This article can be cited as J. M. Núñez-Tabaless, F. J. Rey-Carmona and J. M. Caridad y Ocerin, Commercial Properties Prices Appraisal: Alternative Approach Based on Neural Networks, International Journal of Artificial Intelligence, vol. 14, no. 1, pp. 53-70, 2016.

Commercial Properties Prices Appraisal: Alternative Approach Based on Neural Networks

Julia M. Núñez-Tabales[1]; **Francisco J. Rey-Carmona**[2]

and **José Mª Caridad y Ocerin**[3]

[1]Faculty of Economics, University of Cordova,
Puerta Nueva s/n, 14071, Spain
Email: es2nutaj@uco.es

[2]Faculty of Economics, University of Cordova,
Puerta Nueva s/n, 14071, Spain
Email:td1recaf@uco.es

[3]Faculty of Economics, University of Cordova,
Puerta Nueva s/n, 14071, Spain
Email:ccjm@uco.es

ABSTRACT

There are different approaches to obtain the market price of a dwelling. Several agents, such as buyers and sellers, or local or tax authorities need to estimate the value of properties. Econometric hedonic models encounter several theoretical and practical difficulties when applied to the real estate market, such as downward biases in the estimation of hedonic prices, subjective decisions in the measurement process of categorical attributes, frontier problems related to an imperfect information framework, uniequational specification. Many of these are linked to the parametric approach. A lot of papers have been produced in the academic literature for such purposes, but, these are, almost always, oriented to estimate hedonic prices of residential properties, such as houses or apartments. Here these methodologies are used in the field of estimate market price of commercial premises, using Artificial Intelligence (AI) techniques. A case study is developed in Cordova –a medium city in the South of Spain-. Artificial Neural Networks (ANN) are an attractive alternative to the traditional hedonic modelling approaches, as they are better adapted to non-linearities of causal relationships and they also produce smaller valuation errors. It is also possible, from the ANN model, to obtain implicit prices associated to the main attributes that can explain the variability of the market price of commercial properties.

Keywords: *Commercial properties; prices; Artificial Neural Networks (ANN); valuation.*

Mathematics Subject Classification: 68T01.

Computing Classification System: J2; J4.

1. INTRODUCTION

In our days, it is necessary to know the value of the assets involved in economic transactions, though properties valuation has been present throughout most of the history of mankind. Real estate appraisal frequently affects the citizens, since it is present in many ordinary situations as, for example, at the payment of taxes, the purchase of a property or, in a mortgage loan application. From the perspective of the public administration, the valuation of real estate is also of capital importance, as part of the tax receipts are based on the real estate value of the dwellings owned by the tax payers.

Within the real estate market, its main component is the housing sector. As a result, the estimation of the value of houses and apartments has had, in literature concerning the real estate valuation, a preponderant role, while the rest of the properties' studies, such as the corresponding to commercial premises, have been marginal. Nevertheless, the second type of real estate valuations carries out is the commercial property. Accordingly, in this paper, these commercial premises have been chosen to be the object of study, and for the following two reasons: (i) the lack of empirical studies existing around this type of property - despite their importance in the structure and in the configuration of the city- (ii) the fact that this type of property is the most frequently valued, after the residential dwellings.

Today, the real estate valuation is configured as a multidisciplinary activity that combines different techniques and methodologies of study in order to establish the precise value of a real estate product. There have been several authors who have attempted a classification of the diversity of existing methods for the valuation of a property. Among the various proposals, one well known is due to Pagourtzi et al. (2003), as can be deduced from the literature about this subject -see, for example, Selim (2009) and Kusan et al. (2010). Pagourtzi classifies the valuation methods into two categories: traditional methods and advanced methods. The first basically matches the technical valuation methods, although it allows to include any classical mathematical techniques such as regression analysis between them. Among the advanced methods are included artificial neural networks, the hedonic prices approach, methods of spatial analysis, fuzzy logic and time series Box-Jenkins models or Autoregressive Integrated Moving Average (ARIMA).

Alternatively, Gallego (2008) and Aznar (2012), propose a methodological classification. The first makes a distinction between traditional and automated valuation methods. Within the first group, he includes those primarily based on the criterion of experts, as they are supposed to show a high degree of accuracy, but they are characterized by an excessive level of subjectivity; also, these methods are 'non scalable', and thus show a low productivity, in, for example for a company oriented to obtain values of properties for its customers. These traditional methods are, nowadays, well accepted by individuals, companies, banks, and tax authorities. In a second group are the so-called automated valuation methods, mainly characterized by the use of mathematical techniques for the estimation of the value of a particular property. They use this type of techniques, together in a more systematic procedure, and, thus, are considered more 'scientific'. Of course, they tend to be more objective, and can be applied to multiple properties with a quasi null marginal effort (once the models have been tested). Aznar (2012) performs a classification of the valuation methods based on those included in the international valuation standards, and he includes multi-criteria methods and additional procedures for the assessment of assets and environmental resources. However, the most frequently used in practice, are the methods for correction of its two variants, the ratio of assessment method, the method Beta, the hedonic regression analysis, a financial method based on the actual value of future incomes, and the replacement cost method with the residual value after discounting incomes.

Table 1: Classification of the valuation methods of real estate properties.

GROUP	METHOD
TECHNICAL	• Cost • Compared to the market • Update of incomes • Residual - Static - Dynamic
ADVANCED	• Hedonic prices • Artificial intelligence (ANN, Fuzzy logic, Expert Systems, Genetic algorithms) • Spacial analysisl • K-means clustering and regression trees • Multicriteria decision methods • Time series models

Source: Elaborated by authors

In an attempt to unite the possible classifications, in Table 1 are grouped the various methods of valuation of real estate into two categories, differentiating between technical and advanced methods. As already stated, the contributions in the literature about the assessment of commercial properties are quite limited. In fact, it is possible to cite an application of the hedonic prices method carried out by Humarán (2008) on the main Catalan cities and, on the other hand, an application of spatial analysis to the city of Toledo by Montero and Larraz (2011), using the technique known as krigeage. This paper develops an empirical study of commercial premises located in the city of Cordova (Spain), making use of advanced techniques of assessment (Artificial Intelligence).

2. ARTIFICIAL INTELLIGENCE (AI) IN REAL ESTATE VALUATION

Following Martin and Sanz (2006) AI could be defined as a set of algorithms that aims to imitate human reasoning through a deductive logic or manipulation of symbols. Within AI, it is possible to distinguish two large areas, according to Isasi and Galvan (2004). One is the *symbolic* AI, which deals with the construction of systems that can be defined as "intelligent". The other large area of AI is the *sub-symbolic*. This is the case in which there are not *ad-hoc* designs of systems capable of solving full problems. But, the procedures used are general and adaptable to particular needs; they use the well known techniques of the discipline, but based on generic systems which are adapted and being built up from general tools, that, themselves, can be used in a model specification procedure capable of solving the original problem. In this case the design is bottom - up. The Artificial Neural Networks (ANN) can be included in the *sub-symbolic* area. The ANNs are also linked to the other AI techniques such as Fuzzy Logic, Expert Systems, or Genetic Algorithms.

The first studies on real estate valuation using AI procedures date from the beginning of the 1990s, with works of Borst (1991) in New England; therefore, the application of these techniques in the field of real estate pricing has more than two decades of life. Most of the AI systems applied to this field are ANN. Certain authors carried out a comparison between traditional techniques of hedonic prices (especially - MRA - multiple regression analysis) and artificial intelligence. The vast majority of the works insist on the best properties about the support by the ANN, such as their robustness and ability to detect non-linear relationships between variables. This is the case of Tay and Ho (1992); Do and Grudnitski (1992); Kauko *et al.* (2002); Peterson and Flanagan (2009); while others highlight that the ANN are not necessarily superior to regression models - including those of Allen and Zumwalt (1994), Worzala, Lenk and Silva (1995), Rossini (1997), or Zurada *et al.* (2011). These compare their performance in different scenarios in which they reach different conclusions about the superiority of ANN. For his part, Limsombunchai *et al.* (2004) reached equivalent results applying both methodologies. On the other hand, in Spain the first contributions to this field have more than one decade, pioneering Caridad and Ceular (2001), followed by other notable contributions, as Gallego (2004), García Rubio (2004), Lara (2005) or Landajo *et al.* (2012).

It should be noted that in the case studies analyzed (both nationally and internationally) these focus on dealing with valuation of a particular type of properties: residential housing (either single-family or multi-family), and there is no reference for the application of this methodology in the determination of prices of commercial premises. The type of network used generally is the Multilayer Perceptron (MLP), although some studies are differently oriented, using functions of Radial basis (RBF): García Rubio (2004) and Zurada et al. (2011); or auto-organising maps of Kohonen - Kauko *et al.* (2002) and Kontrimas and Verikas (2011). The first layer of the MLP is the input variables that are determinants of the price of the property. As for the output layer, it will be constituted by a single node or neuron, which is usually the price of the real estate transaction (either total or per square meter), although there are references that 'asking prices' are modelled, as well as valuation prices or rental prices.

3. A CASE STUDY

In this section Artificial Intelligence (AI) procedures are employed to determine the major internal and external factors that determine the price of a specific property, that is commercial premises, located in the city of Cordova (Spain). Then the results will be compared with those obtained by a classical hedonic method.

3.1. Data and specification

The value of a commercial premise can be studied using ANN models, and this is done with a sample of 202 properties in the city of Cordova, a medium size town in the South of Spain. Data have been collected with the collaboration of the main local companies, whose main purpose is the rental of commercial premises. The objective variable is the monthly rental price (before the VAT), and all of them referred to the same period (January 2013). This was a moment when the

economic crisis was still felt in Spain, where the rentals are going up two years later. The sample has been randomly selected to be representative of the commercial property marked in the whole city, and covering all the important areas with a significative economic activity.
The selection of determinant attributes in the assessment was an arduous task; several real estate experts were consulted, as well as real estate agents, managers and directors of leasing companies, and Internet portals. Almost forty variables were considered, and from these, 17 were selected (Table 2). These original exogenous variables could be used in the modelling process, but some were discarded as they did not show enough variability, and in some case, they were difficult to measure in many properties, leading to the final seventeen set, which was cross-examined by the real estate experts consulted, and judged as the most important.

Table 2: Classification of the exogenous variables linked to a commercial property

INTERNAL		EXTERNAL	
BÁSICS	Internal Size Total Size (Real, Secondary and Back-office) Size 'translucent' (Primary and Secondary) Geometric shape	**GENERAL**	Years since renovation
GENERAL	Flooring quality Facade quality		
ECONÓMICS	Market price Use of the premises	**LOCATIÓN**	Zone Quality within zone Corner situation

Most of the exogenous variables are quantitative. However, many have been categorized to be able to be treated statistically: the quality of the flooring and the facade of the building, the location area (which indicates the neighbourhood or shopping area where the property is within the city), the quality of location in its urban area (which measures the commercial attractiveness of its exact location), position or not in a corner, geometrical shape or the use of local.

The original data were complemented with several additional variables in each property that could inform about specific aspects that were potentially important in their price. These new attributes are related to the location of each dwelling. The City Hall Statistical Service, the Chamber of Commerce and the Urban Tax Office provided some data that could be used to estimate the value in each case. Some of them were the following: neighbourhood area, income of the area, zip code, number of commercial activities in each zip code, number of activities and population in each area, and density of population by activity in the zone.

To work with many of the qualitative variables obtained, some indexes were constructed, including information from several attributes, so that their importance could be summarized. These indexes take values ranging between 0 and 1, so they could be treated as scale data and with the aim of homogenization of these new variables. A more favourable situation, related to a property, would push the index towards the upper limit of this interval, while the zero value is linked to absence of positive attributes. Three indexes are elaborated, as shown in Table 3.

The Trajan Artificial Neural Networks 6.0 package was used to estimate the final valuation model. Different types of ANN structures and specifications were compared, and finally, a *Multi*

Layer Perceptron (MLP) was selected as the best tool to use. The network topology needs many trials, with different number of hidden layers, changes in the number of neuron in each of them, and links that define the ANN. Also it is necessary to decide between different activation functions embedded in the package, and also the training algorithm to be employed. In Table 4 the characteristics of the final model are presented.

Table 3: Indexes used for qualitative attributes

ÍNDEX	Included attributes
Conservation index	Flooring and facade
Visual index	Quality of the location within the zone and corner situation
Location index	Area where is located the building and level of income in the area

Table 4: Specifications in the final ANN selected

ARCHITECTURE	5:5-6-1:1
Input layer: exogenous variables	5
Hidden layer: number of neurons	6
Output layer: price	1
Number of Weights	43 (36+7)
Activation function	Linear-Logistic-Logistic
Error function	Error sum of square
Training algorithm	Error back-propagation (BP) and Conjugate gradients (CG), using Levenberg-Marquandt algorithm

The input variables included in the first layer are the exogenous variables considered. These are correlated, so there is redundancy of information between them. It is possible to approach this problem in several ways. The first is to select a subset of variables that are deemed important in the price estimation, and, if statistical tests are available, that result significant. Several alternatives are possible to reduce the dimensionality of the data, as in Sánchez and Caridad (2014): the first principal components of the exogenous set could be used as predictors, or the PLS components, although these procedures are based on linear combinations of the original variables, and we intend to take into account non linearities. In the first approach the selected variables are the size of the premises, the indexes of conservation and location, the age of the building and the interaction of the quality of the location and the situation (or not) as a corner office. Fuzzy logic is suited for pricing on properties on geographical areas, where there are not natural borders between different zones; in this case Valdez et al. (2011) and David et al. (2013) propose several optimization algorithms to estimate models with a limited set of parameters. Also Zävoianu et al (2013) use hybridization of multi-objective evolutionary methods with ANN, to reduce the computational effort.

The topology used is (5:5-6-1:1) with the selected five *inputs* variables in the first layer (surface, age of the building, location index, interaction between the location and the corner situation)[1], six latent variables or neurons in the hidden layer, and the price in the output layer (Figure 1). It is, thus, a model with a single equation (corresponding to the endogenous variable considered, the property price). The number of neurons in the hidden layer was determined after many trials, leading to a 43 weights model, seven of them are thresholds or intercepts.

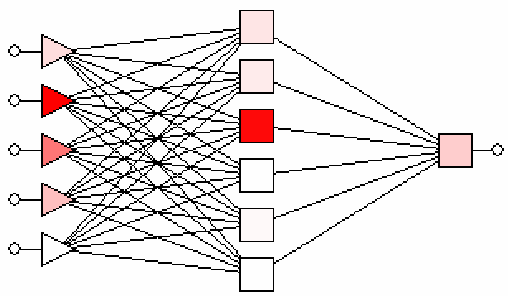

Figure 1. The neural network architecture (5:5-6-1:1)

Sample data were separated randomly in a training set, using 80% of the cases, and a test set with the remaining 20% of the properties. This procedure was carried out several times, selecting different subsets for training of the network, with almost similar results for each batch of cases.

The training of the ANN was carried out using backpropagation and of conjugate gradient, and the Levenberg-Marquandt algorithm, with the following specifications: the network was initialized randomly and maximum number of iterations was fixed at 200 both to the backpropagation and conjugate gradient algorithms; the learning ratio was finally fixed at 0.01 and the moment term at 0.3, after trying several alternative values. Also different topologies were used, with different number of neurons, and also with two hidden layers. The aim is to get a robust model to be used for forecasting prices of future properties, avoiding, thus, overparametrization; finally the chosen network was deemed better for this purpose.

The evolution of the error is plotted along the learning period is plotted in Figure 2.

The final output is thus

$$y_t = g(a'_k) = g(\sum_{j=0}^{6} \omega'_{jk} z_{jt}) = g[\sum_{j=0}^{6} \omega'_{jk} f(\sum_{i=0}^{5} \omega_{ij} x_{it})] + \text{et}, \tag{1}$$

where x_i, i = 1,2,…,5, z_j, j = 1,2,…,6, and y, are the neurons from the layers of the network, and the ω and ω' are the corresponding weights. The exogenous variables are normalized in the first layer. The Levenberg-Marquardt algorithm is designed to minimize the sum-of-squares error function, using a formula that partly assumes that the underlying network function is linear, which is reasonable near the minimum, but not farther away; this method is a compromise between the linear model and a gradient-descent approach; its update formula for the weights is

[1] A similar degree of fit was obtained with another model, including, as explanatory variables in the input layer the size of the property, is age, the location index, the interaction between the length of the outside show-window and depth of the premises, and the number of commercial activities in the area.

$$\Delta\vec{\mathbf{w}}_t = -(\mathbf{Z}'\mathbf{Z} + \lambda\mathbf{I})^{-1}\mathbf{Z}'\vec{\mathbf{e}}_t, \quad (2)$$

where **Z** is the matrix of partial derivatives of the errors $\vec{\mathbf{e}}_t$ with respect to the weights to be estimated. The first term represents the linearized assumption and the second a gradient-descent step, being λ a control parameter related to the relative influence of both terms, and which is decreased by a factor of ten at each iteration when the error is lowered. A stopping rule control the process, until the error drops below a given level or if it fails to improve by a given amount over a predefined number of epochs.

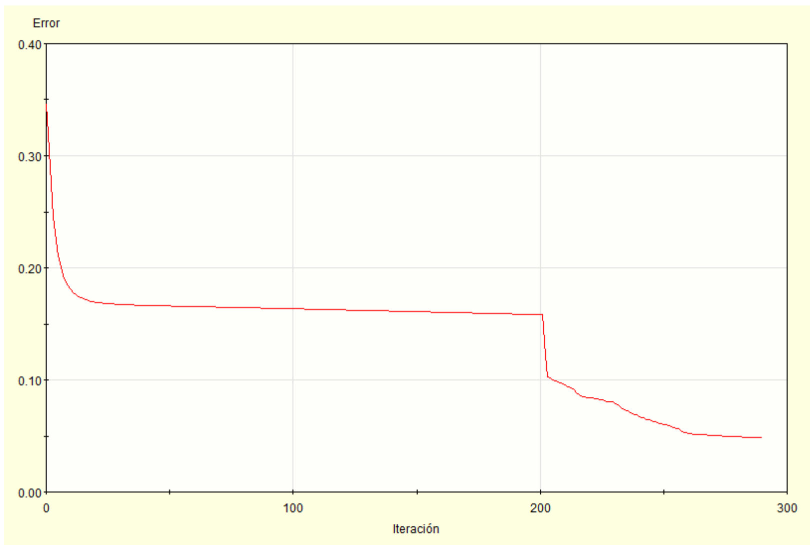

Figure 2. Error evolution during the training phase

3.2. Results

The influence of each input variable on the final price can be analyzed using a sensitivity analysis: the error ratio associated to each input variable is obtained with the error in a truncated model (without the variable tested) and the error in the fully specified model. In Table 5 the ratio obtained for each exogenous variable is included.

As it can be seen, the *size* of the premises is the first (and expected) explanatory variable; its relative error ratio is 2.267; it is, thus, the input with higher degree of predictive power, with a large difference with the rest of the variables. Similar values are associated with the conservation and location indexes, with respective values of 1.557 and 1.537, although the location is also present in the interaction variable between this index and the position of the premises (in a corner place or not in a corner), with relative error of 1.495. The age of the building is related to the perception of the quality of the property, but, this can be influenced by the conservation of the commercial premises also. These variables should be included in the model, and it is not convenient to omit any of them.

Table 5: Sensibility analysis of the exogenous variables in the ANN

INPUT	RATIO	ORDER
SIZE	2,267709	1
CONSERVATION INDEX	1,557319	2

LOCATION INDEX	1,537454	3
LOCATION QUALITY-CORNER EFFECT	1,494639	4
AGE	1,323637	5

In Figure 3 the observed prices (*x*-axis) and the predicted prices (*y*-axis) is presented as a scattergram, and it can be observed the precision of the ARNN estimates, even for the whole range of price considered (this stability of the estimate of different size properties could not be achieved by classical hedonic models). Nowadays, there are three properties whose value is underestimated.

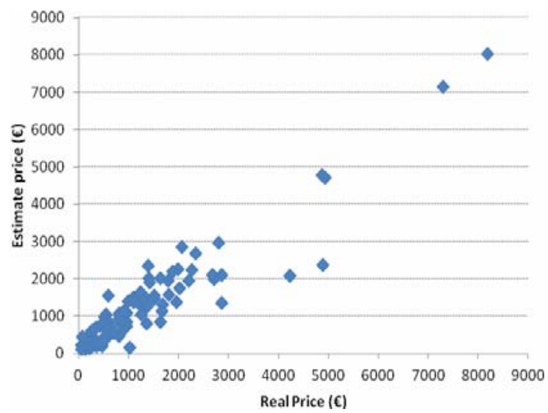

Figure 3. Real vs. estimated prices (Euros) with the ANN

A comparison of the results obtained both with the ANN and with a classical hedonic model (HM) was also made. The results are summarized in Table 6, where the selected ANN produces better estimates than the HM. The neural network has a better degree of fit (measured by its determination coefficient) than the classical model of regression (R^2 of 87.38% against 71.60%), and higher correlation between the price data and their estimates (from 0.8396 to 0.9352), as well as much lower errors.

Table 6: Comparison between the ANN and the HM

	HM	**ANN**
Determination coefficient (R^2)	0.7160	0.8738
Correlation between prices and estimated prices	0.8396	0.9352
Root Mean Square Error (RMSE)	750.64	490.64

To compare the price forecasts obtained with both models, a scattergram of real vs. estimated prices are included in the same graph. The mean square error in both models can be compared, and the ANN produce more precise forecasts that the hedonic models. In Figure 4, the better results of the former are clearly visible. The hedonic modelling is less flexible, and shows more variability.

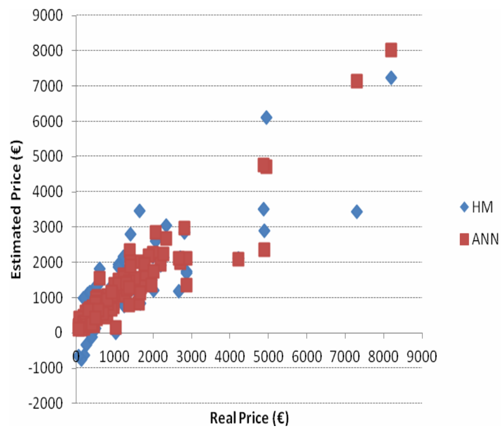

Figure 4. Real vs. estimated prices (Euros) with the ANN vs. HM

The Trajan software produces some graphic output: some response functions to variations of exogenous variables. For example, the response to the size of the premises in square meters and to the localization index (macro-localization within the city). In Figure 5, it is possible to asses that higher prices are linked to larger premises and with better location, as could be expected. The response is more acute with the size of the property, variable that shows a higher sensibility, but the impact on size upon the final price changes with the level of the localization index; for higher values of this index the variability of the surface impact increases, while smaller properties are less prone to be influenced by the situation of the dwelling within the city.

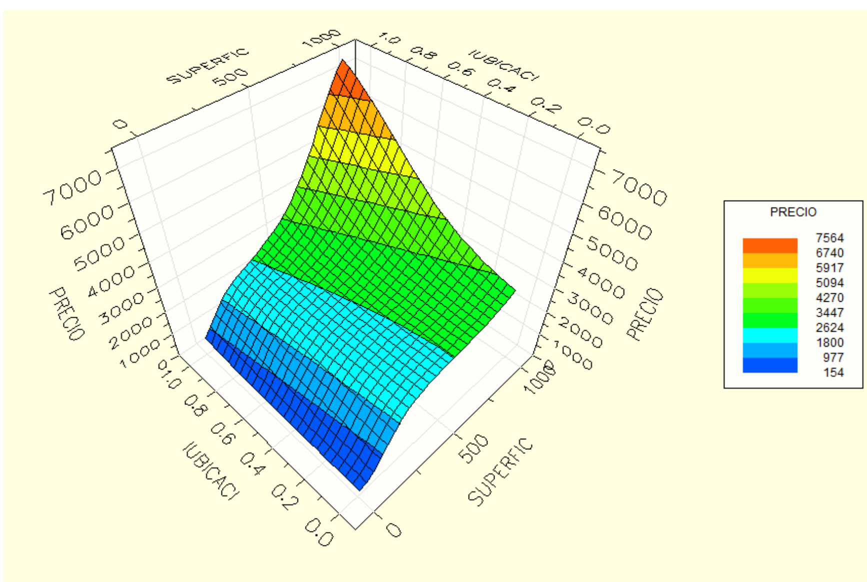

Figure 5. Network response to changes in the inputs size and localization index

Some weight and threshold matrices are produced by Trajan software. In Table 7, the hidden layer is represented by 2.x, being x the x-neuron in this layer.

Table 7: Weights between the input layer and the hidden layer

	2.1	**2.2**	**2.3**	**2.4**	**2.5**	**2.6**
Threshold	1,756416	12,16768	2,53863	1,068854	-1,806674	1,644023
Size	0,2196185	4,626281	0,8645536	-2,424797	-11,84107	0,6873269
Localization index	-0,2759332	2,540661	-0,831218	-0,1748538	-1,342737	0,9653364
Age	-1,368104	3,047515	-1,313245	-0,3981089	-1,256938	2,731032
Interaction between the localization and corner situation	-1,594774	4,066356	-1,237374	-1,948007	-0,4956912	0,906905
Conservation index	-1,974467	2,709747	-1,088052	3,284313	-2,446896	4,350388

In Table 8, the weights linking the hidden and the output layers are presented.

Table 8: Weights between the hidden layer and the output layer

	3.1
Threshold	-0,5769387
2.1	0,8048407
2.2	11,66988
2.3	0,5521522
2.4	-0,8528032
2.5	-6,908442
2.6	-1,497056

The normalization factors are included in Table 9. The final renting price can be obtained using the estimated network.

Table 9: Preprocessing and postprocessing factors for input and output variables

	Origin factor	**Scale factor**
Size	-0,0134048	0,00108717
Localization index	0.000000000	1.000000000
Age	-0,04347826	0,01449275
Interaction between the localization and corner situation	-0,25	0,125
Conservation index	0.000000000	1.000000000
Price	-0,00784658	0,00012306

The implicit prices for the two models are obtained as the marginal influence of each exogenous variable upon the final price. In the case of the ANN the derivative of the estimated price is somewhat cumbersome to obtain, although it can also be deduced using unitary variations in each input. In the case of the HM these values are constant, but in the case of interaction of the exogenous variable with some other in the model, showing, then, a constant rate of change. In the case of ANN, the implicit prices are functions of all the variables in the model, showing the non linearity of the relation between the price of a property and its causal variables. To obtain some visualization of these functions, when obtaining the implicit price for some *x*-variable, it is usual to replace the rest of the exogenous variables with some mean or median value, although it is necessary to take into account the relations existing among many causal variables. For example, holding all other inputs constant at their mean values, several representations are presented. In Figure 6, it can be seen how the increase in the size of the property affects the final price. In the case of the HM this rate of variation is constant (at 319.25€/m^2). From an economic point of view, the limitations of the HM approach are evident, as, the increase in the final renting price is not proportional to the surface of the premises for every type of property. Some nonlinearities are present and are not taken into account by the HM.

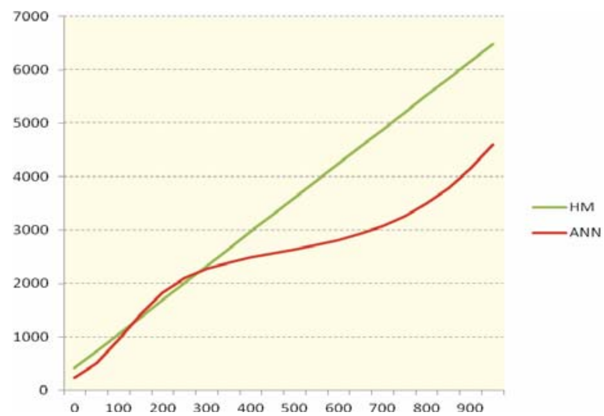

Figure 6. Marginal effect of the surface upon the property estimated price (HM vs ANN)

The ANN reflects that the marginal price that the tenant is willing to pay for a larger office or commercial property presents a peculiar shape with three parts: a concave interval (up to 150 m^2 approximately), a convex part (up to 500 m^2) and, finally, again a concave area for very large surfaces. This two observed cut-points are showing something that can be observed in the market: tenants are willing to pay more for an increase in the local surface up to a certain size of the property (part of the curve with increasing yields); from a certain point, which the increase of the surface produces a proportional rise in the price up to the next cut-point; from the latter, the increase in the marginal prices are less than proportional to the surface. For very larger properties, there is a premium for larger sizes.

Figure 7 compares the response of the ANN and the HM to variations of the variable *age*. Again the hedonic model presents a constant rate €205.53 every ten years. Furthermore, the neural network presents an upward curve of increasing yield at an accelerating rate. It can be seen that prices estimated by the ANN are larger than the corresponding to the HM, to up to 20 years of antiquity approximately, but from that point the situation is reversed, since the network show larger increase in prices. This behaviour can be linked to the development of commercial zones in urban areas, that get consolidated only after a certain interval of time has elapsed.

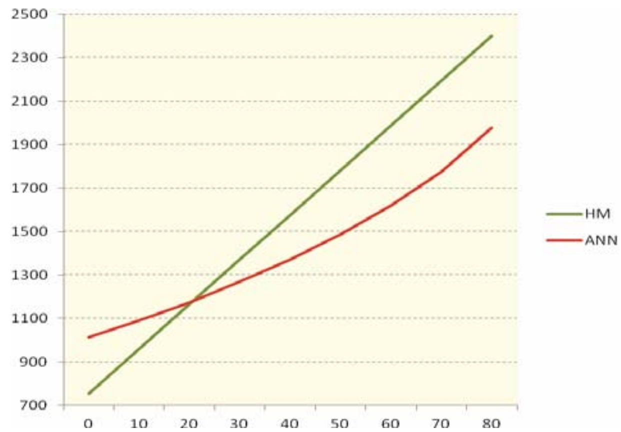

Figure 7. Marginal effect of the age of the building upon the property estimated price

Figure 8 represents the effect of the location index on the estimated price of the property. As noted previously, the index location was obtained weighing the geographical location of the district where the premises are located and the level of income of that particular area; it varies between 0 and 1, and increasing with an improved geographical location and/or a higher level of income of the zone. In regards to the implicit price index of location, the HM reveals that the tenant is willing to pay 137.87€ more for each increment of 0.1 in the index. As the ANN, the effect of location on the price index adopts a nearly linear form, which indicates that the increase in the price of the property is almost the same as the value of the location index increases in 0.1 (increases move in a interval between 92€ and the 86€).

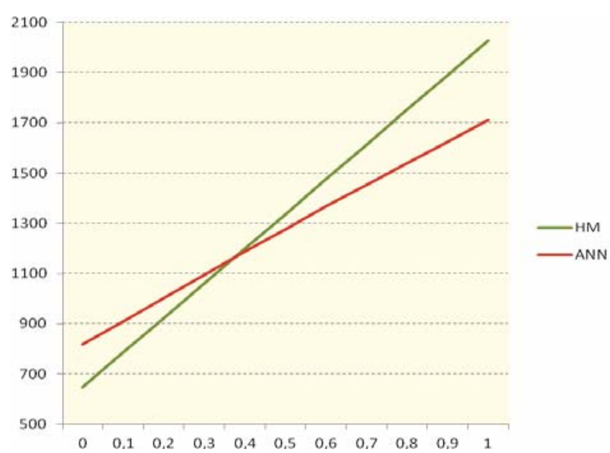

Figure 8. Marginal effect of the location index

The marginal prices determined by the network are larger than those of the hedonic model only up to an index of location of 0.3; from this value, the situation is the opposite. Therefore, once again, the oscillation of the marginal prices reflected by the hedonic model is larger than that corresponding to the network, for properties with increasing values in the location index. Thus, when this index varies from 0 to 1, the marginal prices are up to 1378€ with the HM, but only 893€, when the ANN is used.

The representation of the implicit price related to the interaction between the quality of the location within the area and position or not in a corner (micro-localization) is shown in Figure 9. What is reflected in this case is therefore the product of five levels of quality of the location within a given area, with two levels which reflect whether or not the property is located in a corner position. The estimation of such implicit prices for the HM is such that a unitary increase in the interaction results in constant price increases of 137.52€. For the ANN, the marginal price is a slightly concave curve, reflecting that the increase in this interaction produces an increase in prices more than proportional (constant increases of a unit in the interaction results in increments that increase from 62€ to 133€).

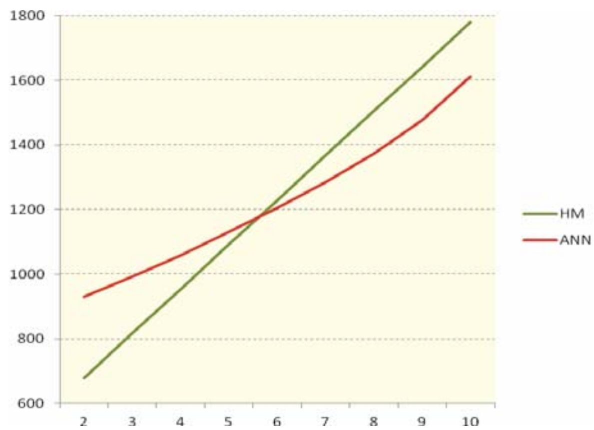

Figure 9. Marginal effects of the interaction between the location and the corner situation

Again, the network estimate higher marginal prices for lower values in the interaction, until a level of interaction of value approximate of five units; from this point, the hedonic estimates exceed those of the network. Therefore, again the network moderates the effect of this variable on the price, since the total variation occurring in the network in the estimation of the price due to this variable is of 681€ against the 1100€ estimate attained by the hedonic model.

In terms of the implicit price for conservation index (built from the state and quality of the facade and floors of the property) it is clear the linear form of the HM reflecting steady increases €88.44 in the estimated price, corresponding to increments of 0.1 in the index (Figure 10). For its part, the ANN marginal price is a convex curve which indicates that the increase in the rate of conservation produces an increase in prices less than proportional (curve of diminishing returns with declines included among the 107€ for lower values of the index to 13€, as it get closer to the value of 1). In summary, up to the index value of 0.5, the hedonic model and the network show similar marginal prices, as clearly shown in the graph, and from this figure the hedonic model continues its proportional climbing, while the network rises slightly to almost stagnate in the highest levels of the index. Level 1 at the index of conservation is valued by the HM €884.4, price which can be excessive, not taking into account the rest of the causal variables, but that is corrected by the estimate made by the network in 594.3€. The relation between the exogenous variables may be the cause, as the improvement of the status and qualities of the façade and the flooring will go in hand with general improvements in other aspects of the building as interior and exterior carpentry, painting, roofing, and the rest of the quality and maintenance characteristics.

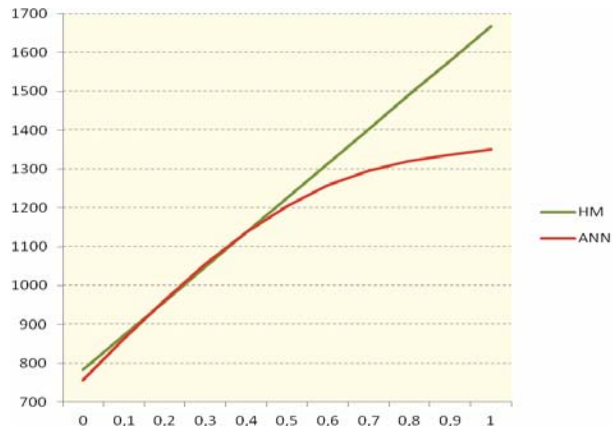

Figure 10. Marginal effects of the conservation index

4. CONCLUSIONS

Artificial Neural Networks provide a more flexible alternative that Hedonic Modelling for price estimation of commercial properties. In this case study, for a medium size city in the South of Spain, neural networks have been applied to obtain evidence of the factors, both external and internal, that determine the price of a specific property type: commercial premises.

The variables that showed a higher incidence in the price of the properties are the following: the surface of the premises, their state of conservation - quantified using an index that reflects the state and quality of the facade and the flooring, the macro-localization within the city - defined through an index that considers the geographical situation of the property (neighbourhood) along with the level of income of the area; another significative variable is the micro-localization -determined by the interaction between the quality of location of the property within the particular area in which it is situated, and the position (being or not in the corner of the building)- and, finally, the age of the building where the property is located. The feature that most influences the determination of the rental price of the commercial premises is therefore a structural factor such as the surface.

However, the presence of the variable referring to the years since the building was erected, as well as its direct relationship with its price, was, in principle, unexpected, since the experts mentioned that this variable was not considered decisive. However, alternative models in which this attribute was not considered presented a sharp decrease in the goodness of fit. The presence of this explanatory variable could be justified with the following arguments: (i) most of the old buildings of the city are located in areas of some tradition or commercial tradition, in which obviously the renting price increases; (ii) the oldest shops, as expected, have experienced serious reforms in the facade, floors, ceilings or carpentry, so their general qualities may outweigh the antiquity of the building, and (iii) in a shop or office, the years of the building do not affect the overall quality of the premises, as in a residential house, because, among other reasons, to access to the premises, you can walk directly in from the street, so it is not possible to appreciate the status and conservation of the main areas of the building.

It is also necessary to highlight the relevance associated to some characteristics related to the same location. The impact of them is twofold: firstly, there has been a very heterogeneous behaviour of the renting prices in the different areas of the city analyzed. Secondly, the quality of the location within the area and its location in corner position has also shown to have a huge impact on the price. ANN methodology compared with hedonic models, shows that the former is a better approach to estimate the formation of the price of a commercial space than the latter. The ANN produces a degree of adjustment far superior to the hedonic methodology, greater correlation between observations and estimates, as well as better indicators of errors.

To obtain the implicit prices corresponding to different inputs or explanatory variables in the HM the specification of the price equation is less flexible. Also you can avoid some classical problems of multicollinearity. In Nuñez et al (2013), some results using the methodology proposed, were used in a different problem: the valuation of flats and apartments. The market shows non linearities for which the ANN are better adapted and in a more consistent way, compared to the hedonic modelling, but, as it proposes models with a larger number of parameters, it is necessary to obtain quite large samples. As a final conclusion, the ANN approach is preferred for the determination of the price of a commercial property, by its superior capacity to recognize the process of formation of prices than the hedonic models.

REFERENCES

Allen, W.C. and Zumwalt, J.K., 1994, *Neural Networks: a word of caution*. Working Paper. Colorado State University.

Aznar Bellver, J., Guijarro Martínez, F., López Perales, A. E., and González Mora, R., 2012, *Valoración inmobiliaria. Métodos y aplicaciones (España e Iberoamérica)*. Valencia (Spain). Ed. Universitat Politècnica de València.

Borst, R.,1991, Artificial neural networks: the next modeling/ calibration technology for the assessment community? *Property Tax Journal, IAAO* **10**(1), 69 – 94.

Caridad, J. M. and Ceular, N., 2001, Un análisis del mercado de la vivienda a través de redes neuronales artificiales. *Estudios de Economía Aplicada* **18**, 67-81.

David, R.-C., Precup, R.-E., Petriu, E.M., Radac, M.-B. and Preitl, S., 2013, Gravitational search algorithm-based design of fuzzy control systems with a reduced parametric sensitivity. *Information Sciences* **247**, 154-173.

Do, A. and Grudnitski, G., 1992, A neural network approach to residential property appraisal. *The Real Estate Appraiser* **58**(3), 38 – 45.

Gallego Mora-Esperanza, J., 2004, La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid. *CT: Catastro* **50**, 51-67.

Gallego Mora-Esperanza, J., 2008, Modelos de valoración automatizada. *CT: Catastro* **62**,7-26.

García Rubio, N., 2004, Desarrollo y aplicación de redes neuronales artificiales al mercado inmobiliario: aplicación a la ciudad de Albacete. PhD Thesis. University of Castilla – La Mancha (Spain).

Humarán, I., Marmolejo, C. and Ruiz, M., 2008, *La formación espacial de los valores comerciales, un análisis para las principales ciudades catalanas*. Communication presented in XXXIV Reunión de Estudios Regionales de la Asociación Española de Ciencia Regional, Baeza (Jaén).

Isasi Viñuela, P. and Galván León, I. M., 2004, *Redes neuronales artificiales: Un enfoque práctico*. Madrid (Spain). Ed. Pearson.

Kauko, T., Hooimaijer, P. and Hakfoort, J., 2002, Capturing housing market segmentation: An alternative approach based on neural network modeling. *Housing Studies* **17**(6), 875 – 894.

Kontrimas, V. and Verikas, A., 2011, The mass appraisal of the real estate by computational intelligence*. Applied Soft Computing* **11**(1), 443-448.

Kusan, H, Aytekin, O. and Özdemir, I., 2010, The use of fuzzy logic in predicting house selling Price. *Expert Systems with Applications* **37**(3), 1808-1813.

Landajo, M., Bilbao, C. and Bilbao, A., 2012, Non parametric neural networks modeling of hedonic prices in the housing market. *Empirical Economics* **42**(3), 989-1009.

Lara Cabeza, J., 2005, Aplicación de las redes neuronales artificiales al campo de la valoración inmobiliaria. *Mapping* **104**, 64-71.

Limsombunchai, V., Gan, C. and Lee., M., 2004, House price prediction: hedonic price model vs. Artificial neural network. *American Journal of Applied Sciences* **3**(1), 193–201.

Martín Del Brío, B. and Sanz Molina, A., 2006, *Redes neuronales y sistemas borrosos.* Madrid: Ed. Ra – Ma.

Montero, J. M. and Larraz, B., 2011, Interpolation methods for geographical data: housing and commercial establishment markets. *Journal of Real Estate Research* **33**(2), 233-244.

Núñez Tabales, J. M., Caridad y Ocerin, J. M., and Rey Carmona, F. J., 2013, Artificial neural networks for predicting real estate price. *Journal of Quantitative Methods for Economics and Business Administration* **15**, 29-44.

Pagourtzi, E., Vassimakopoulos, V.; Hatzichristos, T. and French, N., 2003, Real estate appraisal: A review of valuation methods. *Journal of Property Investment and Finance* **21** (4), 383-401.

Peterson, S. and Flanagan, A., 2009, Neural network hedonic pricing models in mass real estate appraisal. *Journal of Real Estate Research* **31**(2), 147-164.

Rossini, P., 1997, Artificial neural networks versus multiple regression in the valuation of residential property. *Australian Land Economics Review* **3**(1), 1–12.

Sánchez-Rodríguez, M. I. and Caridad y Ocerin, J. M., 2014, Modelling and partial least squares approaches in OODA. *Biometrical Journal* **56** (5), 771-773.

Selim, H., 2009, Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. *Expert Systems with Applications* **36**(2), 2843–2852.

Tay, D.P. and Ho, D.K., 1992, Artificial intelligence and the mass appraisal of residential apartment. *Journal of Property Valuation & Investment* **10**, 525 – 540.

Valdez, F., Melin, P., and Castillo, O., 2011, An improved evolutionary method with fuzzy logic for combining particle swarm optimization and genetic algorithms. *Applied Soft Computing* **11** (2), 2625-2632.

Worzala, E., Lenk, M. and Silva, A.,1995, An exploration of neural networks and its application to real estate valuation. *Journal of Real Estate Research* **10**(2),185 – 201.

Zăvoianu, A. C., Bramerdorfer, G., Lughofer, E., Silber, S., Amrhein, W., and Klement, E. P., 2013, Hybridization of multi-objective evolutionary algorithms and artificial neural networks for optimizing the performance of electrical drives, *Engineering Applications of Artificial Intelligence* **26** (8), 1781-1794.

Zurada, J., Levitan, A. S., and Guan, J., 2011, A comparison of regression and artificial intelligence methods in a mass appraisal context. *Journal of Real Estate Research* ***33***(3), 349-387.