REVISTA DE MÉTODOS CUANTITATIVOS PARA LA ECONOMÍA Y LA EMPRESA (15). Páginas 29–44.
Junio de 2013. ISSN: 1886-516X. D.L: SE-2927-06.
URL: http://www.upo.es/RevMetCuant/art.php?id=67

Artificial Neural Networks for Predicting Real Estate Prices

Núñez Tabales, Julia M.
Faculty of Economics, University of Cordoba (Spain)
Correo electrónico: es2nutaj@uco.es

Caridad y Ocerin, José María
Faculty of Economics, University of Cordoba (Spain)
Correo electrónico: ccjm@uco.es

Rey Carmona, Francisco J.
Faculty of Economics, University of Cordoba (Spain)
Correo electrónico: jnt@uco.es

ABSTRACT

Econometric models, in the estimation of real estate prices, are a useful and realistic approach for buyers and for local and fiscal authorities. From the classical hedonic models to more data driven procedures, based on Artificial Neural Networks (ANN), many papers have appeared in economic literature trying to compare the results attained with both approaches. We insist on the use of ANN, when there is enough statistical information, and will detail some comparisons to hedonic modelling, in a medium size city in the South of Spain, with an extensive set of data spanning over several years, collected before the actual downturn of the market. Exogenous variables include each dwelling's external and internal data (both numerical and qualitative), and data from the building in which it is located and its surroundings. Alternative models are estimated for several time intervals, and enabling the comparison of the effects of the rising prices during the bull market over the last decade.

Keywords: house prices; artificial neural networks (ANN); valuation; econometric modelling.
JEL classification: C45; C51; E37.
MSC2010: 62P20.

Artículo recibido el 16 de noviembre de 2012 y aceptado el 18 de abril de 2013.

Redes neuronales artificiales para la predicción de precios inmobiliarios

RESUMEN

Los modelos econométricos en la valoración de precios inmobiliarios constituyen una herramienta útil tanto para los compradores como para las autoridades locales y fiscales. Desde los modelos hedónicos clásicos hasta los planteamientos actuales a través de redes neuronales artificiales (RNA), han tenido lugar numerosas aportaciones en la literatura económica que tratan de comparar los resultados de ambos métodos. Insistimos en el empleo de RNA en el caso de disponer de suficiente información estadística. En este trabajo se aplica dicha metodología en una ciudad de tamaño medio situada en el sur de España, utilizando una extensa muestra de datos que comprende varios años precedentes a la crisis actual. Las variables utilizadas –tanto cuantitativas como cualitativas– incluyen datos externos e internos de la vivienda, del edificio en el que está localizada, así como de su entorno. Se construyen varios modelos alternativos para distintos intervalos de tiempo, siendo capaces de estimar los efectos de los precios crecientes del mercado alcista durante la década pasada.

Palabras clave: precios de la vivienda; redes neuronales artificiales (RNA); valoración; modelos econométricos.
Clasificación JEL: C45; C51; E37.
MSC2010: 62P20.

1. INTRODUCTION

During last decade, Spain presented a seller's real estate market. There is a strong tradition of home ownership in the country, coupled with fiscal stimuli and low interest rates associated with the euro zone. Prices have continuously gone up, and a growing part of the family's income was devoted to paying mortgages and housing related costs. This trend has come to an abrupt end in 2008, causing turmoil in the country's economy quite dependent on the building industry. In this context, an objective way to assess the prices of properties coming to the market is needed. The case study is presented with data before the downturn, as in the last years prices have been falling and the activity in the markets has slowed to nearly a standstill. Also, the market is distorted by several economic factors that, hopefully, will dwindle in the future: the large number of unsold properties in the hand of building companies and in the banking sector (due to foreclosures and to the bankruptcy of many builders), and the difficulty to access to mortgages and credit. The number of transactions has also fallen restricting the sample sizes available in the following years. The demand for precise valuation models comes from different parties: buyers and sellers of family flats and second homes, agents in the real state sector, investors, realtors and financial institutions, and fiscal authorities at national and municipal levels. Hedonic and Artificial Neural Network (ANN) models could fill the gap with the additional benefit of their usefulness in detecting and measuring price changes, although they must be updated and maintained.

Real estate prices have been studied since the fifties, and two decades later, Rosen (1974) proposed the use of regression models, called 'hedonic models', and applied mainly in urban areas. Hedonic models aim is to estimate the price of a complex good, such as a dwelling, as a function of its characteristics. These characteristics cannot be separated and should be treated as a whole. They include attributes such as the size of the residence, or its internal structure, and variables related to its environment, such as the building properties, its surroundings, and so on. Some authors have also used hedonic models as a tool to obtain price indices. The original methodology was based on classical, usually linear, econometric models. Some time ago, ANN were introduced, to avoid the specification of a functional form relating the exogenous variables to the property's price. A clear advantage was achieved taking into account the non-linearity of the causal relations, and, at the same time, avoiding some econometric problems linked to linear functions, as the multicollinearity among the explanatory variables. Also, as it will be shown, the marginal prices obtained with ANN are more realistic than the classical hedonic prices. The price to be paid to use this more complex methodology is linked both to the mathematical difficulties to obtain the marginal prices with even simple network topologies, and to the need of a fairly large sample to obtain stable parameter estimates.

In Spain there are macroeconomic and sectorial statistics that aim to explain investments in building and in home purchases, but it is unusual to dispose of a model driven methodology to estimate prices. ANN have been applied in several previous studies in the country, and we return to this approach to be used in a dynamic framework, applying it to a case study.

This paper begins with a review of the recent publications on ANN applied to real estate valuation, emphasizing the main works in this field. Then, the methodology is presented with a real data set in a medium sized city in the South of Spain, this case study spans over several years (from 2002 to 2006), to show the feasibility of the methods employed, the stability of the results and the effects of the existing bull market in that period. The models could be employed in the following years to measure the turnaround at the end of the last decade, with falling prices, trend that expands up to the present. The sample includes more than ten thousand transactions of flats and apartments. The procedures can be extended to other type of dwellings, such as industrial and commercial buildings, or to the leasing of properties. The difficulties encountered in the last few years are related to the collapse of the real estate market in most of Spain: the number of transactions has gone down to less than a tenth of any of the years considered, and the credit crunch and falling disposable incomes are affecting dramatically the sale prices lengthening the period to achieve the transactions. Also, the largest segment of the market is absorbed by the banking sector, which is trying to sell its own stock of mortgage foreclosures, contributing to the distortion of the prices. The building industry is almost stalling, and the fiscal authorities are having difficulties with their own valuations.

2. ARTIFICIAL NEURAL NETWORKS: APPLICATION TO REAL ESTATE VALUATION

ANN models were introduced by McCulloch and Pitts (1943) as an alternative to algorithmic programming with some previous work by Karl Lashley in the twenties. In Rumelhart and McClelland (1986), an ANN is presented as a set of operators or neurons, with a small amount of storage capacity, connected numerically one-way by links called axons. Nodes operate with local data supplied through the axons, weighted with some parameters, w_{ij}, linking neurons i and j. A propagation rule establishes how inputs to a neuron are valued and processed. The basic model has an input layer, one or several hidden layers formed by non-observable variables, and an output layer containing the dependent variables. Like any model, the specification is the core task when using ANN: it is necessary to define the network topology (number of hidden layers, and the neurons in each of them), the propagation rule, the transformation of explanatory input variables, and so on. The activation function and the learning rule should also be specified. An excessive number of neurons can originate a lack of forecasting power, due to overparametrization. Computer time to estimate the ANN parameters, learning process, is becoming less relevant with the evolution of the speed in the equipment. There are several learning procedures to estimate the parameters; a widely used technique is the back-propagation method.

An ANN is like a non-linear regression or a multivariate regression model, with non-observable linking variables. Once the topology and the parameters of the network are specified, it can be presented as an ordinary statistical or econometric model. Neural networks are used with different purposes, such as the estimation of models, classification, forecasting, and so on. Here it is used as a modelling tool and as a practical alternative to the well-known econometric hedonic models. Also, it aims at forecasting the value of properties, so it can be used to measure the downturn occurred in the last few years in the real estate market.

Here we use a Multi-Layer Perceptron (MLP) network with one hidden layer, an input layer with the exogenous variables used in estimating prices of a property, and an output layer with the real prices. Several alternative topologies have been tried, with two hidden layers, and different number of neurons and activation functions. The best results have been obtained with the ANN that will be presented.

Table 1. Use of AI methods in real estate price estimation

AUTHORS	DATE	GEOGRAPHICAL AREA
Borst	1991	New England (USA)
Tay and Ho	1992	Singapore
Do and Grudnitski	1992	California
Collins and Evans	1994	UK
Worzala, Lenk and Silva	1995	Colorado
Mc Cluskey *et al.*	1996	Ireland
Rossini	1997	Australia
Bonissone and Cheetham	1997	USA
Haynes and Tan	1998	Australia
Cechin *et al.*	2000	Brasil
Karakozova	2000	Finland
Nguyen and Cripps	2001	Tennessee (USA)
Kauko *et al.*	2002	Finland
Limsombunchai	2004	New Zealand
Liu, Zhang and Wu	2006	China
Selim	2009	Turkey
Hamzaoi and Hernández	2011	Mexico

Source: Elaborated by authors; from Gallego (2004)

Artificial Intelligence (AI) methods started to be applied to real estate prices two decades ago. Some studies are shown in Table 1. In Spain several papers have been presented by Caridad y Ceular (2001), García Rubio (2004), Gallego (2004), and Lara (2005), and some others using data from different urban areas.

3. MATERIAL AND METHODS

The estimation of the selling price of a flat or apartment is the main objective of the ANN approach, as an alternative to hedonic modelling. For this purpose quite a large sample is obtained in a medium size city of the South of Spain, covering its main urban area. Exogenous variables include internal data both about the dwelling and of the building and its location. In 2001 the population is composed of 130,563 properties in town, and over 75% of these are used as main residence; about 14% unoccupied, and less than 10% can be classified as second or holiday residences. This set has augmented during the period analysed to 135,920 properties at the end of 2006.

Price data is recorded by the National Statistics Institute (INE), and by the municipality, for fiscal purposes; however they are not focused on precise valuation of individual properties, but on a spatial classification where average values are considered. The market intermediaries are the main source of reliable data, as they record the real selling price which can differ from the fiscal official figure. The main real estate company in the urban area considered is *Grupo Barin*, and it has supplied 10,124 cases of transactions distributed as follows: in 2002, 772 cases, in 2003, 1,685, in 2004, 1,399, in 2005, 3,380, and in 2006, 2,888. This firm cover the market in the whole city area, with 18 offices scattered through the town. A complementary sample was obtained from several smaller realtors. Table 2 shows the exogenous variables used. The price of the transactions are the real market price (not the declared price, nor the offer price), avoiding, thus duplicities in the data used.

Table 2. Exogenous variables and attributes

<table>
<tr><th colspan="3">INTERNAL</th><th colspan="2">EXTERNAL</th></tr>
<tr><td>BASIC</td><td colspan="2">Area
Bedrooms
Bath
Complimentary baths
Terrace (*)
Communications (*)
Wardrobes (*)
Garage (*)
Storage room(*)
Climatization</td><td>GENERAL</td><td>Building year
Lift (*)
Laundry (*)</td></tr>
<tr><td rowspan="2">GENERAL</td><td>QUALITY</td><td>Floors
Window's type
Interior wooden
Furniture (*)
Kitchen furnit. (*)</td><td rowspan="2">EXTRAS</td><td rowspan="2">Pool (*)
Tennis (*)
Garden (*)</td></tr>
<tr><td>REFORMATIONS</td><td>Reformed (*)</td></tr>
<tr><td>ORIENTATION</td><td colspan="2">Orientation (*)</td><td rowspan="2">LOCATION</td><td rowspan="2">Zone</td></tr>
<tr><td>ECONOMICS</td><td colspan="2">Common expenses
Market price</td></tr>
</table>

Source: Elaborated by authors

Some of these variables are numerical (for example, the number of bedrooms or bathrooms), while categorical information has been represented as binary (*) variables denoting the presence (1) or absence (0) of a facility, such as a garage, a store room, or different important characteristics. With the categorical data, several numerical indices (normalized in the 0-1 range) have been elaborated[1], to represent the general situation of the property, its quality or extensions, or its location (Table 3). The results have been validated and by real estate firms, consulted for this purpose, as a procedure of summarizing this information.

The ANN model for 2006 is presented in the next section. Then the corresponding developments for the previous four years are summarized, and all are compared with the results

[1] Richardson (1973); Saura (1995); Jaén and Molina (1995).

obtained during this period with the classical hedonic models. The model described, for the last year in the sample can be applied to measure the effect on prices with data from 2008 to 2012, while the models estimated for the previous years are used for temporal comparisons and to show the stability of the methodology.

Table 3. Indexes associated to each property

	VARIABLES USED
QUALITY INDEX	Floors, windows type, kitchen furniture and reformations
FACILITIES INDEX	Pool, tennis, garden
BUILDING INDEX	Age, lift, laundry
EXTERNAL DATA INDEX	Orientation, terrace
EXTRAS INDEX	Garage, storage
LOCATION INDEX	Geographical position within the city

Source: Elaborated by authors

The classical hedonic models introduced by Rosen (1974) can be used as an alternative method. The estimation of prices of complex goods composed by several intermingled characteristics has been analysed using regression methods. This approach is useful when there is not enough statistical information, as neural networks usually require quite large samples. Some econometric problems arouse usually in hedonic models, been the more worrisome the multicolinearity always present when many variables linked to each property are introduced. The usual solutions to round the corresponding estimation difficulties, like using principal component regressions, or partial least squares procedures, or others, are an alternative, with a price to be paid related to the statistical properties of the estimators. Heteroskedasticity is also present in these regressions, but the main difficulty in applying hedonic models is related to the specification of the causal relations between some explanatory variables and prices. These relations are non-linear in part of the range of some of these, and there are interactions that are influenced by previously introduced variables in the models. ANN provide an attractive solution, easy to implement, and showing better results that hedonic models for different properties.

Trajan Neural Network software was employed in model estimation, as an alternative to Neuroshell 2. The former is considered more flexible and powerful when working with large set of data. For internal comparisons with classical hedonic models, EViews 7 was used. Some alternative ANN software, included with R language was also employed, and this can be a practical and cheap alternative to elaborate price models.

4. MODEL SPECIFICATION

A *MLP*, with one hidden layer[2] 6:6-6-1:1 was used with the following input variables: size of the property measured in square meters, age of the building, location index, extras index,

[2] The network topology selected, MLP, was decided after several alternatives had been discarded, following also Freeman and Skapura (1993), Haykin (1999), García Rubio (2004). Different number of

community expenses and quality index. They were selected after an identification process with several alternatives. The output layer includes only the transaction price, and there are six neurons in the hidden layer.

Trajan software was used with a linear activation function in the input layer and logistic activation function in the output layer. The error function selected was the residual sum of squares. The training set includes 80% of the sample, corresponding, in the last period, to 952 transactions, leaving the remaining 237 cases as a test set[3].The training method employed is the Backpropagation (BP) procedure, with the following options: a random starting point was used for the parameters, with the default limits, a maximum of 5,000 iterations are allowed, with a learning rate of 0.1, and a moment term equal to 0.1. The error evolution is showed in Figure 1. As the training time can cause an excess in the fitting process, limiting the forecasting power of the ANN, the estimation process has been stopped at iteration 4,624, when the minimum residual sum of squares was almost attained. Tables 4 and 5 present the weights and thresholds between the input and the hidden layer neurons (N1 to N6) and between the hidden and output layer.

Figure 1. Evolution of the residual error during the training process

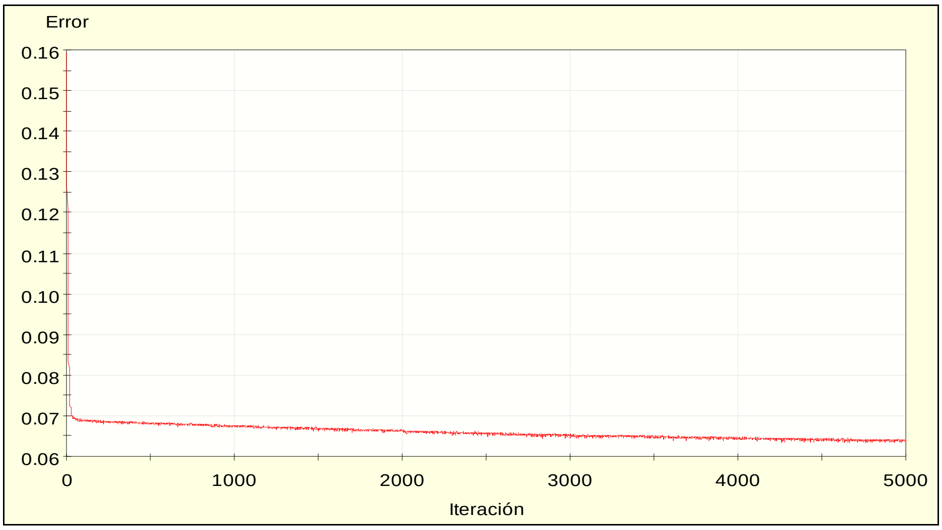


Table 4. Parameters between input and the hidden layer

	N1	N2	N3	N4	N5	N6
Threshold	-0.8637	-5.3193	0.8587	1.3877	-0.8773	-2.9494
Area	1.4581	-2.4698	-3.1902	0.8425	-0.3820	-0.6004
Common expenses	3.4714	-1.6389	1.7720	-0.5584	0.6896	-1.0893
Age	-0.0090	-0.1852	0.8528	0.7126	-1.2623	2.4097
Extras index	2.1949	-1.4684	-0.0846	-0.2235	-2.8127	1.4203
Location index	-1.1403	-0.7805	-5.0060	-1.5345	0.3433	-2.8736
Quality index	0.3414	-0.7721	0.2906	1.4205	-1.0161	-0.0384

hidden layers and neurons were also used; the results selected correspond to the considered best topology, and, also, it was stable for all the yearly subsamples.

[3] All cases had no missing data -1189-.

Table 5.Parameters between the hidden and output layer

	Prices
Threshold	-0.9370
N1	3.9678
N2	-3.9225
N3	-3.7264
N4	1.8743
N5	0.9303
N6	-2.5079

5. RESULTS

Sensitivity analysis allows the evaluation of the influence of each exogenous variable, using its errors ratio (Table 6), obtained as the error in the model without one of the explanatory variables compared to the error in the model including all the variables. As expected, the most important variable is area[4] (1.2953), followed by the location index (1.2009), and the common expenses (1.1577) as shown in Table 6. The other three variables are less important, but if taken out of the model, the final results are less realistic. The determination coefficient (R^2 = 86.05%) and the root mean square error (*RMSE* = €39540.36) are some goodness of fit measures. The mean absolute error (*MAE*) is €28551.34, measured in relative terms to the properties prices is 13.69%. In the scattergram (Figure 2) of the real versus the estimated price, it can be seen that the neural network model produces a good fit all over the spectrum of properties (this can seldom be achieved with classical hedonic models), without distortion at both extremes.

Table 6. Sensitivity analysis

INPUT	RATIO	ORDER
Area	1.2953	1
Common E.	1.1577	3
Age	1.0395	6
Extras index	1.0414	5
Location index	1.2009	2
Quality index	1.0644	4

Figure 2. Real and estimated prices

[4] Different combinations of the exogenous variables were employed; in all cases, as expected, the property's size and location index presented the highest ratios.

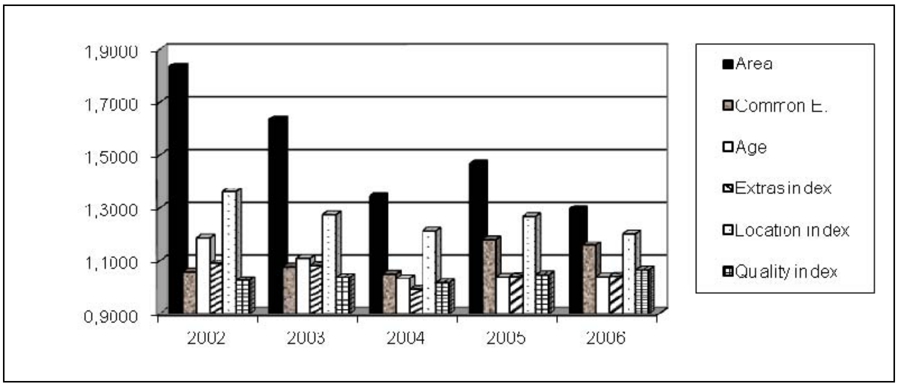
The ANN for the remaining years of the data set produces similar results, albeit with different values of the parameters. The topology of each network is stable over this period, as are the statistical measures of goodness of fit and absolute mean errors. Sample sizes employed are presented in Table 7; in all the models, 80% of the cases are used as training set, with similar tuning parameters.

Table 7. Sample size for the yearly models

	CASES	TRAINING	TEST
2002	470	376	94
2003	914	731	183
2004	791	633	158
2005	1686	1349	337
2006	1189	952	237

The endogenous variable in each network was the transaction price, and the exogenous variables were stable during the whole period. Sensitivity analysis allows the evaluation the input variable influence upon the dwelling's market price. In Figure 3, the variables are ordered with their relative influence, with some temporal variability. In each model the area variable appears in the first place, followed by the location of the property. Then a third variable that can be either the age of the building, the maintenance cost or the quality index. The rest of the variables show less influence.

Figure 3. Comparative sensibility analysis of the inputs



Different goodness of fit statistics is presented in Table 8, corresponding the sample from each year considered; the results are similar during the period considered. During the last years, with the collapse of the real estate market, the number of data available has fallen to such levels as to make the temporal comparisons difficult.

The determination coefficients (R^2) are similar every year, except for 2005. They are always larger than the R^2 of the classical hedonic econometric models (Table 9). The root mean square errors follow a similar pattern with mode variability at the end of the period considered, just before the market turnaround. Mean absolute error shows a growth trend, but the relative errors (*RME*) remain stable, due to the increasing prices of the properties, in what was still a bull market.

Table 8. Goodness of fit measures for the yearly ANN models

	R^2	***RMSE***	***Residual SE***	***MAE***	***RME***
2002	90.21%	21652.25	21634.04	14900.48	13.78%
2003	84.75%	25884.77	25738.99	18999.93	14.75%
2004	90.23%	30983.78	28908.32	24157.36	16.26%
2005	81.12%	32825.44	32755.04	23562.59	13.82%
2006	86.05%	39540.36	39102.13	28551.34	13.69%

RMSE: Root mean square error; Residual SE: Residual standard error; MAE: Mean absolute error; RME: Relative mean error

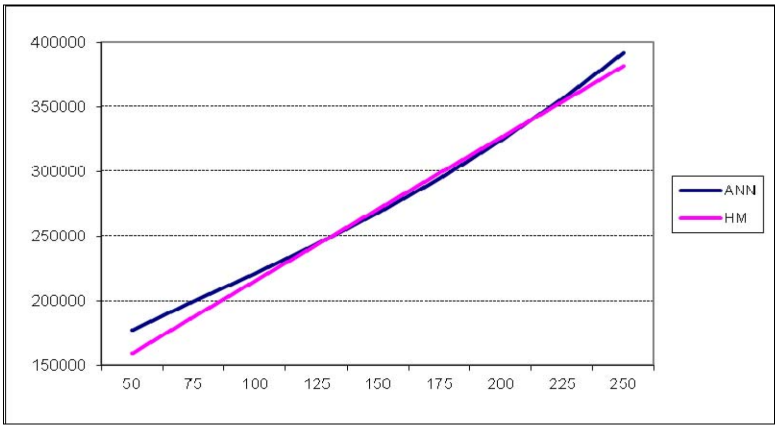
The classical hedonic models are estimated using the same exogenous variables. The estimations have been carried out with EViews 7 software, and it can be appreciated than the ANN approach produces better results. Of course, the neural networks have less degrees of freedom due to the larger number of parameters compared with the classical hedonic models. In this case, with quite large sample sizes, the ANN is to be recommended.

Table 9. Goodness of fit measures for the yearly classical hedonic models

	R^2	***RMSE***	***Residual SE***	***MAE***	***RME***
2002	82.57%	23887.05	23884.92	16118.19	14.91%
2003	80.73%	27951.62	27944.72	19861.95	15.42%
2004	82.27%	29414.24	29401.19	21734.29	14.63%
2005	75.08%	35156.87	35044.47	25124.25	14.74%
2006	77.38%	41645.43	41911.91	30579.18	14.45%

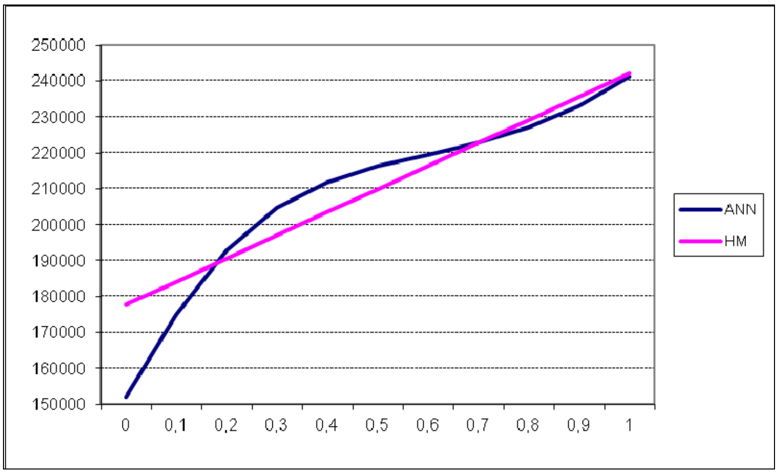
The advantage of the ANN approach over the classical hedonic methodology lies, not only in the more precise estimates of the transaction prices, but also on the results for the marginal prices associated to each of the characteristics of a property. In the classical econometric hedonic models the implicit prices are constant, which is not what a buyer or seller of a property can observe in the market. ANN, as non-linear adaptive models are more flexible. Once the topology is defined, marginal prices of the exogenous variable x_j can be obtained evaluating the function: $\partial Price / \partial x_j$. It is a cumbersome process that, nevertheless, allows the study of the evolution of the hedonic prices attributed to each exogenous variable. For example, to estimate the influence of an additional square metre upon the property price, as will be shown in Figure 4. The conclusions are in line with the results of García Rubio (2004) in a smaller city. The adaptability of the networks to the situations at both extremes of the size range is not possible with classical hedonic models.

Figure 4. Implicit prices (euros) associated to the dwelling surface (squared meters)



The location index takes into account the urban zone and the income level. Upper values correspond to better surroundings and higher income areas. In the hedonic model, an increase in 0.1 units in this index is valued at €6429.73. Being constant this marginal price is not very realistic, so it is necessary to estimate the corresponding implicit price using the neural network, as can be seen in Figure 5, compared to the classical hedonic price.

Figure 5. Implicit prices (Euros) associated to the location index



The influence of the location of the property shows a decreasing rate of increase at lower values of the location index, and then a stabilisation until an almost constant value for the best urban areas is reached.

The common expenses influence the valuation of each dwelling in a non-linear way; an additional euro in these expenses is linked to nearly 1.3 thousands euros in the value of the property. The neural network shows a stabilising influence. When the common expenses increase over a certain level, the implicit price is almost constant (Figure 6). The additional facilities at the property, such as an additional storage room and a parking place, are valued in a different way with the neural network than with the hedonic model, with more realistic values using ANN (Figure 7). When the number of years, since the property was built, is considered, the decrease in the implicit price is a decreasing concave function, with the neural network, while the hedonic price does not show the real depreciation perceived by the buyers, showing a linear decrease over time (Figure 8).

Figure 6. Implicit prices (euros) associated to the common expenses (euros)

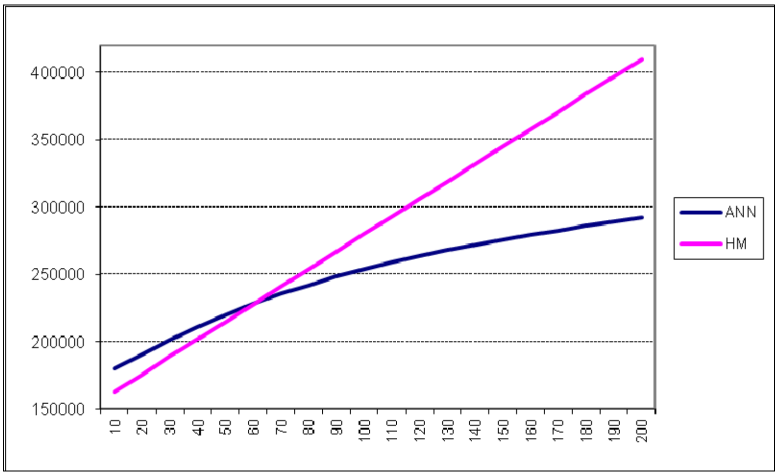


Figure 7. Implicit prices (euros) associated to the annexes (storage and garage)

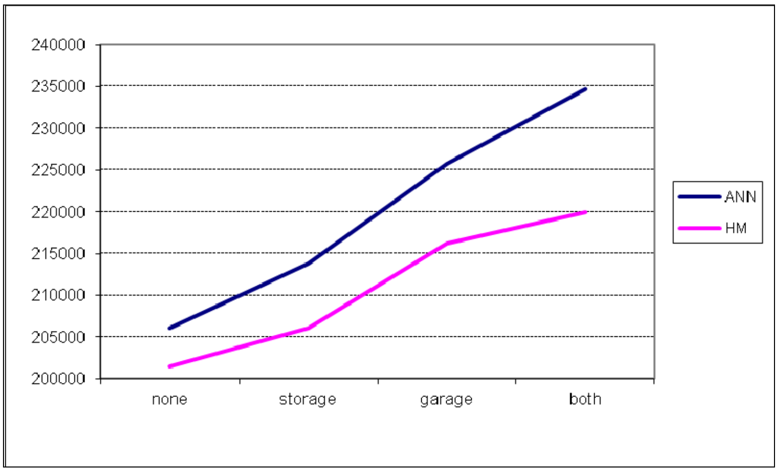
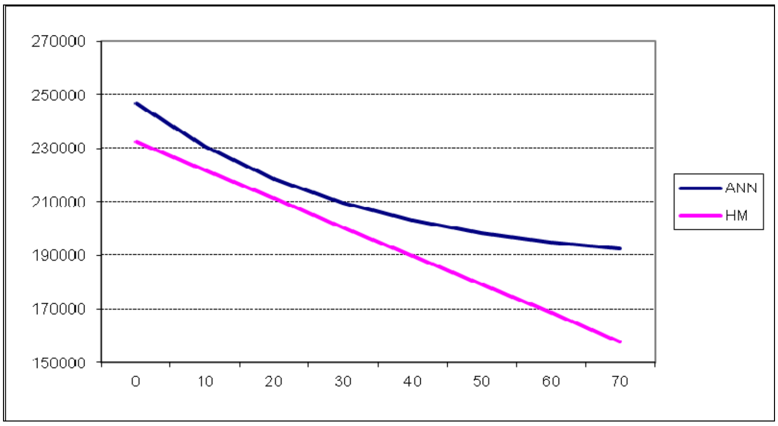


Figure 8. Implicit prices (euros) associated to the age of the building (years)



In general, the ANN approach presents clear advantages over the classical hedonic methodology, in spite of the increases in complexity and the need of large samples to obtain reasonable forecasting capabilities.

6. CONCLUSIONS

Classical hedonic models in real estate valuation have been used for several years as a procedure to estimate prices of such complex goods. The availability of neural network software packages has allowed the use of this alternative methodology, with significantly better results. The use of ANN is more flexible than classical econometric models, when enough sample data are available. Some common problems, in hedonic models, such as non-linearities in the extreme range of prices in real estate market, are well solved by using ANN, even adapting to properties that could be labelled as outliers, although, some authors[5] criticize this 'black box' approach. In our case study with a MLP, estimates fit well with the real market situation, and have even been able to forecast some extreme cases. Neural networks allow a better goodness of fit with the usual measures such as the determination coefficient or with error related statistics *(RMSE*, *MAE*, and so on). The 'dark side' of this methodology is the sampling effort needed. This can be realistically overcome, only with the collaboration of realtors with a broad presence in the markets because their internal databases contain the real data of the properties, where a random sample can be selected, and the uncertainty about the exact identification of each dwelling is eliminated, allowing the inclusion of the particular data of each building and its location.

Most papers include, as explanatory variables, the house size, its location, age and availability of garage[6], and some additional data. Here, a large set of useful characteristics of each property are considered, summarising some of them in several specialized indices. Sensibility analysis confirms these assertions and, due to redundancy in the variables, a subset of explanatory variables are selected, and are confirmed to be stable over the period analyzed. The distance to the city center is useful in monocentric towns. Variability is present even in the same street or building. In clogged Spanish cities, the availability of a garage dramatically enhances the final price. Common expenses are related to the facilities provided by the building, so this variable is highly collinear with other used, and less restrictive in the transactions.

In the validation process, ANN presents more uncertainties, although in hedonic models, multicollinearities introduce a significant difficulty in the interpretation of hedonic prices. In the models proposed, these hedonic prices are obtained as marginal non-linear functions derived once the topology of the network is specified; it is a somewhat cumbersome process to obtain the corresponding derivatives, but the non-linearities of the marginal implicit prices conforms better to reality (for example, a decreasing implicit price with the property's size).

---

[5]Allen and Zumwalt (1994) or Worzala, Lenk, and Silva (1995).

[6]Lara (2005); Do and Grudnitski (1992); García Rubio (2004); Lisombunchai (2004); Worzala *et al.* (1995).

The same network topology is used in different years, with determination coefficients lying between 0.81 and 0.91. The model obtained for 2006 is applied to new data from 2008 and 2010 to forecast the prices of properties. Due to the changing conditions of the market, the estimated prices are 12% to 20% higher than the real transaction data, showing the clear downturn of real estate prices in Spain.

These models could be useful not only to potential buyers/sellers and their agents, but also for foreclosure valuations, not uncommon in the actual bear market, and for fiscal purposes. Also in sectorial studies, to assess the importance of a structural change due to turning points in the cyclical evolution of the building industry.

REFERENCES

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

Bonissone, P.P. and Cheetham, W. (1997): "Financial applications of fuzzy case-based reasoning to residential property valuation". *Fuzz- IEEE*, 1, pp. 37–44.

Borst, R. (1991): "Artificial Neural Networks: The Next Modelling/ Calibration Technology for the Assessment Community?". *Property Tax Journal*, *IAAO*, 10(1), pp. 69–94.

Caridad, J.M. and Ceular, N. (2001): "Un análisis del mercado de la vivienda a través de Sistemas de Redes Neuronales". *Revista de Estudios de Economía Aplicada*, 18, pp. 67–81.

Cechin, A., Souto, A., and Aurelio, M. (2000): "Real estate value at Porto Alegre city using Artificial Neural Networks". *Sixth Brasilian Symposium on Neural Networks Proceedings*, 22-25 November, pp. 237–242.

Collins, A. and Evans, A. (1994): "Artificial Neural Networks: an application to residential valuation in the U.K". *Journal of Property Valuation and Investment*, 11(2), pp. 195–204.

Do, A. and Grudnitski, G. (1992): "A Neural Network Approach to Residential Property Appraisal". *The Real Estate Appraiser*, 58(3), pp. 38–45.

Freeman, J. and Skapura, D.M. (1993): *Redes neuronales algoritmos, aplicaciones y técnicas de programación*. Ed. Wilmington Addison-Wesley.

Gallego, J. (2004): "La inteligencia artificial aplicada a la valoración de inmuebles. Un ejemplo para valorar Madrid". *Revista CT/Catastro*, 50, pp. 51–67.

García Rubio, N. (2004): *Desarrollo y aplicación de redes neuronales artificiales al mercado inmobiliario: aplicación a la ciudad de Albacete*. Tesis Doctoral. Universidad de Castilla-La Mancha (España).

Hamzaoi, Y.E. and Hernández, J.A. (2011): "Application of Artificial Neural Networks to predict the selling Price in the real estate valuation". *10th Mexican International Conference on Artificial Intelligence*, November 26-December 04, pp. 175–181.

Haykin, S. (1999): *Neural networks: A comprehensive foundation*. Ed. Prentice-Hall.

Haynes, J.D. and Tan, C.N.W. (1993): *An Artificial Neuronal Network real estate price predictor*. IEEE Computer Society Press: USA.

Jaén, M. and Molina, A. (1995): *Modelos econométricos de tenencia y demanda de vivienda*. Servicio de Publicaciones de la Universidad de Almería (España).

Karakozova, O.A. (2000): *Comparison between neural network and multiple regression approaches: An application to residential valuation in Finland*. Swedish School of Economics and Business Administration.

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

Lara, J. (2005): "Aplicación de las redes neuronales artificiales al campo de la valoración inmobiliaria". *Mapping*, 104, pp. 64–71.

Lashley, K. (1929): *Brain mechanisms and intelligence*. University of Chicago Press.

Limsombunchai, V., Gan, C., and Lee, M. (2004): "House price prediction: Hedonic Price Model vs. Artificial Neural Network". *American Journal of Applied Sciencies*, 1(3), pp. 193–201.

Liu, J., Zhang, X., and Wu, W. (2006): "Application of fuzzy neural network for real estate prediction". *LNCS*, 3973, pp. 1187–1191.

McCluskey, W., Dyson, K., McFall, D., and Anand, S. (1996): "Mass appraisal for property taxation: an artificial intelligence approach". *Land Economics Review*; 2(1), pp. 25–32.

McCulloch, W.S. and Pitts, W. (1943): "A logical calculus of the ideas immanent in nervous activity". *Bulletin of Mathematical Biophysics*, 5, pp. 115–133.

National Statistics Institute (Instituto Nacional de Estadística –INE-): http://www.ine.es

Nguyen, N. and Cripps A. (2001): "Predicting housing value: a comparison of multiple regression analysis and artificial neural networks". *Journal of Real Estate Research*; 22(3), pp. 314–336.

Richardson, H.W. (1973): *Economía Regional. Teoría de la localización, estructuras urbanas y crecimiento* regional. Ed. Vicens Vives. Barcelona.

Rosen, S. (1974): "Hedonic Prices and Implicit Markets: Product Differentiation in Pure competition". *Journal of Political Economy*, 82, pp. 34–55.

Rossini, P. (1998): "Improving the results of artificial neural network models for residential valuation". *Four Annual Pacific-Rim Real Estate Society Conference. Perth, Western Australia*, 19-21 January.

Rumellhart, D. and McClelland, J. (1986): *Parallel distributed processing: Explorations in the microstructure of cognition*. Cambridge: MIT.

Saura, P. (1995) *Demanda de características de la vivienda en Murcia*. Secretariado de Publicaciones de la Universidad de Murcia.

Selim, H. (2009): "Determinants of house prices in Turkey: Hedonic regression versus artificial neural network". *Expert Systems with Applications*, 36, pp. 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, pp. 525–540.

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), pp. 185–201.