Real Estate valuation and forecasting in non-homogeneous markets: A case study in Greece during the financial crisis

Antonios K. Alexandridis[a,*], Dimitrios Karlis[b], Dimitrios Papastamos[c]

[a]*Kent Business School*
University of Kent
[b]*Department of Statistics*
Athens University of Economics and Business.
[c]*Eurobank Property Services S.A*

Abstract

In this paper we develop an automatic valuation model for property valuation using a large database of historical prices from Greece. The Greek property market is an inefficient, non-homogeneous market, still at its infancy and governed by lack of information. As a result modelling the Greek real estate market is a very interesting and challenging problem. The available data covers a big range of properties across time and includes the Greek financial crisis period which led to tremendous changes in the dynamics of the real estate market. We formulate and compare linear and non-linear models based on regression, hedonic equations, spatial analysis and artificial neural networks. The forecasting ability of each method is evaluated out-of-sample. Special care is given on measuring the success of the forecasts but also to identify the property characteristics that lead to large forecasting errors. Finally, by examining the strengths and the performance of each method we apply a combined forecasting rule to improve performance. Our results indicate that the proposed methodology constitutes an accurate tool for property valuation in non- homogeneous, newly developed markets.

Keywords: Forecasting, Property valuation, Real Estate, Neural Networks

---

*Corresponding author
Email address: `A.Alexandridis@kent.ac.uk.` (Antonios K. Alexandridis)

Preprint submitted to FEBS 2017 *May 15, 2017*

1. Introduction

In recent years big financial institutions are interested in creating and maintaining property valuation models. The main objective is to use reliable historical data in order to be able to forecast the price of a new property in a comprehensible manner and provide some indication for the uncertainty around this forecast. In this paper we develop an automatic valuation model for property valuation using a large database of historical prices from Greece. The Greek property market is an inefficient, non-homogeneous market, still at its infancy and governed by lack of information. As a result modelling the Greek real estate market is a very interesting and challenging problem.

The global crisis led to a significant decline in house prices. For years financial institutions were eager to lend money to home buyers. As a result when the house market collapsed financial institutions were the ones most affected with major financial losses. The global financial crisis was followed by the Greek crisis and a long period of recession. At the moment the Greek market is experiencing an unprecedented situation regarding the current valuations and the future trends. The "domino effect" of the financial crisis became apparent to Greece at the end of 2008. The residential market in Greece has experienced significant contraction over the last 8 years. It is highlighted that since the start of the financial crisis the private construction activity in Greece, as this depicted by the number of building permits reduced by almost 80% and the house prices showed a cumulative decrease[1] of 41% with the corresponding drop on prices, in metropolitan areas such as Athens and Thessaloniki, being 43.5% and 45.1 respectively. At the period 2008q1-2015q4, the ratio of non-performing loans to total bank loans increased by 30.9ppts (and by 38.4ppts if restructured loans are also taken into consideration), hitting 35.6% (and 43.5%, respectively) at the end of that period.

The need for unbiased, objective, systematic assessment of real estate property has always been important, and never more so than now where on the one hand banks need assurance that they have appraised a property on a fair value before issuing a loan and on the other hand the government needs to know the market value of a property in order to determine accordingly the annual property tax[2]. Furthermore, valuations determined for real estate property have further significant tax implications for current and new home owners and have to be verified in the courtroom in extreme cases.

Forecasters in the real-estate sector have to take into consideration the unique characteristics of property (Hoesli and MacGregor, 2000) such as heterogeneity, fixed locations, illiquidity, and the absence of a central marketplace. These characteristics make the real-estate market inefficient. For the above reasons automatic mass appraisal approaches could assist in the science of valuation especially in a world where there is increased availability and use of data, and where failure to achieve an opinion of value which takes proper and balanced account of such information and analysis may result in greater exposure to expensive litigation. Automatic valuation models (AVMs) or mass appraisal systems can enhance experts' valuation with data-drives estimates. They can provide model-based valuations for properties using information about the property's location and characteristics, appropriate for risk management and big-data analytics. Last but not least AVMs can be used for the redesign of the appraisal process. The automation features of the AVM can reduce the need for manual data collection and manipulation by the appraiser, while at the same time providing an independent estimate value. The role of the appraiser would be to evaluate the findings of the AVM in light of his own physical inspection of the property, verification of comparables and knowledge of local market conditions.

Traditional valuation methods include various expressions of linear regression including multiple, stepwise, quantile, robust and additive regression approaches using hedonic models incorporating features of the property such as its age, square feet of living space, number of bedrooms, plot size, and others, (Brunauer et al., 2013; Caples et al., 1997; Coleman and Larsen, 1991; Janssen et al., 2001; Isakson, 2001; Mark and Goldberg, 1988; Pagourtzi et al., 2003; Reichert, 1990; Schulz et al., 2014). The underlying hypothesis of these models is that the valuation of the residence can be related to a specific set of the propertys characteristics, (Kummerow, 2000).

Recently more advanced methodologies have been employed including neural networks, machine learning, fuzzy logic, multi-criteria decision analysis and spatial analysis (Ahn et al., 2012; Antipov and Pokryshevskaya, 2012;

[1]Monetary Policy Report 2015 - 2016, originally published in Greek (in June 2016, Chapter IV, Section 2).

[2]In 2016, Greeks were called to pay seven times more in property taxes compared to 2009, even though they had to deal with a 25 percent drop in GDP and similar unemployment percentages. Greece is one of the countries with the highest taxation of real estate as a percentage of GDP. According to European Commission figures for 2015, the only countries with higher property taxes are France and Britain. In particular, in Greece, property owners are required to pay taxes that exceed 2.5% of GDP, when in Germany the figure is no more than 0.5% while citizens of neighbouring countries, such as Italy, Cyprus, Bulgaria and Turkey enjoy less property taxes.

Atkinson and Crocker, 1987; Aznar and Guijarro, 2007; Aragonés-Beltrán et al., 2008; Ball and Srinivasan, 1994; Bitter et al., 2007; Brown and Uyar, 2004; Chica-Olmo, 2009; Helbich et al., 2014; Kaklauskas et al., 2007; Kilpatrick, 2011; Kontrimas and Verikas, 2011; Kusan et al., 2010; Landajo et al., 2012; Lins et al., 2005; McCluskey et al., 2013; Narula et al., 2012; Park and Bae, 2015; Peterson and Flanagan, 2009; Selim, 2009; Worzala et al., 1995; Zurada et al., 2011).

Although various studies have been published on mass appraisal systems, previous studies focus on large and already developed markets. Furthermore, the analysis is usually based on small samples (less than 500 properties) at regional level, (Landajo et al., 2012; Kilpatrick, 2011; Kusan et al., 2010; Selim, 2009; Kontrimas and Verikas, 2011; Narula et al., 2012; Brasington and Hite, 2008). In this study the proposed AVMs are tested in a new market still at its infancy with lots of unique characteristics. As it was already mentioned the Greek market is an inefficient, non-homogeneous market governed by lack of information. Also, there are declining prices due to the recession while the properties' characteristics are diverse both at regional and country level, for example differences in urban and rural areas or touristic areas of high demand. We examine the forecasting performance of the proposed methods in a very large data set (over 35,000 properties) in country level. Given the large size of our data set, we expect to derive significant conclusions about the strengths and weaknesses of each method as well as the dynamics that govern the Greek real estate property market. As this application illustrates AVMs can be applied to both case-by-case valuations and batch processing of thousands of properties.

In this study we develop three mass appraisal systems and we compare their forecasting power in 4 non-overlapping out-of-sample sets. The systems are based on hedonic characteristics and professional property valuations. Very few papers have examined the accuracy of professional forecasts in real estate, (Papastamos et al., 2015). The first method is linear and based on multiple linear regressions. The second valuation method uses spatial information. It is based on similarity measures and geographical distances in order to derive the price of a property. Finally, we apply a non-linear automatic valuation method based on machine learning. More precisely, we apply an optimised Neural Network (NN) in order to forecast real-estate prices based on hedonic characteristics. We give extra care in the construction of the NN. We apply statistical methods in order to select the appropriate number of hidden units as well as the statistical significant variables. Furthermore, we fine tune the NN using regularization methods in order to avoid over-fitting.

The rest of the paper is organised as follows. In section 2 we present the three valuation methods. More precisely we present three methods based on 1) Similarity Valuation Method, 2) Multiple Linear Regression and 3) NNs. The data are described in section 3. In section 4 we present our results. More precisely in section 4.1 the forecasting ability of each method is discussed while in section 4.2 an in-depth analysis of how the forecasting error changes when the characteristics of the properties change is presented. Section 5 concludes the paper.

2. Methodology

2.1. Multiple Regression Analysis

The model used was a typical hedonic regression model. Hedonic models assume that the price of a product reflects inherent characteristics valued by some implicit prices. In empirical studies, these implicit characteristic prices are coefficients that relate prices and the underlying attributes in a regression model.

The model takes the form

$$Y_i = \beta_0 + \sum_{j=1}^{k} \beta_j X_j^i + \epsilon_i$$

where X_i^j is the value of the j-th explanatory variable/characteristic for the i-th property, Y_i is the logarithm of the value of the property translated to value of the present period and $\beta_j,\ j = 0, \ldots, k$ are regression coefficients associated with the explanatory variables. The usual assumptions for the errors apply, namely zero mean and constant variance.

Estimation of the model was done using standard OLS approach. The variable selection approach however was not standard, and we will explain it later. There are many reasons for considering such a simple model. Regression models while simple they can reveal useful information about the underlying structure. Being simple offers certain advantages as a) it is easy to use and interpret, b) provides easy and stable variable selection approaches, c) modification is rather simple, the same for its update and generalization, and finally d) inference is simple and hence insight can be generated rather easily.

We applied a variable selection approach in order to find the variables that are predictive for the value of the property. The aims were: first to check existing work and whether it needs simplification with simpler models to attain parsimony, second to end up with a meaningful model to use and third to be able to derive a comprehensive and simple model in order to see the variables that are deemed useful for the purpose of the prediction.

Since the aim of the approach was to predict new unseen properties, we modified the forward selection approach in order to use it for creating a predictive model. Standard forward selection selects the new variables to add at the model among the significant ones. The reason is that the model building aims at producing a descriptive (exploratory) model that can help to identify the variables with relationship to the response. In our case the interest lies on prediction and hence we want to find a model that predicts well while it is not necessarily the best for exploring the existing data.

Hence our approach is the following

1. Start from a model with only the constant.
2. Select as the variable to enter the model the one that minimizes the mean of the relative absolute prediction error for the model k, defined as

$$MAPE_k = \sum_{i=1}^{n_t} \left| \frac{y_i - \hat{y}_i}{y_i} \right|$$

   where y_i is the observed value of the property and $\hat{y}_i$ is the predicted value of the model. In the above n_t is the cardinality of a validation.
3. With the selected variable in the model we go back to step 2 to find among the other candidates the one that minimizes the MAPE
4. Stop when no further decrease of MAPE is possible.

The approach mimics typical forward approaches but uses a criterion that selects a predictive model. An interesting note is that usually the MAPE after few steps almost stabilizes and further covariates create a small decrease. For predictive purposes this needs some care because it is known that for prediction the more covariates the more over-fitting is achieved and hence the model may loose its value very quickly. We tried to keep the model parsimonious, i.e. without many covariates. Finally, note that other variable screening/selection approaches like LASSO could have been used but since mainly the covariates are categorical special amendments for the LASSO were needed.

2.2. Similarity Measure Valuation

In this section the Similarity Measure Valuation (SMV) method is presented. The SMV method is based on a representative asset (RA). The RA is the "average" property derived from the database. This is a standard procedure in sales comparison methods since comparables have different characteristics, (Kauko and d'Amato, 2009). The value of each property is converted to a Hedonic Value (HV) based on the characteristics of the property and the Index area. The role of the HV is to convert all properties into a representative property in terms of characteristics. So, each comparable in our database has each own HV based on their characteristics compared to the RA and is given by:

$$HV_i = X_8^{RA} \exp\left[\ln\left(\frac{UV_i}{X_8^i}\right) + \sum_{j=1}^{J} \beta_{kj}\left(X_j^{RA} - X_j^i\right)\right] \tag{1}$$

where β_{kj} is the hedonic coefficient of variable j for the index area k where property i is located, X_j^i is the value of variable j for the property i, X_8^i is the size of the property in square meters and X_j^{RA} is the value of variable j for the RA. Finally, UV_i, is the updated value of property i. Our database consists of historical valuations, V_i, performed by experts. We are interested in updating each property's historical value to the current time where our method is used. Residential indices by region are used to update the values of comparables. The UV_i is given by:

$$UV_i = V_i \frac{ind_1}{ind_3}\left(\frac{ind_1}{ind_2}\right)^{\frac{m_1 - m_2}{3}} \tag{2}$$

where ind_1 is the residential index at the current quarter, ind_2 is the residential index of the previous quarter, ind_3 is the residential index of the initial quarter and m_1 and m_2 are the month of the quarter of valuation and the month of the quarter of the initial valuation respectively.

All available properties in the database are ranked based on their similarity with the property under consideration. A metric, W_{ij}, is defined to quantify the similarity:

$$W_{ij} = \exp\left[w_1 \ln\frac{c_1}{d_{ij} + c_1} + w_2 I_{ij}(X_7) + w_3 I_{ij}(X_8)\right] \tag{3}$$

The above formula assesses the similarity of property i to another property j from the database by considering the geographical distance between properties i and j, the administrative sector and the type of the property where:

- d_{ij} is the geographical distance between properties i and j.
- X_7, X_8 are the main characteristics of the properties as defined in section 3.
- $I_{ij}(x)$ is a $0-1$ indicator which equals 1 if properties i and j are identical in terms of their characteristic x and 0 if they differ on that characteristic.
- w_i are weighting coefficients, which sum up to 1; they indicate the relative importance of the different characteristics of the properties in defining the above similarity metric.
- c_1 is a scaling parameter for the distance, it is used to map the difference between the properties being compared on a similarity scale common to all characteristics. They are scalars used to avoid numerical problems.

The weights and the scaling parameters are adjusted differently for each administrative index area and they have been defined on the basis of inputs obtained from experts. The higher the similarity metric W_{ij} is, the stronger is the similarity between properties i and j.

After selecting the most suitable properties, a weighted RA value, $WRAV$, is obtained based on the following formula:

$$WRAV = \frac{\sum_{i=1}^{n} w_i HV_i}{\sum_{i=1}^{n} w_i} \tag{4}$$

Finally, we need to convert the $WRAV$ into the weighted value based on the property's under valuation characteristics.

$$SMV = X_8^i e^{\ln\left(\frac{WRAV}{X_8^{RA}}\right) + \sum_{j=1}^{J} \beta_{kj}\left(X_j^{RA} - X_j^i\right)} \tag{5}$$

where, as before, β_{kj} is the hedonic coefficient of variable j for the index area k where property i is located and X^i_j is the value of variable j for the property i,

2.3. Neural Networks

In this paper we treat NNs as the eminent expression of non-linear regression, which constitutes a very powerful approach, especially for financial applications. The main characteristic of NNs is their ability to approximate any non-linear function without making a priori assumptions about the nature of the process that created the available observations. A multilayer perceptron (MLP) is a feed-forward NN that utilizes a back-propagation learning algorithm in order to enhance the training of the network (Rumelhart et al., 1986).

For this study we propose a three-layer NN. The lower layer is called the input layer and consists of the input variables. The middle layer is the hidden layer and consists of hidden units (HUs). Finally, the upper layer is called the output layer where the approximation of the target values is estimated. Often more hidden layers can be used. Each node in one layer connects to each node in the next layer with a weight w_{ij}, where ij is the connection between two nodes in adjacent layers within the network. The units of each layer receive their inputs from the units of the layers immediately below and send their outputs to the units of the layers lying directly above. The flow of information is done through the connections. A sigmoid activation function is used in the hidden layer while there is a linear connection between the neurons and the output nodes, (see Cybenko (1989)).

On each pass through, the NN calculates the loss between the predicted output $\hat{y}_n$ at the output layer and the expected output y_n for the n^{th} iteration (epoch). The loss function used in this paper is the sum of squared errors, given by:

$$L_n = \frac{1}{2}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 \tag{6}$$

where N represents the total number of training points. Once the loss has been calculated, the back-propagation step begins by tracking the output error back through the network. The errors from the loss function are then used to update the weights for each node in the network, such that the network converges. Therefore, minimising the loss function requires w_{ij} to be updated repeatedly using gradient descent, so we update the weights at step $t+1$, $w_{ij,t+1}$, by using:

$$w_{ij,t+1} = w_{ij,t} - \eta\frac{\partial L}{\partial w_{ij,t}} + \mu\Delta w_{ij,t} \tag{7}$$

2.3.1. Parameter tuning for neural network generalisation improvement

A small number of HUs will lead to underfitting of the NN to the data while a very large number of HUs will lead to overfitting. In this study the model selection algorithms presented in Zapranis and Refenes (1999) and Alexandridis and Zapranis (2013, 2014) were adapted. One method for improving network generalization is to use a network that is just large enough to provide an adequate fit. Unfortunately, it is difficult to know beforehand how large a network should be for a specific application. In this study two methods for improving generalization are implemented: regularization and early stopping.

The default method for improving the generalization ability of a NN is called early stopping. In early stopping a relative large number of HUs is used in the construction of the network. The number of weights roughly defines the degrees of freedom of the network. If the training phase continues more than the appropriate iterations and the weights grow very large on the training phase then the network will start to learn the noise part of the data and will become overfitted. As a result the generalization ability of the network will be lost. Hence, it is not appropriate to use such a NN in predicting new unseen data. On the other hand, if the training is stopped at an appropriate point, it is possible to avoid overfitting.

A common practice to overcome the above problems is to use a validation sample. The in-sample data consists of property valuations in the period January 2012 - December 2015. In order to train a neural network the in-sample data were split into two samples. The first one is called the training sample which is used for computing the gradient and updating the network weights and biases as described in the previous section. The second subset is the validation set and is used to measure the generalisation ability of the network. The data were split randomly. The train sample consists of 85% of the in-sample data while the validation set of the 15% of the in-sample data. This ratio allows for a large enough sample for training for all Index areas and a large enough set for validation.

At each iteration, the NN is trained using the training sample. Then the cost function between the training data and the network output is estimated and it is used for the adjustment of the weights. The generalization ability of the network is measured using the validation sample. More precisely, the network is used to forecast the target values of the validation sample using the unseen input data of the validation sample. The error between the network output and the target data of the validation sample is calculated. At the beginning of the training phase the errors of both the training and the validation sample will start to decrease as the network weights are adjusted to the training data. After a particular iteration the network will start to learn the noise part of the data. As a result the error of the validation sample will start to increase. This is an indication that the network is starting to lose its generalization ability and the training phase must be stopped, (Anders and Korn, 1999; Dimopoulos et al., 1995). The network weights and biases are saved at the minimum of the validation set error. The network is trained using the Levenberg-Marquardt (LM) algorithm, (Samarasinghe, 2006). The LM alogorithm is very fast but is less efficient for a large network (with thousands of weights) as it requires a lot of memory. In this study the proposed network is very small and only few parameters are used for the minimisation of the fitness function.

Another approach to avoid overfitting is regularization. In regularization methods the weights of the network are trained in order to minimize the loss function plus a penalty term. Regularization is attempting to keep the overall growth of weights to a minimum by allowing only the important weights to grow. The rest of the weights are pulled towards zero, (Samarasinghe, 2006). This method is often called “weight decay”, (Samarasinghe, 2006). The regularization method tries to minimizes the sum:

$$W = L_n + \delta \sum_{j=1}^{J} w_j^2 \tag{8}$$

where the second term is the penalty term, w_j is a weight, J is the total number of weights in the network architecture and δ is a regularization parameter. The penalty term is not restricted to the above choice, (Anders and Korn, 1999; Samarasinghe, 2006).

It is desirable to determine the optimal regularization parameters in an automated fashion. We apply the Bayesian framework of MacKay (1992). A discussion of Bayesian regularization is beyond the scope of this study. For detailed expositions of the Bayesian regularizations and the LM training algorithm we refer to, for example, Dan Foresee and Hagan (1997). The Bayesian regularization provides a measure of the number of the weights that are being effectively used by the network. The effective number of weights should remain approximately the same, no matter how large the number of parameters in the network becomes. This assumes that the network has been trained for a sufficient number of iterations to ensure convergence.

3. Data description

This study focuses in the Greek property market. The available data covers a big range of properties across time and include the financial crisis period in Greece which led to tremendous changes in the dynamics of the real estate market.

The data were provided by the Eurobank Property Services S.A.[3]. The database represents the hedonic characteristics of real estate properties[4]. We have enriched our dataset with new variables by transformations and interactions between the initial variables [5].

The sample consists of $36,527$ properties that have been professionally evaluated in the period 2012 – 2016. In the majority of the studies regarding real estate value forecasting, very small datasets are used for testing the various methodologies (Landajo et al., 2012; Kilpatrick, 2011; Kusan et al., 2010; Selim, 2009; Kontrimas and Verikas, 2011; Narula et al., 2012; Brasington and Hite, 2008). An exception are the studies from Zurada et al. (2011); Peterson and Flanagan (2009). Given the large size of our datasets, we expect to derive significant conclusions about the strengths and weaknesses of each method as well as the dynamics that govern the Greek real estate property market.

The properties belong to 240 different administrative sectors covering all areas in Greece. In Figure 1 a map with all the properties used in this study is presented. The various characteristics of each property are presented in Tabel 1. We are interested in forecasting the valuation price of each property (V06). In Table 2 the descriptive statistics of each variable are presented. The values of the properties range from €$15,000$ to €$1,000,000$. Similarly, there is a lot of variation in the year of construction (from 1800 to 2016) and the size of the properties (from $12m^2$ to $400m^2$).

Valuations in the period of 2012 to December 2015 will constitute the in-sample period that will be used for the estimation and fitting of all models. Then, property prices in the first quarter of 2016 are forecasted. Finally, we follow a recursive procedure to forecast the property prices in the remaining quarters of 2016. Hence, the in-sample consists of $32,477$ properties while the out-of-sample contains $4,050$ properties.

Some sectors contain only a few observations, so we group the data into 32 aggregated administrative areas (Index areas). In Table 3 the number of properties per Index area, year of valuation, urban classification and type of property are presented. A closer inspection reveals that the majority of the properties are located in the capital or in large cities. This is expected as around 50% of the population in Greece lives in two cities, the capital – Athens – and Thessaloniki. Similarly, around 84.5% of the properties are flats while 6.5% are houses, 5.4% maisonettes and only the 3.6% of properties are of type duplex.

[3]Eurobank Property Services S.A. is the real estate arm of Eurobank group and is one of the largest real estate service providers in Greece as well as in South East Europe.

[4]We feel the need to stress that only the hedonic characteristics of each property were obtained and not any personal data of the clients.

[5]Due to confidentiality reasons we cannot report the new variables that have been resulted due to the transformations or the interaction between the initial variables. The original variables are reported in Table 1.

Table 1: Explanation of the initial set of variables

Code	Characteristic	Value
V01	Record code	Code
V02	Year of valuation	Year
V03	Month of valuation	Month no.
V04	Administrative sector	Code value
V05	Urban classification	Code value
V06	Survey value	Euro
V07	Type of residence	Code value
V08	Usable residence area	Sq. m.
V09	Land area	Sq. m.
V10	Year of construction	Year
V11	Distance from CBD	km
V12	Floor	Number
V13	Total number of floors	Number
V14	Existence of parking space	Yes/no (1/0)
V15	Type of parking	Yes/no (1/0)
V16	Type of heating	Code value (0-3)
V17	Quality of construction	Code value (0-3)
V18	Number of bedrooms	Number
V19	Touristic hotspot	Yes/no (1/0)
V20	Elevator	Yes/no (1/0)
V21	View	Yes/no (1/0)
V22	Number of bathrooms	Number

Table 2: Descriptive statistics

Var	Mean	St.Dev	Max	Median	Min	Skewness	Kurtosis
V06	123,592.14	112,403.08	1,000,000.00	90,000.00	15,000.00	3.20	17.01
V07	2.08	0.56	4.00	2.00	1.00	1.89	8.76
V08	97.64	52.42	400.00	87.30	12.00	2.07	9.31
V09	79.87	882.82	86000.00	0.00	0.00	57.29	4774.88
V10	1988	16.88	2016	1989	1800	-0.64	4.53
V11	28.33	45.38	321.20	8.30	0.00	2.75	12.21
V12	1.81	1.59	14.00	1.00	-1.00	1.01	4.05
V13	3.44	1.75	25.00	3.00	0.00	0.41	4.36
V14	0.14	0.36	3.00	0.00	0.00	2.44	8.30
V15	0.03	0.17	1.00	0.00	0.00	5.62	32.62
V16	1.16	0.58	3.00	1.00	0.00	0.67	4.36
V17	1.71	0.59	3.00	2.00	0.00	-0.02	2.70
V18	2.10	0.87	6.00	2.00	0.00	0.46	3.45
V19	0.14	0.34	1.00	0.00	0.00	2.14	5.57
V20	0.31	0.46	1.00	0.00	0.00	0.84	1.70
V21	0.08	0.27	1.00	0.00	0.00	3.12	10.72
V22	1.46	0.68	6.00	1.00	0.00	1.44	5.55

Table 3: Number of observations per Index area, year, urban classification and type of property

Index	All	In	Out
14	3977	3620	357
18	925	849	76
19	3442	2961	481
20	3157	2831	326
21	3354	2966	388
22	1063	919	144
23	2657	2355	302
24	425	390	35
25	1652	1426	226
27	1999	1845	154
28	1016	922	94
29	641	548	93
30	309	275	34
31	1441	1311	130
32	1106	996	110
33	517	477	40
34	216	201	15
35	1363	1177	186
36	595	536	59
37	820	755	65
38	413	360	53
39	1250	1027	223
40	153	140	13
41	838	702	136
42	387	309	78
43	73	68	5
44	832	765	67
45	384	347	37
46	507	461	46
47	432	398	34
48	253	235	18
49	330	305	25

Year	All	In	Out
2012	3951	3951	-
2013	13316	13316	-
2014	10017	10017	-
2015	5193	5193	-
2016	4050	-	4050

Urban Classifications	All	In	Out
Rural	1976	1615	361
Small town	2959	2580	379
Small city	2894	2571	323
Medium size city	3857	3483	374
Larger city	4510	4105	405
Capital	20331	18123	2208

Type of Property	All	In	Out
House	2395	2037	358
Flat	30838	27657	3181
Duplex	1315	1108	207
Maisonette	1979	1675	304

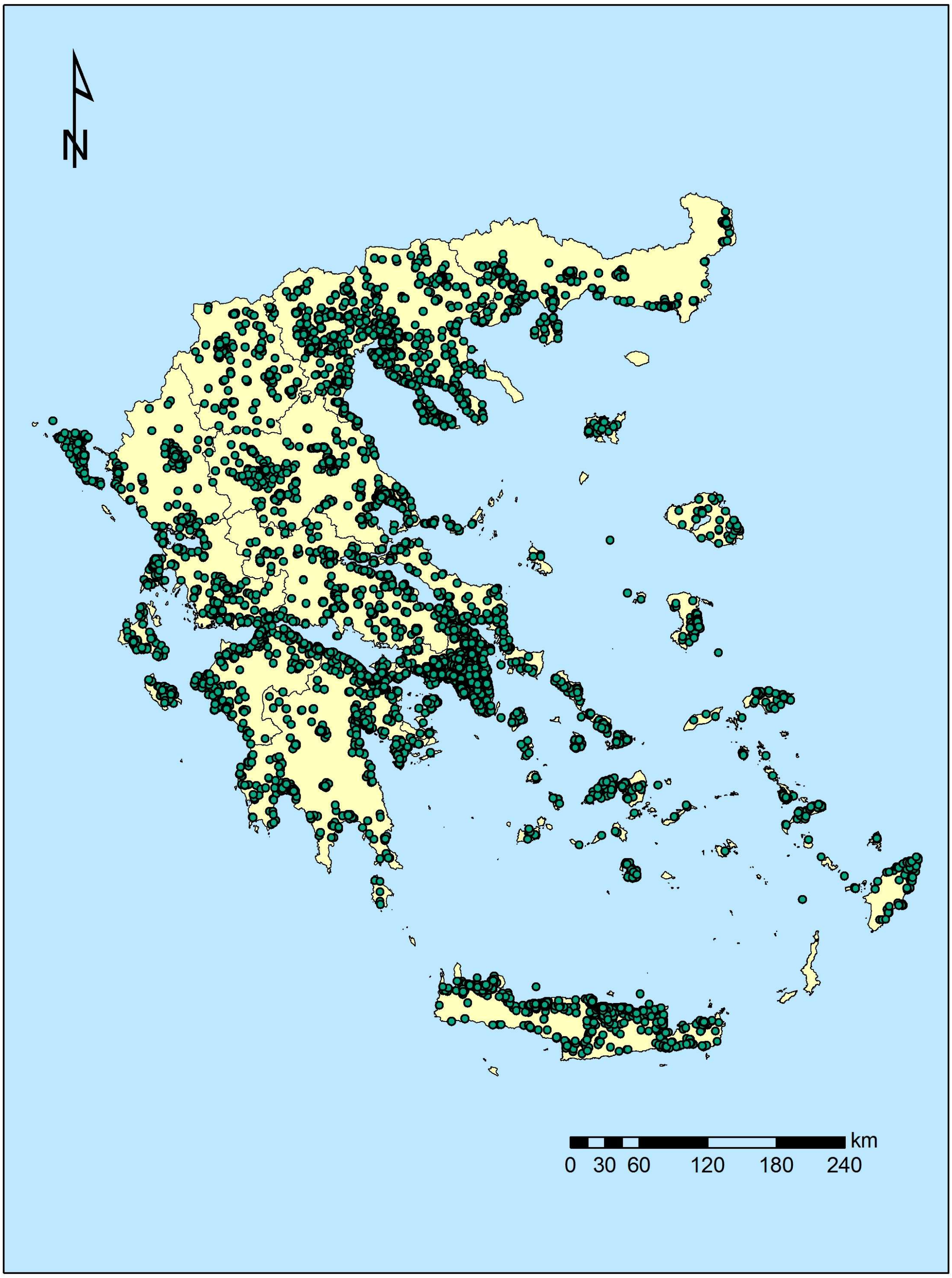


Figure 1: Geographical location of the real estate properties in Greece

4. Results

4.1. Initial results

In this section an out-of-sample validation of the proposed methodologies is provided. The three models are evaluated to 4 non-overlapping samples corresponding to the four quarters of 2016. We apply a recursive window forecasting scheme. Initially, the in-sample data consists of the property's valuations between the 1^{st} quarter of 2015 and the 4^{th} quarter of 2015 ($2012q1$ – $2015q4$). The out-of-sample data consists of the property's valuations that took place during the 1^{st} quarter of 2016 ($2016q1$). In the next step, $2016q1$ is included in the in-sample data set and we forecast the 2^{nd} quarter of 2016 ($2016q2$). Similarly for the remaining quarters. The total out-of-sample set consists of 4,050 observations.

Three error criteria are used for the evaluation of the forecasting ability of each method. The first one is the Mean Absolute Percentage Error (MAPE) and it is given by

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_i - \hat{y}_i}{y_i}\right|$$

The second error criterion, denoted by $P20$, measures the percentage of the cases where the MAPE is less than 20%. This is a standard metric used in real-estate valuations, (Rossini and Kershaw, 2008) and it is given by

$$P20 = \frac{100}{N}\sum_{i=1}^{N} 1_{|PE_i|<0.2}$$

where PE is the percentage error and it is given by

$$PE = \frac{y_i - \hat{y}_i}{y_i}$$

and $1_{|PE_i|<0.2}$ is an indicator function where

$$1_{|PE_i|<0.2} = \begin{cases} 1 & if \quad |PE_i| \leq 0.2 \\ 0 & if \quad |PE_i| > 0.2 \end{cases}$$

Finally, we calculate the squared correlation coefficient, R^2, between the predicted and the real prices.

In Table 4 a summary of the results is presented. More precisely, the Average PE, the standard deviation of the PE, the average MAPE, the average P20 and the R^2 are presented for each of the four quarters of 2016.

A closer inspection of Table 4 reveals that NN constantly outperforms the alternative methodologies. Interestingly, there is an indication that the results from NNs are more stable. They produce similar forecasting errors for all quarters. The MAPE ranges from 15.05% in the first quarter to 17.67% in the last. The MRA produce similar but slightly worse results for the 1^{st} quarter but the MAPE increases significantly for the remaining quarters. More precisely the MAPE increases from 15.34% to 20.72%. Finally, SMV seems to produce the largest out-of-sample forecasting errors ranging from 18.15% in the third quarter to 22.64% in the fourth.

Similarly, the P20 is always higher when NN are used, followed by MRA while SMV ranks last. With an exception of the last quarter, the P20 is always above 70% for the NN while for the MRA is above 70% for the first two quarters. Finally, it is always below 70% for the SMV. The R^2 is always higher when the NN is used. The MRA has a higher R^2 for the first and fourth quarter but it is lower for the second. Finally, MRA and SMV have the same R^2 in the third quarter.

In general, the MAPE increases in the third and fourth quarter indicating a change in the dynamics of the Greek housing market. However, one must be careful in the interpretation of the results since in each quarter a different set of properties is used. For example, in the 4^{th} quarter the number of properties with land area is doubled and they have significantly larger land area on average[6]. As it is shown in the next section, for all methods the MAPE is higher when properties with land area are considered. This is due to the fact that only 6.5% of the properties have land area.

[6]As an example we report a property with 86,000m^2 land area while in the historical data set the average land area was around $1,200m^2$. Properties with zero land area where excluded from the calculation of the mean.

Next we focus on combining the results. Two simple averaging approaches were used. The first approach is to take the average of the two best methods, the NN and MRA, as SMV give significantly higher errors. The second is to compute the average of all three methods.

In general both averaging schemes improved the results. Surprisingly, including the SMV method, further reduces the MAPE. More precisely, the NN+MRA produce the best results, outperforming the NN, in the first quarter with a MAPE of only 14.54% and a P20 of 77.00%. For quarters 2 and three the best results are given by the SMV+NN+MRA approach with a MAPE of 15.70% and 15.89% respectively while the P20 is 71.76% and 71.33%. Finally, for the last quarter neither averaging technique can outperform the NN with respect to the MAPE although the P20 when all three methods are used was increased to 69.03% from 68.28% in the case of NN.

The MAPE values for each index for all 5 approaches are presented in Tables 5- 8. Similarly the values of the P20 for each index for all 5 approaches are presented In Tables 9- 12.

Table 13 shows how many times each method had the best predictive performance with (bottom) and without (top) averaging methods, i.e. in how many index areas each method outperforms all the others in each quarter. In summary, the NN outperformed the alternative methods in 63 cases out of the 126. The MRA method produced the most accurate forecasts 39 times while the SMV only 24. A closer inspection of Table 13 reveals that the NN outperform the other methods in all quarters while MRA and SMV always rank second and third respectively. Taking into consideration the two averaging techniques SMV+NN+MRA produces the lower MAPE in 40 cases while NN in 38. The MRA outperform all the other methods in 25 cases while the SMV in 16. Finally, the NN+MRA give the best forecasts only in 9 index areas. Breaking down our results by quarter we observe that in the first quarter the NN method ranks first and the NN+MRA+SMV ranks second while the opposite holds in the fourth quarter. In the second and third quarter both methods rank first.

Recently artificial intelligence based methods have been proposed as an alternative for mass assessment. However, there are mixed results. Guan et al. (2009) finds no improvement when advance machine learning techniques are used while Worzala et al. (1995) find that NN based methods to be inferior to traditional regression methods. More precisely Worzala et al. (1995) find that NN-based methods do not produce results that are notably better than those of MRA except when more homogeneous data are used.

In contrast in this study where NN were fine tuned and extra care was taken to avoid overfitting our results indicate that NN can significantly outperform traditional vaulation methods.

4.2. Analysis of the forecasting errors

In this section we analyse the forecasting errors of each methodology. More precisely, we examine how the forecasting error changes when the characteristics of the properties change [7]. Due to space limitations we focus on the characteristics where the analysis is more interesting. Our analysis is based on the complete out-of-sample period (2016q1-2016q4) and consists of 4,050 properties.

In Figure 2 the MAPE for each Index area is presented. In addition the number of observations for each Index area is depicted in Figure 2. A closer inspection of Figure 2 reveals that the variation of the MAPE is greater in the case of SMV compared to the other methods. For all methods the MAPE is greater when only few observations are present while the lower MAPE for all indices is obtained when the average of the three methods is considered. Finally, the MAPE is similar across all indices for the NN, the MRA and the averaging method while is quite different for the SMV method.

The MAPE per urban classification is presented in Figure 3. The performance of all methods is the same. The MAPE is higher for rural areas and small towns while it is significantly lower for small, medium and large cities and the capital. While the number of observations per urban classification is the same in the out-of-sample set (except the capital) this is not the case in the in-sample. Table 5 shows that the majority of the properties are located in the capital (18,123) while there are 3,483 and $2,571$ in large and medium size cities. On the other hand there are only 1,615 properties in rural areas in the in-sample where the out-of-sample MAPE is greater. Similar results are presented in Figure 4 where the MAPE per type of property is presented. More precisely the MAPE is lower for flats while it is large for houses. The majority of the properties are flats, 3,181, while only 356, 256 and 304 properties are of type house, duplex and maisonette respectively.

[7] Due to space limitation we cannot present the results for all the characteristics, however the results are available from the authors upon request.

Table 4: Out-of-sample performance for the four quarters of 2016

	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
2016q1					
Av. PE	6.93%	2.05%	1.25%	1.65%	3.41%
Std. PE	0.1400	0.0423	0.0452	0.0392	0.0510
MAPE	19.73%	**15.05%**	15.34%	**14.54%**	14.86%
P20	66.63%	**75.54%**	75.42%	**77.00%**	75.99%
R^2	81.13%	**86.98%**	86.85%	**88.31%**	88.11%
2016q2					
Av. PE	0.67%	1.86%	1.01%	1.43%	1.18%
Std. PE	0.0722	0.0532	0.0581	0.0500	0.0474
MAPE	18.30%	**16.22%**	17.46%	16.19%	**15.70%**
P20	67.27%	**72.06%**	68.06%	71.06%	71.76%
R^2	81.71%	**85.71%**	78.18%	84.14%	85.62%
2016q3					
Av. PE	3.20%	1.61%	0.10%	0.85%	1.63%
Std. PE	0.0661	0.0511	0.0604	0.0502	0.0487
MAPE	18.15%	**16.67%**	18.13%	16.48%	**15.89%**
P20	66.19%	**70.97%**	65.95%	69.65%	**71.33%**
R^2	84.44%	**85.64%**	84.44%	87.03%	**87.70%**
2016q4					
Av. PE	10.18%	3.91%	2.45%	3.18%	5.51%
Std. PE	0.7240	0.2448	0.2807	0.2390	0.3227
MAPE	22.64%	**17.67%**	20.72%	17.80%	18.10%
P20	65.33%	**68.28%**	60.80%	67.15%	**69.03%**
R^2	78.65%	**88.25%**	80.08%	87.75%	**88.29%**

Next, in Figure 5 the MAPE per usable residence area is presented. For the SMV the error is minimised for properties between $50m^2$ and $80m^2$ while it is significantly larger for any other category. For the NN and the MRA the results are similar. The MAPE is lower for properties up to $120m^2$ and then it increases as the area increases. Finally, the MAPE for the NN is smaller for every category.

When the land area is considered (Figure 6) all methods produce significantly higher errors. However, the SMV produces significantly higher errors. More precisely, when land area is included the MAPE for the SMV is 0.40 while it is only 0.29 for the remaining methods. When the properties do not have any land the MAPE falls to 0.18 and 0.17 for the SMV and the MRA respectively while it is only 0.15 for the NN and the averaging method. Again the lower errors for each category are obtained by the NN and the averaging method.

Next, in Figure 7 we examine the effect of the age of the property to the forecasting ability of the models. It is clear that the MAPE is higher for properties constructed before 1970. Also the variation for the SMV is higher compared to the other methods while it is relative stable for the remaining methods. Again, the lower MAPE per category is obtained by the NNs.

Finally, in Figure 8 the MAPE per number of bedrooms is presented. The MAPE is high for the SMV, MRA and the averaging method when properties with 0 or 5 bedrooms are considered. On the other hand, for the NNs the MAPE increases for properties with 5 or 6 bedrooms. A closer inspection reveals that the majority of the properties have 1–3 bedrooms. Again, the best results for all categories are obtained for the NN and the averaging methods.

Table 5: 2016q1 MAPE per Index

Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	17.24%	13.91%	17.82%	14.40%	14.17%
18	10.42%	15.00%	12.85%	13.93%	12.39%
19	17.86%	16.38%	13.90%	14.34%	14.89%
20	16.91%	12.91%	11.72%	12.14%	13.51%
21	13.09%	12.14%	13.58%	12.25%	11.25%
22	24.97%	14.57%	19.62%	16.27%	15.23%
23	11.67%	10.29%	10.42%	9.69%	9.55%
24	28.35%	15.51%	10.03%	10.90%	14.53%
25	18.75%	14.66%	15.78%	14.65%	14.68%
27	16.56%	14.50%	15.20%	14.49%	12.90%
28	12.58%	6.38%	9.49%	7.49%	7.49%
29	17.28%	17.71%	13.12%	14.76%	14.78%
30	106.28%	26.12%	14.53%	19.17%	46.40%
31	22.35%	22.26%	22.13%	22.03%	21.66%
32	17.30%	13.84%	11.33%	11.51%	12.14%
33	26.83%	11.66%	14.00%	12.14%	16.49%
34	6.05%	28.38%	21.55%	17.18%	9.43%
35	21.10%	18.32%	19.37%	17.98%	16.94%
36	27.27%	18.92%	20.42%	19.35%	19.06%
37	20.67%	18.52%	22.60%	20.56%	18.62%
38	23.79%	15.12%	17.43%	13.94%	12.91%
39	18.33%	11.74%	12.39%	12.04%	11.79%
40	42.75%	20.97%	24.79%	22.88%	21.88%
41	23.45%	23.29%	26.00%	24.19%	22.39%
42	45.19%	27.72%	22.13%	24.88%	29.69%
44	10.87%	12.47%	10.06%	10.37%	9.38%
45	11.62%	11.41%	5.85%	6.93%	7.38%
46	23.11%	10.25%	14.53%	10.88%	14.71%
47	23.37%	10.43%	15.07%	11.74%	14.66%
48	10.45%	12.37%	9.71%	8.81%	8.77%
49	13.00%	10.11%	12.54%	10.18%	11.12%
Average	19.73%	15.05%	15.34%	14.54%	14.86%

Table 6: 2016q2 MAPE per Index

Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	18.38%	16.99%	24.32%	19.87%	18.31%
18	10.85%	14.88%	14.05%	13.92%	10.59%
19	15.70%	16.09%	15.71%	15.55%	14.58%
20	16.97%	17.34%	17.77%	17.13%	16.39%
21	16.90%	11.40%	14.93%	11.99%	12.31%
22	12.93%	14.07%	17.75%	15.64%	13.36%
23	19.34%	16.41%	12.89%	14.26%	15.28%
24	15.74%	19.90%	15.86%	14.78%	12.49%
25	22.62%	23.51%	22.54%	22.19%	21.77%
27	14.77%	13.33%	16.31%	13.94%	12.79%
28	19.95%	12.59%	10.78%	11.58%	12.25%
29	13.07%	11.32%	16.27%	13.44%	13.03%
30	26.60%	25.28%	28.88%	26.67%	25.99%
31	16.76%	17.24%	18.53%	17.83%	15.02%
32	16.88%	15.40%	16.13%	14.99%	14.85%
33	10.96%	15.25%	5.90%	10.15%	10.38%
34	32.75%	30.30%	48.18%	39.24%	37.07%
35	20.97%	15.77%	16.23%	15.24%	15.28%
36	14.69%	14.17%	20.70%	16.86%	13.92%
37	21.77%	19.34%	22.65%	20.68%	19.98%
38	12.53%	20.63%	24.81%	22.06%	16.71%
39	15.00%	18.70%	21.36%	19.73%	17.90%
40	41.98%	31.46%	22.49%	26.79%	27.62%
41	29.58%	13.58%	13.86%	12.53%	17.09%
42	26.62%	22.76%	26.92%	22.84%	23.38%
43	42.83%	6.36%	11.72%	9.04%	20.30%
44	16.99%	11.38%	10.83%	10.85%	11.73%
45	22.93%	8.90%	12.35%	8.70%	11.60%
46	11.31%	19.99%	9.55%	13.75%	9.86%
47	22.11%	22.15%	12.92%	17.06%	16.24%
48	30.25%	10.48%	16.47%	13.47%	15.24%
49	20.64%	25.86%	16.83%	20.35%	18.59%
Average	18.30%	16.22%	17.46%	16.19%	15.70%

Table 7: 2016q3 MAPE per Index

Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	14.39%	14.40%	17.23%	14.16%	12.87%
18	14.96%	11.85%	14.75%	12.75%	12.68%
19	15.50%	14.98%	14.27%	13.81%	13.63%
20	19.39%	19.10%	18.35%	18.61%	18.47%
21	14.70%	12.91%	14.65%	13.18%	12.25%
22	21.02%	17.64%	21.74%	19.46%	18.79%
23	18.35%	15.84%	16.93%	15.88%	16.21%
24	22.90%	18.71%	25.73%	15.99%	15.56%
25	24.29%	20.86%	20.35%	19.69%	20.25%
27	23.69%	16.21%	21.49%	18.32%	19.23%
28	18.23%	10.82%	14.99%	12.13%	13.53%
29	13.54%	15.83%	16.89%	15.24%	13.92%
30	25.80%	13.43%	14.61%	14.02%	17.95%
31	17.27%	19.65%	23.96%	21.18%	18.22%
32	14.65%	10.23%	10.75%	10.32%	10.72%
33	15.41%	26.22%	23.69%	24.86%	20.98%
34	46.84%	51.64%	49.15%	50.40%	49.21%
35	13.44%	17.07%	20.40%	18.06%	15.67%
36	9.46%	14.11%	12.81%	11.86%	7.45%
37	14.21%	14.18%	14.24%	13.76%	12.75%
38	26.75%	19.92%	25.58%	19.35%	16.93%
39	19.27%	17.93%	22.74%	18.66%	16.15%
40	25.81%	20.75%	20.12%	17.60%	13.54%
41	25.58%	23.69%	20.69%	19.65%	21.01%
42	36.75%	34.07%	31.33%	29.55%	29.20%
44	15.65%	16.82%	20.48%	18.52%	17.41%
45	17.45%	10.86%	10.43%	10.46%	11.47%
46	25.84%	16.64%	17.49%	16.99%	19.68%
47	14.95%	9.84%	22.02%	14.98%	13.83%
48	49.74%	62.98%	45.57%	54.27%	52.76%
49	20.79%	16.93%	29.46%	21.30%	19.48%
Average	18.15%	16.67%	18.13%	16.48%	15.89%

Table 8: 2016q4 MAPE per Index

Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	19.38%	17.11%	28.91%	21.26%	18.92%
18	23.13%	20.93%	22.57%	21.12%	20.42%
19	16.26%	16.95%	17.26%	16.17%	15.40%
20	17.39%	15.28%	17.25%	15.53%	14.98%
21	15.32%	14.35%	16.95%	14.91%	14.05%
22	14.33%	13.57%	23.25%	16.84%	14.93%
23	15.42%	11.95%	19.61%	15.18%	14.80%
24	25.78%	20.86%	16.06%	13.89%	13.19%
25	20.33%	17.61%	17.87%	16.60%	16.11%
27	19.29%	13.03%	21.52%	16.71%	15.51%
28	22.51%	15.42%	21.11%	17.38%	17.66%
29	21.12%	14.18%	14.14%	12.96%	13.85%
30	79.61%	20.15%	31.52%	23.88%	42.04%
31	93.61%	27.30%	21.04%	21.96%	45.65%
32	19.18%	18.94%	22.78%	19.29%	18.68%
33	10.77%	16.92%	6.19%	9.35%	9.04%
34	45.20%	75.29%	38.58%	45.86%	41.02%
35	27.75%	16.93%	25.08%	19.41%	20.83%
36	45.64%	33.10%	31.42%	26.22%	32.24%
37	31.93%	26.82%	24.29%	25.45%	26.26%
38	73.64%	38.25%	27.14%	28.07%	42.41%
39	16.15%	16.87%	19.24%	16.96%	15.64%
40	29.62%	47.43%	17.85%	27.69%	14.79%
41	28.59%	18.31%	34.35%	24.86%	24.73%
42	27.58%	27.66%	33.34%	28.92%	24.71%
43	15.29%	42.27%	50.21%	41.37%	32.68%
44	17.91%	14.16%	14.91%	12.80%	12.83%
45	13.47%	13.37%	19.63%	14.46%	12.77%
46	14.50%	15.54%	15.88%	13.65%	12.73%
47	8.57%	21.57%	18.80%	16.79%	10.66%
48	80.91%	34.30%	16.90%	22.79%	42.16%
49	43.99%	31.32%	3.65%	13.95%	8.60%
Average	22.64%	17.67%	20.72%	17.80%	18.10%

Table 9: 2016q1 P20 per Index

Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	67.61%	71.83%	59.15%	70.42%	77.46%
18	86.67%	93.33%	86.67%	86.67%	93.33%
19	68.49%	69.86%	75.34%	69.86%	72.60%
20	70.59%	81.18%	83.53%	85.88%	82.35%
21	81.16%	81.16%	78.26%	81.16%	84.06%
22	59.38%	75.00%	75.00%	71.88%	81.25%
23	84.00%	84.00%	84.00%	82.00%	86.00%
24	50.00%	75.00%	75.00%	75.00%	75.00%
25	67.27%	74.55%	80.00%	78.18%	76.36%
27	68.57%	80.00%	77.14%	82.86%	80.00%
28	87.50%	100.00%	87.50%	95.83%	95.83%
29	73.68%	57.89%	68.42%	73.68%	68.42%
30	44.44%	55.56%	77.78%	55.56%	55.56%
31	51.28%	56.41%	61.54%	61.54%	56.41%
32	65.22%	82.61%	82.61%	82.61%	78.26%
33	53.85%	84.62%	76.92%	76.92%	76.92%
34	100.00%	50.00%	50.00%	50.00%	100.00%
35	56.00%	66.00%	70.00%	74.00%	70.00%
36	43.48%	65.22%	73.91%	69.57%	60.87%
37	68.75%	81.25%	75.00%	81.25%	81.25%
38	50.00%	75.00%	75.00%	75.00%	75.00%
39	66.10%	84.75%	88.14%	88.14%	76.27%
40	0.00%	50.00%	50.00%	50.00%	50.00%
41	60.87%	65.22%	52.17%	56.52%	56.52%
42	34.48%	41.38%	48.28%	48.28%	31.03%
44	80.95%	90.48%	95.24%	95.24%	95.24%
45	85.71%	85.71%	100.00%	85.71%	100.00%
46	44.44%	100.00%	77.78%	77.78%	77.78%
47	44.44%	77.78%	77.78%	88.89%	77.78%
48	100.00%	100.00%	66.67%	100.00%	100.00%
49	83.33%	100.00%	66.67%	100.00%	100.00%
Average	66.63%	75.54%	75.42%	77.00%	76.0%

Table 10: 2016q2 P20 per Index

Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	63.54%	70.83%	53.13%	58.33%	60.42%
18	86.96%	78.26%	69.57%	73.91%	82.61%
19	71.03%	70.09%	71.96%	70.09%	73.83%
20	67.61%	64.79%	64.79%	64.79%	63.38%
21	78.82%	84.71%	75.29%	83.53%	80.00%
22	77.27%	75.00%	68.18%	77.27%	77.27%
23	71.62%	78.38%	72.97%	75.68%	74.32%
24	60.00%	70.00%	70.00%	70.00%	80.00%
25	63.64%	65.91%	65.91%	68.18%	68.18%
27	64.86%	75.68%	72.97%	81.08%	83.78%
28	73.33%	73.33%	83.33%	83.33%	80.00%
29	76.00%	84.00%	68.00%	76.00%	80.00%
30	45.45%	45.45%	36.36%	45.45%	27.27%
31	69.44%	66.67%	55.56%	61.11%	75.00%
32	75.00%	70.45%	75.00%	77.27%	79.55%
33	80.00%	80.00%	100.00%	80.00%	80.00%
34	25.00%	25.00%	0.00%	25.00%	25.00%
35	51.16%	69.77%	65.12%	65.12%	62.79%
36	66.67%	77.78%	66.67%	55.56%	88.89%
37	61.90%	61.90%	66.67%	61.90%	57.14%
38	88.89%	55.56%	55.56%	55.56%	55.56%
39	77.27%	65.91%	52.27%	61.36%	72.73%
40	33.33%	33.33%	66.67%	66.67%	66.67%
41	32.73%	83.64%	81.82%	83.64%	72.73%
42	56.25%	56.25%	56.25%	62.50%	50.00%
43	0.00%	100.00%	100.00%	100.00%	0.00%
44	71.43%	85.71%	78.57%	85.71%	92.86%
45	42.86%	85.71%	71.43%	85.71%	85.71%
46	90.00%	80.00%	90.00%	80.00%	80.00%
47	50.00%	40.00%	90.00%	80.00%	60.00%
48	60.00%	80.00%	60.00%	80.00%	80.00%
49	77.78%	55.56%	77.78%	55.56%	77.78%
Average	67.27%	72.06%	68.06%	71.06%	71.76%

Table 11: 2016q3 P20 per Index

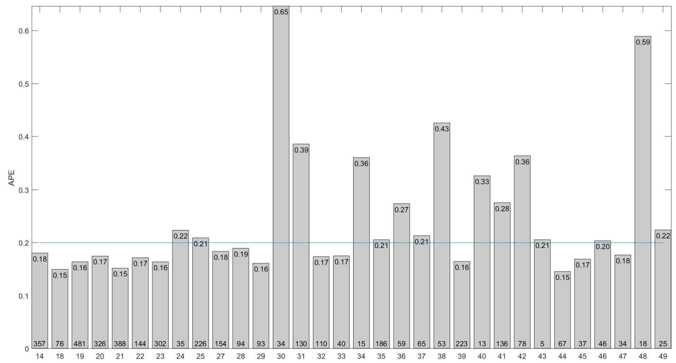
Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	76.67%	66.67%	70.00%	75.00%	81.67%
18	65.00%	85.00%	65.00%	75.00%	80.00%
19	69.81%	71.70%	75.47%	72.64%	73.58%
20	61.29%	64.52%	66.13%	67.74%	64.52%
21	73.91%	81.52%	75.00%	80.43%	83.70%
22	63.16%	68.42%	57.89%	63.16%	68.42%
23	73.08%	78.21%	65.38%	70.51%	75.64%
24	50.00%	66.67%	66.67%	83.33%	83.33%
25	56.52%	63.04%	65.22%	65.22%	60.87%
27	51.85%	62.96%	59.26%	59.26%	55.56%
28	81.82%	90.91%	72.73%	81.82%	81.82%
29	75.00%	75.00%	68.75%	68.75%	68.75%
30	50.00%	100.00%	100.00%	100.00%	100.00%
31	66.67%	61.90%	57.14%	52.38%	71.43%
32	66.67%	100.00%	88.89%	94.44%	88.89%
33	57.14%	28.57%	42.86%	42.86%	42.86%
34	50.00%	50.00%	0.00%	0.00%	0.00%
35	82.00%	70.00%	62.00%	62.00%	68.00%
36	90.91%	90.91%	72.73%	81.82%	100.00%
37	70.59%	76.47%	76.47%	76.47%	76.47%
38	40.00%	60.00%	50.00%	50.00%	50.00%
39	67.39%	65.22%	54.35%	71.74%	76.09%
40	33.33%	66.67%	33.33%	33.33%	66.67%
41	42.86%	54.29%	65.71%	71.43%	54.29%
42	30.77%	46.15%	38.46%	30.77%	46.15%
44	66.67%	60.00%	40.00%	60.00%	46.67%
45	61.54%	84.62%	84.62%	76.92%	84.62%
46	35.29%	70.59%	52.94%	58.82%	47.06%
47	66.67%	100.00%	50.00%	83.33%	100.00%
48	0.00%	0.00%	0.00%	0.00%	0.00%
49	57.14%	71.43%	42.86%	57.14%	57.14%
Average	66.19%	70.97%	65.95%	69.65%	71.33%

Table 12: 2016q4 P20 per Index

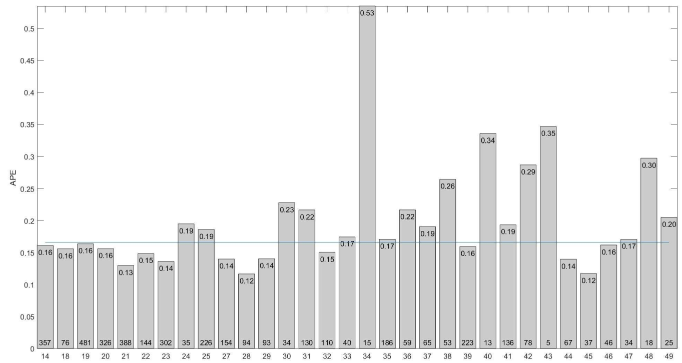
Index	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
14	70.00%	65.38%	45.38%	55.38%	60.77%
18	61.11%	66.67%	50.00%	55.56%	66.67%
19	70.77%	69.23%	67.18%	71.79%	73.33%
20	64.81%	75.00%	77.78%	77.78%	80.56%
21	71.13%	78.17%	67.61%	74.65%	80.99%
22	79.59%	79.59%	44.90%	63.27%	71.43%
23	77.00%	80.00%	54.00%	73.00%	73.00%
24	73.33%	66.67%	80.00%	73.33%	86.67%
25	58.02%	71.60%	72.84%	74.07%	71.60%
27	58.18%	76.36%	65.45%	69.09%	67.27%
28	58.62%	72.41%	58.62%	68.97%	75.86%
29	63.64%	69.70%	78.79%	81.82%	75.76%
30	8.33%	50.00%	16.67%	50.00%	16.67%
31	50.00%	50.00%	50.00%	44.12%	41.18%
32	76.00%	56.00%	56.00%	72.00%	68.00%
33	86.67%	60.00%	100.00%	86.67%	93.33%
34	28.57%	0.00%	14.29%	28.57%	42.86%
35	53.49%	72.09%	51.16%	65.12%	60.47%
36	37.50%	43.75%	43.75%	37.50%	43.75%
37	36.36%	36.36%	63.64%	45.45%	45.45%
38	27.27%	27.27%	36.36%	36.36%	36.36%
39	72.97%	70.27%	60.81%	70.27%	74.32%
40	40.00%	20.00%	40.00%	40.00%	80.00%
41	52.17%	60.87%	52.17%	65.22%	52.17%
42	50.00%	35.00%	35.00%	25.00%	35.00%
43	75.00%	25.00%	0.00%	25.00%	25.00%
44	76.47%	76.47%	70.59%	70.59%	70.59%
45	80.00%	80.00%	80.00%	90.00%	90.00%
46	70.00%	70.00%	60.00%	80.00%	70.00%
47	100.00%	66.67%	66.67%	66.67%	88.89%
48	0.00%	22.22%	66.67%	44.44%	11.11%
49	33.33%	66.67%	100.00%	66.67%	100.00%
Average	65.33%	68.28%	60.80%	67.15%	69.03%

Table 13: Number of best predictive performance of the three main methods with (bottom) and without (top) averaging methods

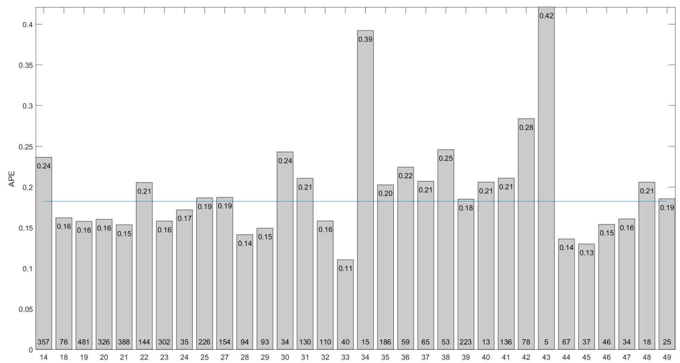
Main Methods	SMV	NN	MRA		
Q1	2	18	11		
Q2	8	15	9		
Q3	8	15	8		
Q4	6	15	11		
All methods	SMV	NN	MRA	NN + MRA	SMV+NN+MRA
Q1	2	11	8	1	9
Q2	3	9	8	3	9
Q3	6	10	3	2	10
Q4	2	8	7	3	12



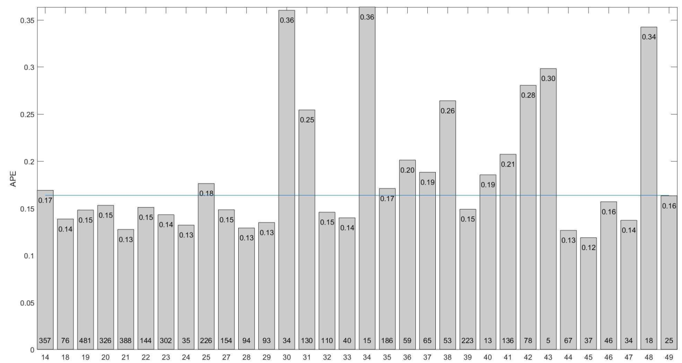
(a) SMV



(b) NN

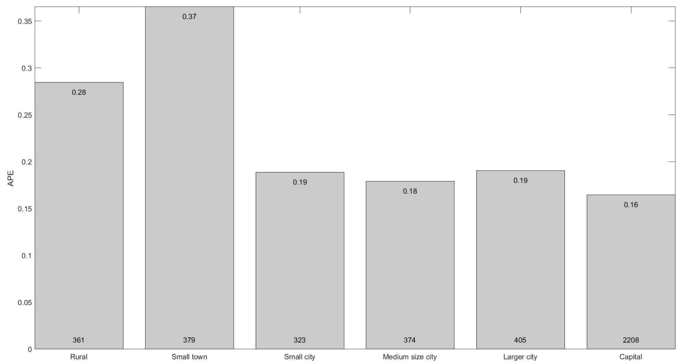


(c) MRA

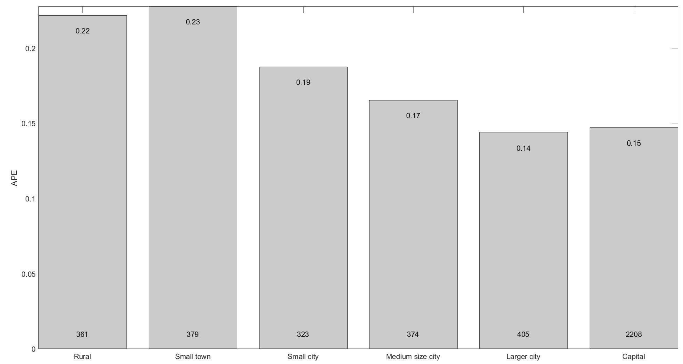


(d) SMV+NN+MRA

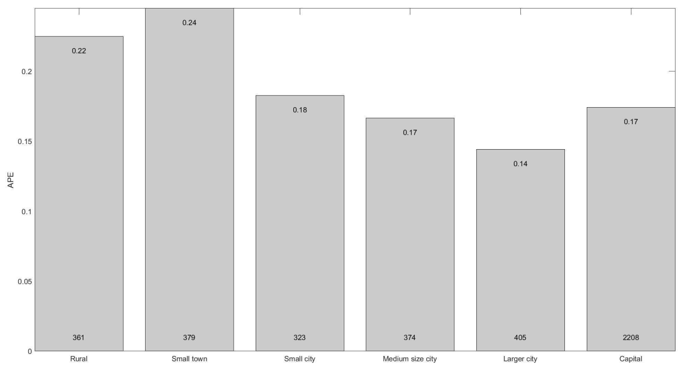
Figure 2: Mean absolute percentage error per index area. The horizontal lines is the average error across all index areas.



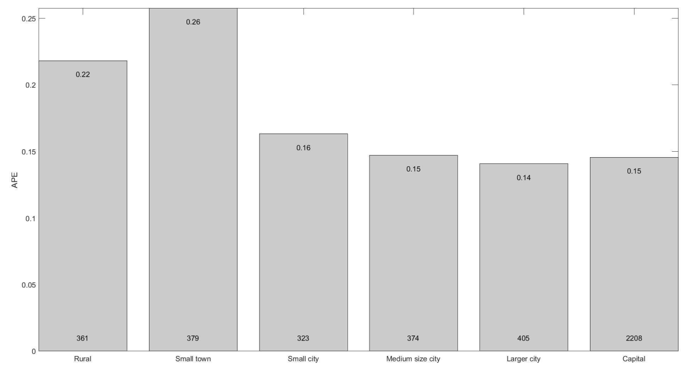
(a) SMV



(b) NN

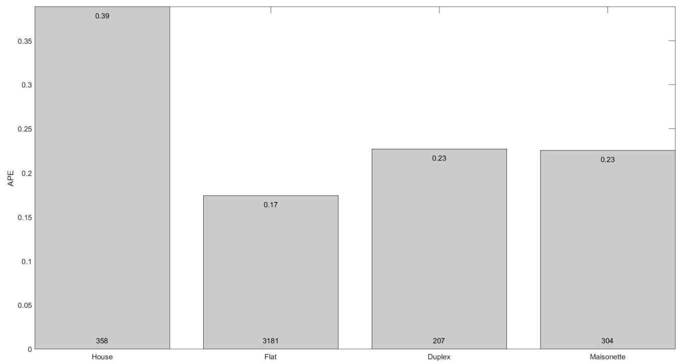


(c) MRA

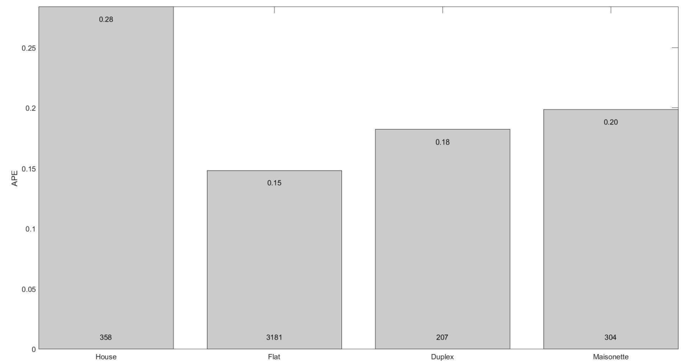


(d) SMV+NN+MRA

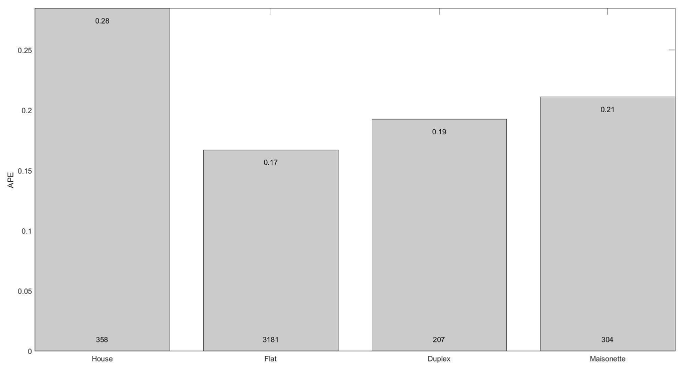
Figure 3: Mean absolute percentage error per urban classification (V05).



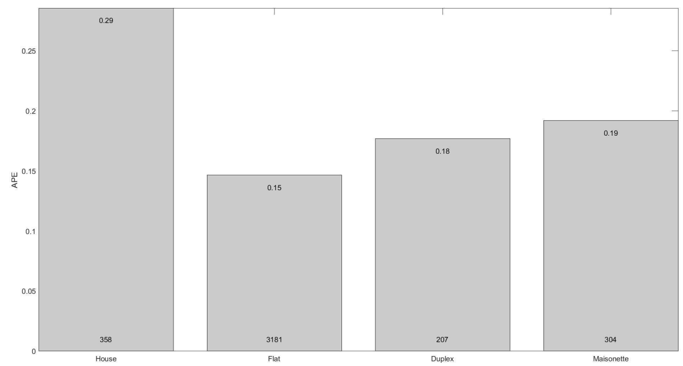
(a) SMV



(b) NN

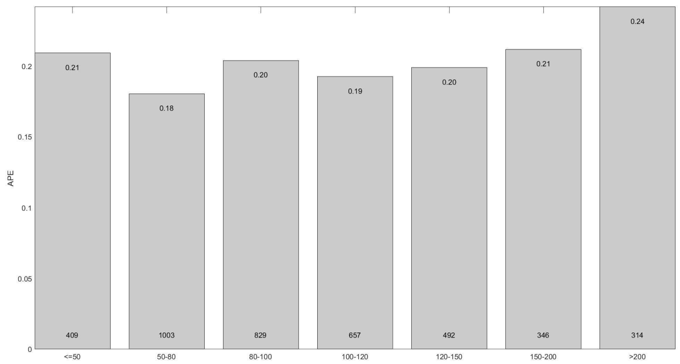


(c) MRA

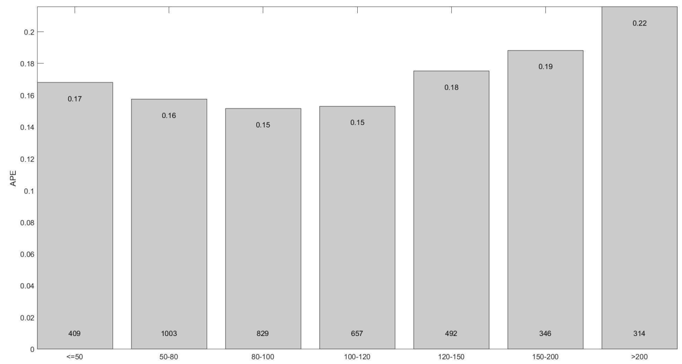


(d) SMV+NN+MRA

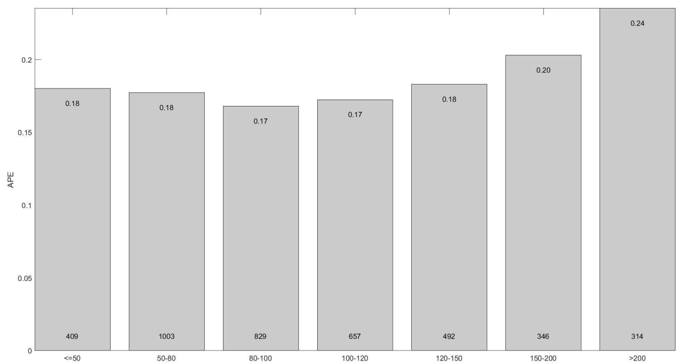
Figure 4: Mean absolute percentage error per type of residence (V07).



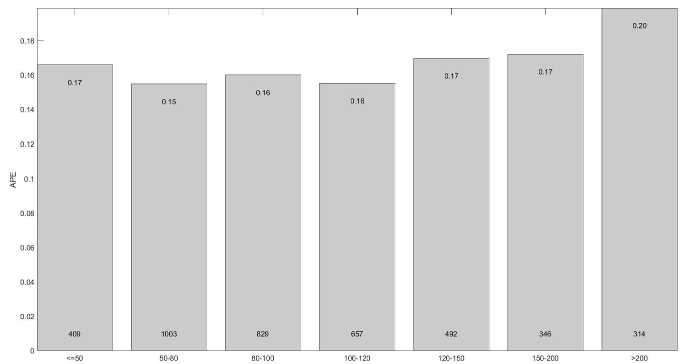
(a) SMV



(b) NN

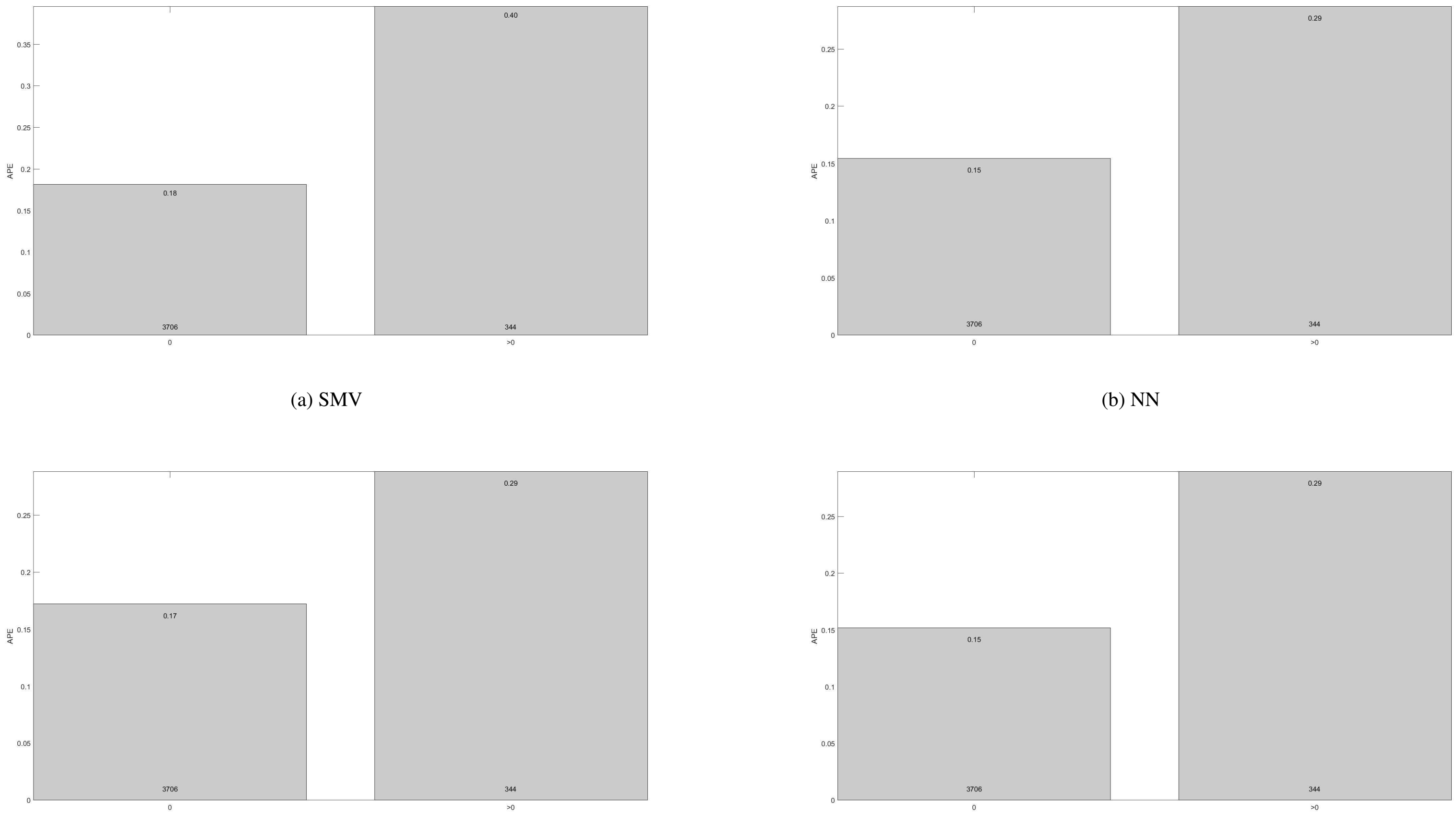


(c) MRA



(d) SMV+NN+MRA

Figure 5: Mean absolute percentage error per residence area (V08).



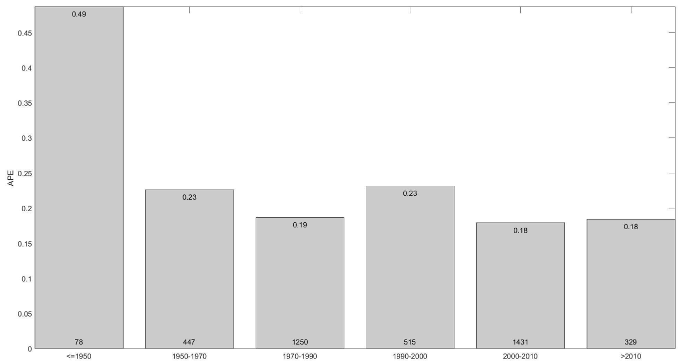
(a) SMV

(b) NN

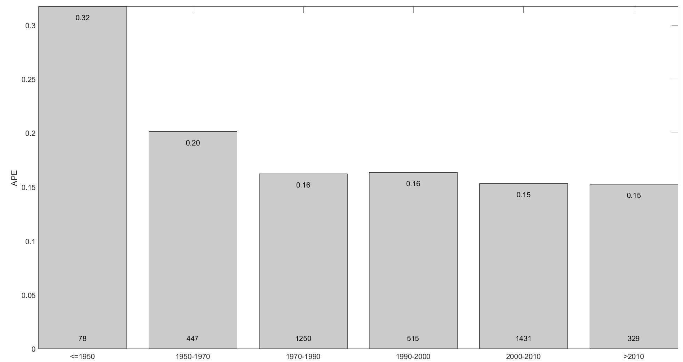
(c) MRA

(d) SMV+NN+MRA

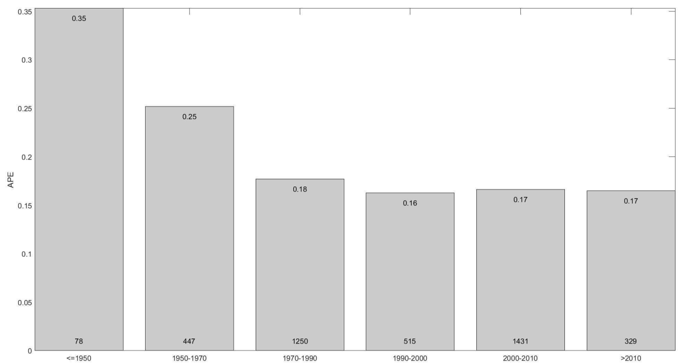
Figure 6: Mean absolute percentage error per land area (V09).



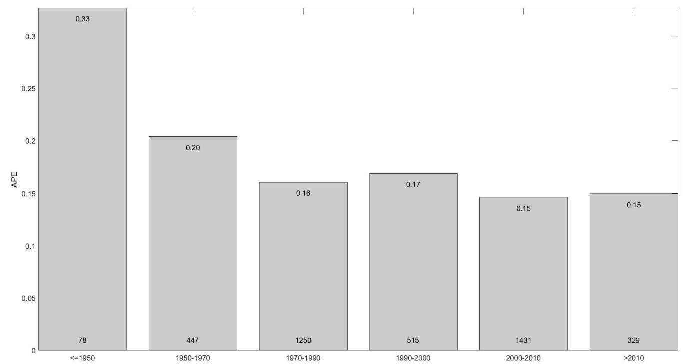
(a) SMV



(b) NN

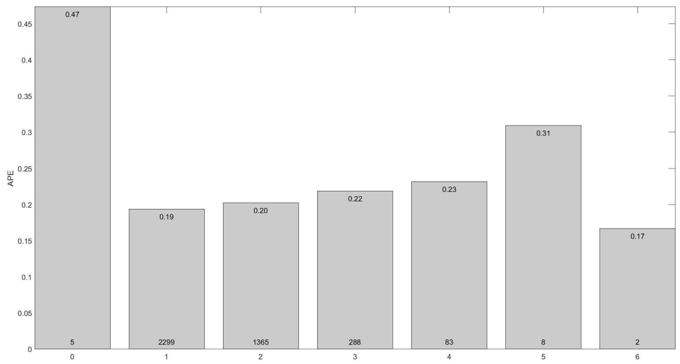


(c) MRA

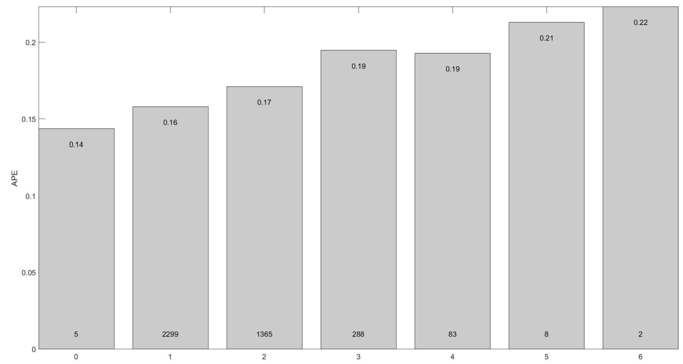


(d) SMV+NN+MRA

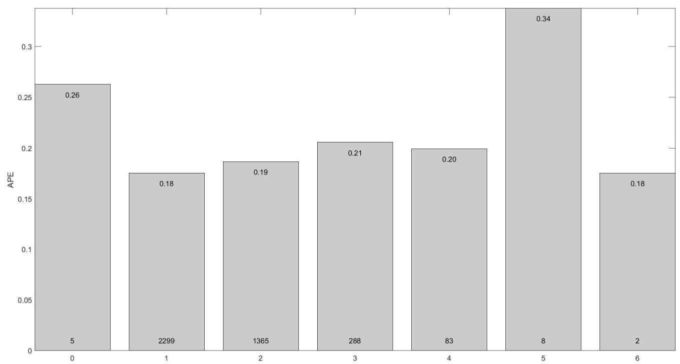
Figure 7: Mean absolute percentage error per year of construction (V10).



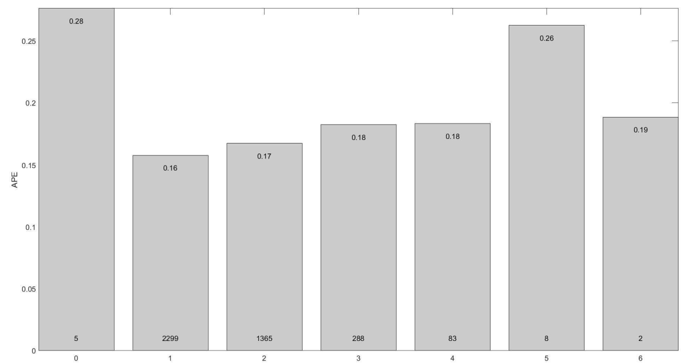
(a) SMV



(b) NN



(c) MRA



(d) SMV+NN+MRA

Figure 8: Mean absolute percentage error per number of bedrooms (V22).

5. Conclusions

In this study we developed three mass appraisal systems for the automatic valuation of real estate properties in Greece. The Greek property market is a new market still at its infancy with lots of unique characteristics. It is an inefficient, non-homogeneous market governed by lack of information. Also, there are declining prices due to the recession while the properties' characteristics are diverse both at regional and country level. We formulate and compare linear and non-linear models based on regression, hedonic equations, spatial analysis and artificial neural networks.

We perform an extensive out-of-sample analysis in four non-overlapping data sets. In contrast to previous studies, our results indicate that NNs constantly outperform traditional valuation methods. In this study the proposed NN was fine tuned and extra care was taken to avoid overfitting. The MRA method ranks second while the SMV method ranks third. The forecasting accuracy can be further improved by employing averaging techniques. A simple average of the three methods performs as well as, and in some cases outperforms, the NN.

Finally, we try to identify the property characteristics that lead to large forecasting errors. Our results indicate that the forecasting error increases when the residence area is above $120m^2$ or the property is a house or large land area is included. Similarly, very old properties (build before 1950) lead to larger forecasting errors. However, is it worth to mention that our analysis revealed that NNs are less sensitive to the changes of these characteristics compared to the SMV or the MRA. Our results indicate that the proposed methodology constitutes an accurate tool for property valuation in non- homogeneous, newly developed markets.

The proposed AVM can be adapted in applications such as mortgage quality control or appraisal review, loss mitigation analysis, portfolio valuation and appraisal process redesign.

References

Ahn, J. J., Byun, H. W., Oh, K. J., Kim, T. Y., 2012. Using ridge regression with genetic algorithm to enhance real estate appraisal forecasting. Expert Systems with Applications 39 (9), 8369–8379.

Alexandridis, A. K., Zapranis, A. D., 2013. Wavelet neural networks: A practical guide. Neural Networks 42, 1–27.
URL `http://www.sciencedirect.com/science/article/pii/S0893608013000129`

Alexandridis, A. K., Zapranis, A. D., 2014. Wavelet Networks: Methodologies and Applications in Financial Engineering, Classification and Chaos. Wiley, New Jersey, USA.

Anders, U., Korn, O., 1999. Model selection in neural networks. Neural Networks 12 (2), 309–323, 177DH Times Cited:81 Cited References Count:37.

Antipov, E. A., Pokryshevskaya, E. B., 2012. Mass appraisal of residential apartments: An application of random forest for valuation and a cart-based approach for model diagnostics. Expert Systems with Applications 39 (2), 1772–1778.

Aragonés-Beltrán, P., Aznar, J., Ferrís-Oñate, J., García-Melón, M., 2008. Valuation of urban industrial land: An analytic network process approach. European Journal of Operational Research 185 (1), 322–339.

Atkinson, S. E., Crocker, T. D., 1987. A bayesian approach to assessing the robustness of hedonic property value studies. Journal of Applied Econometrics 2 (1), 27–45.

Aznar, J., Guijarro, F., 2007. Estimating regression parameters with imprecise input data in an appraisal context. European Journal of Operational Research 176 (3), 1896–1907.

Ball, J., Srinivasan, V. C., 1994. Using the analytic hierarchy process in house selection. The Journal of Real Estate Finance and Economics 9 (1), 69–85.

Bitter, C., Mulligan, G. F., Dallerba, S., 2007. Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method. Journal of Geographical Systems 9 (1), 7–27.

Brasington, D. M., Hite, D., 2008. A mixed index approach to identifying hedonic price models. Regional Science and Urban Economics 38 (3), 271–284.

Brown, K., Uyar, B., 2004. A hierarchical linear model approach for assessing the effects of house and neighborhood characteristics on housing prices. Journal of Real Estate Practice and Education 7 (1), 15–24.

Brunauer, W., Lang, S., Umlauf, N., 2013. Modelling house prices using multilevel structured additive regression. Statistical Modelling 13 (2), 95–123.

Caples, S. C., Hanna, M. E., Premeaux, S. R., 1997. Least squares versus least absolute value in real estate appraisals. Appraisal Journal 65, 18–24.

Chica-Olmo, J., 2009. Prediction of housing location price by a multivariate spatial method cokriging. Journal of Real Estate Research 29 (1), 91–114.

Coleman, J. W., Larsen, J. E., 1991. Alternative estimation techniques for linear appraisal models. The Appraisal Journal 59 (4), 526–32.

Cybenko, G., 1989. Approximation by superpositions of a sigmoidal function. Mathematics of Control, Signals and Systems 2 (4), 303–314.

Dan Foresee, F., Hagan, M., 1997. Gauss-newton approximation to bayesian learning. In: International Conference on Neural Networks. Vol. 3. pp. 1930–1935.

Dimopoulos, Y., Bourret, P., Lek, S., 1995. Use of some sensitivity criteria for choosing networks with good generalization ability. Neural Processing Letters 2 (6), 1–4, tj374 Times Cited:33 Cited References Count:7.

Guan, J., Zurada, J., Levitan, A., 2009. An adaptive neuro-fuzzy inference system based approach to real estate property assessment. Journal of Real Estate Research 30 (4), 395–420.

Helbich, M., Brunauer, W., Vaz, E., Nijkamp, P., 2014. Spatial heterogeneity in hedonic house price models: the case of austria. Urban Studies 51 (2), 390–411.

Hoesli, M., MacGregor, B. D., 2000. Property investment: principles and practice of portfolio management. Pearson Education.

Isakson, H. R., 2001. Using multiple regression analysis in real estate appraisal excerpt from papers and proceedings. Appraisal Journal 69 (4), 424–430.

Janssen, C., Soderberg, B., Zhou, J., 2001. Robust estimation of hedonic models of price and income for investment property. Journal of Property Investment & Finance 19 (4), 342–360.

Kaklauskas, A., Zavadskas, E. K., Banaitis, A., Satkauskas, G., 2007. Defining the utility and market value of a real estate: a multiple criteria approach. International Journal of Strategic Property Management 11 (2), 107–120.

Kauko, T., d'Amato, M., 2009. Mass appraisal methods: An international perspective for property valuers. John Wiley & Sons.

Kilpatrick, J., 2011. Expert systems and mass appraisal. Journal of Property Investment & Finance 29 (4/5), 529–550.

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

Kummerow, M., 2000. Thinking statistically about valuations. Appraisal Journal 68 (3), 318–326.

Kusan, H., Aytekin, O., Ozdemir, I., 2010. The use of fuzzy logic in predicting house selling price. Expert systems with Applications 37 (3), 1808–1813.

Landajo, M., Bilbao, C., Bilbao, A., 2012. Nonparametric neural network modeling of hedonic prices in the housing market. Empirical Economics 42 (3), 987–1009.

Lins, M. P. E., Novaes, L. F. d. L., Legey, L. F. L., 2005. Real estate appraisal: A double perspective data envelopment analysis approach. Annals of Operations Research 138 (1), 79–96.

MacKay, D. J., 1992. Bayesian interpolation. Neural computation 4 (3), 415–447.

Mark, J., Goldberg, M. A., 1988. Multiple regression analysis and mass assessment: A review of the issues. Appraisal Journal 56 (1).

McCluskey, W., McCord, M., Davis, P., Haran, M., McIlhatton, D., 2013. Prediction accuracy in mass appraisal: a comparison of modern approaches. Journal of Property Research 30 (4), 239–265.

Narula, S. C., Wellington, J. F., Lewis, S. A., 2012. Valuating residential real estate using parametric programming. European journal of operational research 217 (1), 120–128.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., French, N., 2003. Real estate appraisal: a review of valuation methods. Journal of Property Investment & Finance 21 (4), 383–401.

Papastamos, D., Matysiak, G., Stevenson, S., 2015. Assessing the accuracy and dispersion of real estate investment forecasts. International Review of Financial Analysis 42, 141–152.

Park, B., Bae, J. K., 2015. Using machine learning algorithms for housing price prediction: The case of fairfax county, virginia housing data. Expert Systems with Applications 42 (6), 2928–2934.

Peterson, S., Flanagan, A., 2009. Neural network hedonic pricing models in mass real estate appraisal. Journal of Real Estate Research.

Reichert, A. K., 1990. The impact of interest rates, income, and employment upon regional housing prices. The Journal of Real Estate Finance and Economics 3 (4), 373–391.

Rossini, P., Kershaw, P., January 2008 2008. Automated valuation model accuracy: some empirical testing.

Rumelhart, D. E., McClelland, J. L., of California San Diego. PDP Research Group., U., 1986. Parallel distributed processing : explorations in the microstructure of cognition. Computational models of cognition and perception. MIT Press, Cambridge, Mass., 85024073 David E. Rumelhart, James L. McClelland, and the PDP Research Group. 24 cm. Vol. 2 by James L. McClelland, David E. Rumelhart, and the PDP Research Group. "A Bradford book." Bibliography: p. [507]-516. Includes index. v. 1. Foundations – v. 2. Psychological and biological models.

Samarasinghe, S., 2006. Neural Networks for Applied Sciences and Engineering. Taylor & Francis Group, New York, USA.

Schulz, R., Wersing, M., Werwatz, A., 2014. Automated valuation modelling: a specification exercise. Journal of Property Research 31 (2), 131–153.

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

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

Zapranis, A., Refenes, A. P., 1999. Principles of Neural Model Indentification, Selection and Adequacy: With Applications to Financial Econometrics. Springer-Verlag.

Zurada, J., Levitan, A., Guan, J., 2011. A comparison of regression and artificial intelligence methods in a mass appraisal context. Journal of Real Estate Research.