economic valuation of maintenance tasks for real estate assets, and therefore to determine the degree of immediacy of these works as carried out by finance entities, thus optimizing the workflow (see Fig. 5).

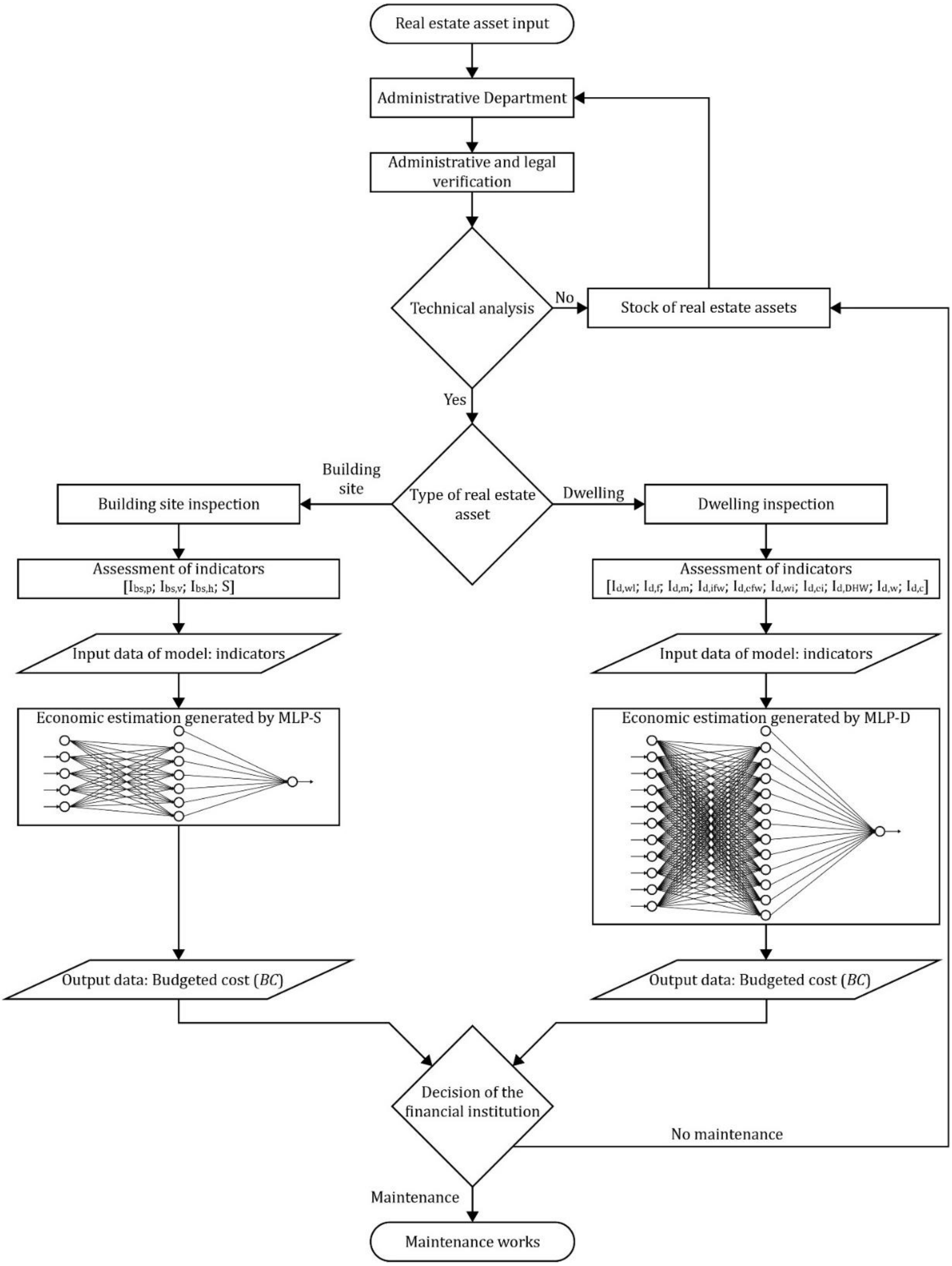


Fig. 5. Workflow process performance.

4.- CONCLUSIONS

A method has been proposed to estimate the maintenance cost for real estate assets owned and managed by financial entities. To this end, two multilayer perceptrons (MLPs) for the two more common typologies of real estate assets were generated: building sites (MLP-BS) and dwellings (MLP-D). The key findings are as follows:

- In the training and validation phase, the architecture of MLP-BS with 6 nodes in the hidden layer and the architecture of MLP-D with 12 nodes were those providing the best yield, according to the values obtained for the statistical parameters considered (R2, MAE and RMSE). In MLP-BS, R2 of 98.71%, MAE of 270.05, and RMSE of 357.18 were obtained, whereas for MLP-D, R2 of 98.64%, MAE of 170.97, and RMSE of 290.14 were obtained.

- From the different input variables in MLP-BS, those that most influenced the network behavior were the status of the perimeter fence index (Ibs,p) and the surface area (S), with deviations with respect to the RMSE of the original MLP of 496% and 487.62%, respectively. Their correct determination is therefore essential to guarantee the prediction of values as precisely as possible. The variable of least influence was the level of health index (Ibs,h), presenting a deviation of 9.78%.

  In MLP-D, variables with highest impact were the index of the state of the electrical installation (Id,ei), the index of moisture pathologies (Id,m), and the index of the state of the walls (Id,wl), with deviations of RMSE of 205.84%, 159.56%, and 97.83%, respectively. The impact of these three variables on the network behavior was due to the high cost associated with work under these indexes, whereas the index of cleanliness level (Id,c) is associated with works of low cost, and was the one with the lowest impact on the network, with a percentage deviation of 5.79%.

- To determine the feasibility of using MLPs, the results obtained by the models generated were compared with those from multiple linear regression (MLR), because this is widely used in other fields of economic estimation. Estimation by MLP gave a better adjustment than that obtained by regression analysis, with a reduction greater than 25% in terms of error parameters.

- MLPs had valid results in the testing of new case studies, with average deviations of 9.95% in MLP-BS and 8.98% in MLP-D, with R2 greater than 99%, and a reduction in MAE and RMSE of more than 9% with respect to the results obtained in the training phase. Only in one case study, considering each model, percentage deviations higher than 11% were obtained: 24.23% in MLP-BS and 14.37% in MLP-D.

The MLPs described here can be used by companies responsible for the management of the real estate stock owned by finance entities, in order to assess, with an adequate degree of adjustment, the cost associated with maintenance tasks. Furthermore, those results can be useful to establish an order of priority in the performances as well as to optimize the workflow. Thus, it is an opportunity to optimize works in this field of the building engineering. The models have been designed to offer good behavior in new case studies, with a low computational resource. Future steps on this research are the comparison of the developed model with other regression typologies, such as support vector machines or regression trees, as well as the determination of new models for other typologies of real estate assets, such as factory or business premises.

REFERENCES

[1] N. Baba, F. Packer. "From turmoil to crisis: Dislocations in the FX swap market before and after the failure of Lehman Brothers". Journal of International Money and Finance. 2009. Vol. 28. p.1350–1374. (doi: http://dx.doi.org/10.1016/j.jimonfin.2009.08.003).
[2] M.B. Aalbers. "European Mortgage Markets Before and After the Financial Crisis". 2012.
[3] S. Fernández Muñoz, L. Collado Cueto. "What has happened in Spain? The real estate bubble, corruption and housing development: A view from the local level". Geoforum. 2017. Vol. 85. p.206–213. (doi: http://dx.doi.org/10.1016/j.geoforum.2017.08.002).
[4] F. Carballo-Cruz. "Causes and consequences of the Spanish economic crisis: Why the recovery is taken so long?". Panoeconomicus. 2011. Vol. 58. p.309–328. (doi: http://dx.doi.org/10.2298/PAN1103309C).

[5] C. Di Feliciantonio, M.B. Aalbers. "The Prehistories of Neoliberal Housing Policies in Italy and Spain and Their Reification in Times of Crisis". Housing Policy Debate. 2017. Vol. 28. p.135–151. (doi: http://dx.doi.org/10.1080/10511482.2016.1276468).
[6] A. Fernández-Tabales, E. Cruz. "Análisis territorial del crecimiento y la crisis del sector de la construcción en España y la Comunidad Autónoma de Andalucía". Eure. 2013, Vol. 39, p. 5–37, (doi: http://dx.doi.org/10.4067/S0250-71612013000100001).
[7] J. Crump, K. Newman, E.S. Belsky, et al. "Cities Destroyed (Again) For Cash: Forum on the U.S. Foreclosure Crisis", Urban Geography. 2008. Vol. 29. p.745–784. (doi: http://dx.doi.org/10.2747/0272-3638.29.8.745).
[8] R. Waldron, D. Redmond. "We're just existing, not living!" Mortgage stress and the concealed costs of coping with crisis". Housing Studies. 2017. Vol. 32. p.584–612. (doi: http://dx.doi.org/10.1080/02673037.2016.1224323).
[9] A. Gutiérrez, S. Vives. "Acumulación de viviendas por parte de los bancos a través de los desahucios: geografía de la desposesión de vivienda en Cataluña". Eure. 2018. Vol. 44. p.5–26.
[10] A.M. Cruz Valdivieso. "Definición de un Modelo de Gestión para los Activos Inmobiliarios mediante Procesos Estructurados". Universidad de Alicante, 2015.
[11] J. Zurita Ussia. "La reforma del sector bancario español: de la correcta valoración de los activos a la recuperación del crédito". Boletín de Estudios Económicos. 2014. Vol. 212. p.263–290.
[12] R. Waldron. "Capitalizing on the State: The political economy of Real Estate Investment Trusts and the 'Resolution' of the crisis". Geoforum. 2018. Vol. 90. p.206–218. (doi: http://dx.doi.org/10.1016/j.geoforum.2018.02.014).
[13] M. Byrne. "Bad banks: the urban implications of asset management companies". Urban Research and Practice. 2015. Vol. 8. p.255–266. (doi: http://dx.doi.org/10.1080/17535069.2015.1050208).
[14] B. Fung, J. George, S. Hohl, G. Ma. "Public Asset Management Companies in East Asia: A comparative study". 2004.
[15] S. Climent Serrano. "La reestructuración del sistema bancario español tras la crisis y la solvencia de las entidades financieras. Consecuencias para las cajas de ahorros". Revista de Contabilidad. 2013. Vol. 16. p.136–146. (doi: http://dx.doi.org/10.1016/j.rcsar.2013.07.003).
[16] P. Quintas, P. Lefrere, G. Jones. "Knowledge Management: a Strategic Agenda". Long Range Planning. 1997. Vol. 30. p.385–391. (doi: http://dx.doi.org/10.1016/S0024-6301(97)90252-1).
[17] F.F. Farshad, J.D. Garber, J.N. Lorde. "Predicting temperature profiles in producing oil wells using artificial neural networks". Engineering Computations. 2000. Vol. 17. p.735–754. (doi: http://dx.doi.org/10.1108/02644400010340651).
[18] S.A. Kalogirou. "Applications of artificial neural-networks for energy systems". Applied Energy. 2000. Vol. 67. p.17–35. (doi: http://dx.doi.org/10.1016/S0306-2619(00)00005-2).
[19] J.M.N. Tabales, F.J.R. Carmona, J.M. Caridad. "Redes neuronales (RN) aplicadas a la valoración de locales comerciales". Informes de la Construcción. 2017. Vol. 69. e179. (doi: http://dx.doi.org/10.3989/ic.15.053).
[20] A. Leśniak, M. Juszczyk. "Prediction of site overhead costs with the use of artificial neural network based model". Archives of Civil and Mechanical Engineering. 2018. Vol. 18. p.973–982. (doi: http://dx.doi.org/10.1016/j.acme.2018.01.014).
[21] B. Hola, K. Schabowicz. "Estimation of earthworks execution time cost by means of artificial neural networks". Automation in Construction. 2010. Vol. 19. p.570–579. (doi: http://dx.doi.org/doi:10.1016/j.autcon.2010.02.004).
[22] X. Luo, C. Wu, D. Rosenberg, D. Barnes. "Supplier selection in agile supply chains: An information-processing model and an illustration". Journal of Purchasing and Supply Management. 2009. Vol. 15. p.249–262. (doi: http://dx.doi.org/10.1016/j.pursup.2009.05.004).
[23] H.L. Yip, H. Fan, Y.H. Chiang. "Predicting the maintenance cost of construction equipment: Comparison between general regression neural network and Box-Jenkins time series models". Automation in Construction. 2014. Vol. 38. p.30–38. (doi: http://dx.doi.org/10.1016/j.autcon.2013.10.024).
[24] J. Aznar Bellver, R. González Mora, F. Guijarro Martínez, A.A. López Perales. "Valoración Inmobiliaria. Métodos y Aplicaciones (España e Iberoamérica)". Polytechnic University of Valencia, Valencia, Spain, 2012.
[25] D.E. Rumelhart, G.E. Hinton, R.J. Williams. "Learning representations by back-propagating errors". Nature. 1986. Vol. 323. p.533–536. (doi: http://dx.doi.org/10.1038/323533a0).
[26] Y.N. Wang. "A neural network adaptive control based on rapid learning method and application". Advance in Molding & Analysis. 1994. Vol. 46. p.27–34.
[27] P. Werbos. "Beyond Regression: New Tools for Prediction and Analysis in the Behavior Science". Harvard University, 1974.
[28] R. Fletcher. "Practical methods of optimization". John Wiley&Sons, Chichester - New York - Brisbane - Toronto, United States, 1980.
[29] A. Ramírez de Arellano. "Presupuestación de obras". Universidad de Sevilla, Sevilla, España, 2010.
[30] COAATGU. Base de Precio Centro de la Construcción en España ", (2017). http://preciocentro.com/ (accessed February 3, 2017).