**Neural networks support vector machine for mass appraisal of properties**

Yacim, Joseph Awoamim[1] Boshoff, Douw Gert Brand[2]

[1]Department of Construction Economics,
Faculty of Engineering, Built Environment and Information Technology,
University of Pretoria, Hatfield 0028, South Africa; or Dept of Estate Management and Valuation,
School of Environmental Studies, Federal Polytechnic, Nasarawa, Nasarawa 962101, Nigeria
[2]Urban Real Estate Research Unit, Department of Construction Economics and Management,
University of Cape Town, Private Bag X3, Rondebosch 7701, South Africa;

**Abstract**

The paper introduced the use of a hybrid system of neural networks support vector machines (NNSVMs) consisting of artificial neural networks (ANNs) and support vector machines (SVMs) to price single-family properties.

The mechanism of the hybrid system is such that its output is given by the SVMs which utilise the results of the ANNs as their input. The results are compared to other property pricing modelling techniques including the standalone ANNs, SVMs, geographically weighted regression (GWR), spatial error model (SEM), spatial lag model (SLM) and the ordinary least squares (OLS). The techniques were applied to a dataset of 3225 properties sold during the period, January 2012 to May 2014 in Cape Town, South Africa.

The results demonstrate that the hybrid system performed better than ANNs, SVMs and the OLS. However, in comparison to the spatial models (GWR, SEM and SLM) the hybrid system performed abysmally under with the SEM favoured as the best pricing technique.

The findings extend the debate in the body of knowledge that the results of the OLS can significantly be improved through the use of spatial models that correct bias estimates and vary prices across the different property locations. Additionally, utilising the result of the hybrid system is thus affected by the black-box nature of the ANNs and SVMs limiting its use to purposes of checks on estimates predicted by the regression-based models.

Keywords: Artificial intelligence, Hedonic regression models, Pricing of properties

## 1. Introduction

This study presents a comparative analysis of methods for the pricing of residential properties. The main motivation is to discover a suitable technique from the arrays of pricing methods in terms of their predictive accuracy, consistency, transparency of estimates and ease of usage within the mass appraisal setting. There exist a plethora of research carried out with the global models (e.g. hedonic regression models based on the OLS estimator (HRMs-OLS)), global model with spatial autoregressive term (e.g. simultaneous autoregressive), local models (e.g. geographically weighted regression (GWR)), artificial neural networks (ANNs) and the support vector machines (SVMs) but to date there has been no comparative study that investigates the performance of these techniques. Again as Bourassa, Cantoni and Hoesli (2010) observe, the techniques used to arrive at optimal results in previous studies cannot be implemented in other climes or regions because of differences in the data. Therefore this study is undertaken to contribute to the growing literature on modelling property prices using the perspective of the Cape Town property market.

There exist some criteria designed to assess methods suitability in mass appraisal valuation. Kryvobokov (2004) suggests five criteria for assessing model suitability in Ukraine including (i) clearness of method; (ii) measurement of the outcome; (iii) importance of the result; (iv) market orientation; and (v) simplicity rather than the precision of the technique. These were expanded in d' Amato and Kauko (2008) to include the following groups and subgroups of criteria: a. *institutional criteria,* (i) suitability of the method to specific market context; (ii) specific path-dependence; b.

*methodological criteria,* (i) precision of independent assessments; external or out-sample validation; (ii) conceptual reliability; (iii) analysis of valuation variance; (iv) internal checks and reliability of the model structure; (v) nature of the adjustment; (vi) strength of the model; and (vii) practicability.

However, whilst it is imperative to assess method suitability using the criteria above, the study of McCluskey (1997); McCluskey, Davis, Haran, McCord and Mcllhatton (2012) and McCluskey, McCord, Davis, Haran, and McIlhatton (2013) added the explicit explainability of the method so that appraisers can easily defend estimates before a tribunal or in a formal court. The ability of a model to provide detail explanation of the appraisal process is germane hence the latter criterion is seen as fundamental to the assessment of model suitability. Traditionally, the global hedonic regression models (HRMs) based on the ordinary least squares (OLS) is used for pricing of residential properties. The OLS is simple and suitable to a straightforward assessment of the association between a response variable and the different explanatory variables but problems such as nonlinearity, functionality, spatial dependence and spatial heterogeneity among attributes and observations ensue if the appraisal work incorporates more aspects (Kauko, 2003).

The limitations of OLS triggered the development of several alternative pricing techniques. For instance to capture spatial effects and allow variation of coefficient across the geographic space the GWR and spatial expansion method are among others used, similarly, the spatially lagged (SLM), the spatially mixed (SAM) and the spatial error models (SEM) are (three models in the simultaneous autoregressive (SAR) that signify areas the spatial autoregressive process is said to occur (Dormann *et al*., 2007; Kissling and Carl, 2008)) used to deal with autocorrelation in the property data. The GWR and SAR extend the OLS to work differently in addressing spatial effects. Whereas the GWR creates a distinct regression model for every sale point (property location) and assign weights to observations relative to their distance to the regression point thereby permitting single marginal-price estimates at each location (Bitter, Mulligan and Dall'erba (2007), the SAR includes an additional term that contains the spatial autocorrelation structure in a given data. Kissling and Carl (2008) report that the additional term is applied with a spatial weight matrix where the neighbourhood of each location and the weight of each neighbour need to be defined (Anselin and Bera, 1998; Fortin and Dale, 2005).

Moving away from the regression-based methods is the interest in the use of artificial intelligence-based techniques, particularly the ANNs and SVMs. The interest on these models (ANNs and SVMs) is borne out of their high computing abilities to recognise patterns in the property data and effectively model property prices devoid of the many parametric restrictions of the OLS. While the SVMs and ANNs have a good history of optimal performance in the mass appraisal industry, they, however, suffer from some limitations. The study of Wiering *et al*. (2013) identified over-fitting and over-training particularly on small datasets as limitations of ANNs which is not a problem with the SVMs. Though the SVMs generalise well, they have also faced with the problem of rigidity in the selection of the appropriate kernel function[1]and inability to handle many outputs in a single-layer architecture inhibiting their use in learning a task, particularly dimensionality reduction[2] (Wiering *et al*., 2013). This study proposes the use of a hybrid system comprising of ANNs and SVMs similar to the one used in Wiering *et al*. (2013). Whereas, the hybrid system in their study was used in eye images experiment relative to standalone ANNs and SVMs, this study utilises the hybrid system in mass appraisal of properties, with an example data from Cape Town, South Africa.

Relative to the South African property market, the OLS technique is commonly used for pricing of properties to levy taxes in Cape Town and other municipalities, but cases of objections leading to appeals as a result of inflated tax amounts had been reported by property owners (KPMG, 2015; LexisNexis, 2018). Thus, discovering a suitable pricing technique that is appropriate for providing estimates that are closer to market value or price, from the arrays of methods utilised in this study is the main motivation. The remainder of the study is structured into five sections including section two providing the literature review, section three contains the methodology and modelling procedures, which is followed by empirical analysis in section four. Section five concludes the study.

## 2. Literature review

The development of property pricing could not be complete without the mention of the hedonic OLS modelling. This is because any pricing model despite its capabilities in modelling prices stemmed from the framework provided in the OLS. The development of the OLS could be traced to the work of Court (1939) that utilised it in the automobile industry and also Lancaster (1966) who relates it to the unit of measurement used to quantify the characteristics possessed by consumer goods from which user(s) derived satisfaction (utility). However, it was the study of Rosen (1974) that provided the framework for measuring the different contribution of individual attributes to the overall price of a property. The OLS assesses property prices based on the physical/structural and locational attributes. Accordingly, Goodman (1978) used a database of 1835 single family property transaction in the New Haven standard metropolitan statistical area (SMSA) to form property price indices with the OLS. Mok, Chan and Cho (1995) used the OLS to estimate the values of 1,027 property sales in Hong Kong. Raymond and Peter (2000) used the hedonic regression on 139 property data to estimate prices in Hong Kong. Yang (2001) used the techniques to estimate the prices of 226 apartments sold in Beijing, China.

Stevenson (2004) used the OLS on 6,441 property data to estimate property prices in Boston and found that at a disaggregated level the model may provide better results. Also, Shimizu (2014) used the OLS on a dataset containing 13,822 single-family dwellings to estimate property prices in Tokyo, Japan and found it to reduce bias when neighbourhood effects are added to the model. McCluskey (2016) used the OLS on a dataset of 40,138 apartments in Kazakhstan and found the model to be useful in estimating property prices for taxation purpose. Hayrullahoğlu, Aliefedioğlu, Tanrivermis and Hayrullahoğlu (2018) used the OLS to price 163 properties in the Çankara, Çukurambar district of Ankara, Turkey and found the stepwise regression to perform better in comparison to other price modelling techniques. Yacim and Boshoff (2018a) used 3242 sales transaction in Cape Town to estimate the price of properties and found that although the OLS did not perform well in comparison to other price modelling techniques, the technique is still relevant in property valuation.

Although the OLS has been in use for pricing property, it is beset with many practical limitations that stem from its parametric rigidity, choice of functional form to use and inability to adequately handle spatial effects. Unfortunately, the economic theory fails to stipulate a particular form to use in relating property price to its characteristics. This led to high reliance on the use of goodness of fit criterion in selecting appropriate functional form by researchers (Crooper, Deck and McConnel, 1988). Yang (2001) report that the Box-Cox transformation is extensively utilised in property-related studies, linear and other logarithmic forms are also used (Henry, Patrick and Yiu-Sun, 1995). However, Borst (2007) notes that several criticisms followed the use of Box-Cox function leading to a situation where some authors directly formulate a model structure without reference to the hedonic function. This study utilises the linear form of the OLS to price property. This is because the linear OLS model is said to sufficiently describe the relationship between attributes more clearly and the results are easily compared to the expert valuations (Kryvobokov and Wihelmsson, 2007). Nonetheless, spatial effects are not sufficiently catered for in the OLS (Páez, Uchida and Miyamoto, 2001; Bitter, Mulligan and Dall'erba, 2007), leading to bias coefficients (Mueller and Loomis, 2008). Spatial effects (spatial dependence and spatial heterogeneity) are twin glitches that are inherently present in property markets; hence, methods of dealing with these will among other issues be the focus of this study.

Des Rosier and Thériault (2008) observed that in the OLS, the contextual variation over space is generally specified with "fixed" coefficients obtained from locational dummies to measure their direct effect on property prices. However, this stable price assumption is not possible where the markets are heterogeneous making it imperative to account for spatial effects within the regression framework (Theriault, Des Rosier, Villeneuve and Kestens, 2003). To tackle spatial heterogeneity, the spatial expansion method pioneered by Casseti (1972; 1997), and GWR are among others used. Accordingly the study of Kestens, Thériault and Des Rosier (2006) using transaction data of 761 single-family dwellings in Quebec City, Canada, employed the spatial expansion method and GWR and found that the two methods provide good results relative to the marginal value given to property and locational attributes based on the preferences of the buyers' household. Farber and Yeates (2006) used the OLS and compare performance with GWR, spatial lag model and moving window regression on a data of 19,007 freehold housing sales in Toronto, Canada and found the GWR to be less spatially biased, account best for the spatial variation in prices but did not support its adoption by assessment community because of limitations in the statistical framework used.

Bitter, *et al*. (2007) used a spatial expansion method and GWR to account for the spatial heterogeneity and prediction of property price in Tucson. The study used the transaction data of 11,732 single-family residences and found that both methods improve the results of stationary coefficient models but in terms of explanatory power and predictive precision the GWR performs better. Páez, Long and Farber (2008) using an estimation sample of 30,145 and a validation sample of 3,349 observations in the city of Toronto, Canada tests the predictive power of moving windows regression (MVR), GWR, kriging and moving windows kriging and found that in terms of predictive power the GWR marginally performed better than the MVR but substantially outperformed the other models. Also, McCluskey and Borst (2011) used the GWR with datasets from Catawba, Sarasota and Fairfax Counties in the USA to identify market segments and conclude that the resultant segments have the propensity of improving predictive accuracy and lowering spatial autocorrelation in the residual errors. Again, the study of McCluskey *et al*. (2013) used the GWR to assess property prices in Northern Ireland, UK. With a sample of 2,694 residential properties, the GWR was found to outperform the other models used for comparison in terms of cost-efficiency, ease of use, and predictive accuracy. Yacim and Boshoff (2019) used 3232 sales data in Cape Town and applied the semi-log OLS and GWR to control for spatial heterogeneity among observations. The study found a significant improvement in the performance of GWR relative to the OLS techniques despite the inclusion of geo-coordinates among input variables.

According to Valente, Wu, Gelfand and Sirmans (2005) spatial effects (dependence) are being modelled by spatial conditional autoregressive (CAR), spatial simultaneous autoregressive (SAR) and kriging (Pace, Barry and Sirmans, 1998) modelling. Spatial dependence or autocorrelation occurs when there is interdependence among observations in a geographical space that violates the assumption of uncorrelated error terms (Osland, 2010). The SAR has been used to account for spatial autocorrelation in several studies. For example, the study of Pace and Gilley (1997) used 506 housing data points from Boston SMSA and found SAR to outperform the OLS. Dubin, Pace and Thibodeau (1999) used a small sample (10) to compare the regression coefficients of OLS and four spatial techniques including SAR, CAR, mixed regressive spatially autoregressive and Gaussian correlogram. The study used the property selling price as a function of the square footage of living area and location and found that apart from the mixed regressive spatially autoregressive model, other spatial models yield parameter estimates that are slightly closer to the true values than the OLS. Wilhelmsson (2002) used 1,377 single-family property transactions in the municipality of Stockholm, Sweden and found the autoregressive model to outperform the OLS. Militino, Ugarte and Garcia-Reinaldos (2004) used the OLS, lattice, geostatistical models on a 293 property dataset in Pamplona, Spain and found that the lattice models (SAR and CAR) coped with the spatial dependence and gives a robust inference similar to other spatial techniques used. Bourassa, Cantoni and Hoesli (2007) used a dataset of 4880 residential property sales in Auckland, New Zealand and found the lattice models (SAR and CAR) to offer less prediction accuracy than the OLS and geostatistical (exponential and robust exponential variogram) models. Again, McCluskey *et al*. (2013) used SAR in Northern Ireland, UK and found it to perform well.

The reviewed studies show that although the OLS is widely accepted for use in property-related analysis, it has many weaknesses which have since been identified and addressed through various improvements. This current study extends the use of techniques that enhanced pricing of properties in the South African market because such improvement, particularly, with spatial dependence modelling techniques do not exist. Yacim and Boshoff (2019) is the known study that provided controls for spatial heterogeneity with an example of data from South Africa. However, the suggestion in de Graaff, Florax, Nijkamp and Reggiani (2001) that techniques dealing with the two spatial effects must be utilised in an assessment to permanently correct the glitches, among others, led to the current study.

Other studies favoured the application of artificial intelligence-based techniques into the mass appraisal environment (Do and Grudnitski, 1992; Tay and Ho, 1992). Several models fell under artificial intelligence techniques; however, this study will concentrate on the ANNs and the SVMs. The SVMs is among the techniques that have in the last decade been introduced into the mass appraisal environment, with little research undertaken on it. To our understanding, the SVMs was first mentioned as a technique for mass appraisal in Lam, Yu and Lam (2009). The study used 4,143 and 21 property transactions in Hong Kong and Nanjing, Mainland China, respectively and found the SVMs to outperform ANNs and OLS in factor weighting and predictive accuracy. Again, Zurada, Levitan and Guan (2011) used the SVMs for mass appraisal in the USA and found it to be among the best performing models. The SVMs used different kernel function to construct a separating hyperplane in high dimension feature space without plainly performing the computation in the feature space (Zurada *et al.,* 2011), but the choice of kernel function is not a straight forward procedure. Cui and Curry (2005) reported that there is no comprehensive meta-theory that serves as a guide in the selection of kernel functions for SVMs. Therefore a high reliance on trial and error process is prevalent until a kernel that provides optimal result is guaranteed. The trial and error procedure implies that a kernel that works well with a particular class of data or market context might not work well with another. The choice of a kernel function is seen in the studies of Lam *et al*. (2009) and Zurada *et al*. (2011). Whereas the study of Zurada *et al.* (2011) used a polynomial kernel function (PLK) for their analysis, Lam *et al.* (2009) utilised radial basis kernel function (RBF) in their analysis. This study used the PLK kernel function in estimating property prices.

The ANNs have been applied in property pricing since the early 1990s in many studies, but with variegated results (McCluskey *et al*., 2012). While several studies establish the model as a valuable tool for assessment of property prices others do not share this view. Again most studies used the OLS as a baseline for comparison with the ANN models. For instance studies of Tay and Ho (1992) in Singapore; Do and Grudnitski (1992), Borst (1995), Nguyen and Cripps (2001) and Peterson and Flanagan (2009) in the USA; Evans *et al*. (1992) in the UK; and Limsombunchai, Gan and Lee (2004) in New Zealand all found the ANNs to outperform the OLS but studies of Worzala, Lenk and Silva (1995); Lenk, Worzala and Silva (1997) in the USA and McGreal, Adair, McBurney and Patterson (1998) in the UK found inconsistent results that do not show clear superiority of ANNs over the OLS and thus cautioned the appraisal community against its use in property pricing.

Other studies used more than two models for comparison. These studies compare the performance of ANNs against other techniques. Specifically, the study of Zurada *et al*. (2011) compares the performance of the OLS, ANNs, additive regression (AR), M5P trees, SVMs, radial basis function neural networks (RBNN) and memory-based reasoning (MBR) in the USA. The study found that non-traditional regression-based models (AR, M5P trees and SVM) performed better in all of the five simulated experiments, mostly with homogenous data, while artificial intelligence (AI) based models (ANNs, RBNN and MBR) performed well with less homogenous datasets. Lin and Mohan (2011) compare the performance of the OLS, ANNs and additive nonparametric regression (ANR) in the USA and found the ANNs to perform better than the OLS and ANR. McCluskey *et al*. (2012) and McCluskey *et al*. (2013) undertook two studies in the UK. While the former investigated the performance of ANNs against three regression-based models (linear, semi-log and log-log) and found the regression models to outperform the ANNs, the latter investigated the performance of the ANNs against the OLS, SAR and GWR and found the GWR to achieve better results than other models.

In the last one to two decades, studies have tended towards building a hybrid system to enhance estimation results relative to standalone model assessments. Such a combination is not new to the field of artificial intelligence. For instance, Kilpatrick (2011) and McCluskey and Anand (1999) observed that hybrid systems evolved to draw strength from different individual techniques. In mass appraisal, McCluskey and Anand (1999) construct a hybrid system of GA and backpropagation trained ANNs and used the *K*- nearest neighbour (*k*-NN) algorithm to select comparable properties and provide transparency, creditability and explainability to the appraisal process. Gonzalez and Formoso (2006) built fuzzy rules based on gross building area (GBA) and location in Porto Alegre, Brazil. In building the fuzzy rule of the former, each membership function is defined using the limits of GBA while the genetic algorithm estimates the fuzzy rules. The fuzzy rule-based on location was constructed similarly but have differences in the procedures utilised. The rules are determined in a specialised manner according to the region of the city but all contribute to the final estimates. Similarly, Guan, Zurada and Levitan (2008) combined fuzzy systems with neural networks to assess property prices in the Midwest region of the USA. While the fuzzy system was used in generating fuzzy rules and parameters for the membership function, the neural network was used to fine-tune the fuzzy rules and found results that are comparable to the regression model. The proposed hybrid system is expected to address, among others, dimensionality problem that is inherent in property price analysis, especially, when property attributes are many and perceived to be value significant.

## 3. Data and methodological approaches

### 3.1 The data

The city valuation office (CVO), Cape Town, South Africa provided the database of 3526 properties with 46 property variables and features. The variables were reduced to 11 including the *x, y* coordinates to ease out the problem of multicollinearity among variables after a sequence of initial regression tests that disclose them to be suitable for this analysis (Table 1). Within the property sample, is variables such as STRAP (property identification number) and use code (single dwelling), which merely describe the property as such were excluded from the analysis. Others such as topography, shape, double volume, attic area, office shop, pergola, tennis court, detached sauna, wooden deck, squash court, etc. were sparsely found in few observations in the sample and thus have insignificant contributions. The neighbourhood and submarket codes found in the sample were used to form location dummies for models devoid of the *x*, *y* coordinates platform. The use of regression technique in the selection of property variables is in line with previous multivariate studies of Bitter *et al*. (2007), Borst (2007), and Páez *et al*. (2008). The *x, y* coordinates are longitude and latitude of each property used to calculate distances between properties in the spatial models. The number of transactions was reduced to 3225 after the removal of outliers, extreme and unrealistic transactions. The dataset used is considered satisfactory for the analysis since the effectiveness of the techniques is not in any case measured by the number of observations (McCluskey and Anand, 1999).

Furthermore, text data were converted to numeric values to aid in the assessment. There are four texts (non-numeric) data used by tax assessor to grade the state of the properties in the study area. These are quality, condition, building style and property views. For example, property quality has five categories namely poor, fair, average, good, very good and excellent with poor assigned the 1 value and excellent assigned the 5 numeric values. The same procedure was done for all other non-numeric variables. The age of construction was not included in the original sample, and thus was not a part of this analysis.

**Table 1.** Property variables

| Variable | Description |
|---|---|
| TASP | Time adjusted sale price of properties measured in Rand (1$USD is ZAR15.35) |
| Beds | Number of bedrooms |
| Condition | Physical state/condition of property |
| Property size | Size of property in square metres |
| Pool | Size of a swimming pool in square metres |
| Quality | The quality grade of construction |
| Storey | Total number of storeys |
| Style | Building architecture style & design |
| Submarket | The locational variable identifying area the property is found |
| View | The type of property view e.g. panoramic |

Depending on the technique used, the data was stratified into 70% modelling (training) and 30% testing sets, respectively. For instance, when modelling with the ANNs, SVMs and the hybrid system, the stratification applies, but when modelling with the GWR, OLS, spatial lag and spatial error models the 100% data was used. The ANNs, SVMs, hybrid system and OLS models used the locational indicator variables to capture 15 submarkets within the city of Cape Town while the spatial models used the discrete explicit *x, y* coordinates. The submarkets used are the creation of the Tax Assessors (TA) in Cape Town. The TA clustered neighbourhoods that are at proximity and located in the same administrative district to form submarkets. The TA generally looks at geographic proximity and average or median sale prices per neighbourhood as the basis of creating submarket areas.

The submarkets are used because representing property location with a total of 181 dummies (for neighbourhoods) as input variables that comprised the 15 segments will result in having a large number of inputs (dummy variables). Zurada *et al.* (2011) observe that the problem of having too many variables as inputs (dummy variables) will increase dimensionality thereby reducing the strength of the models. The submarkets used might not be consistent with actual submarkets in reality (Wilhelmsson, 2002), but mere clustering together of neighbourhoods with homogeneous sales. The study of Borst (2007) defined submarkets and neighbourhoods based on theory to include among others “that submarket comprised of one or more neighbourhoods” and “neighbourhoods are close areal units consisting of one or more properties”. The use of submarkets to account for location is consistent with the study of Wilhelmsson (2002) who used 13 submarket dummy variables from the previously defined administrative parish and Bourassa *et al*. (2007) who utilised 33 submarket dummies to account for location in their study.

To account for time trend, the procedure suggested by Gloudemans (1990); Gloudemans (1999) and adopted in McCluskey *et al* (2013) was used. The procedure involves using a regression approach whereby sale or assessment date is included as an independent variable and regressed against price or value and then dividing the coefficient of sale or assessment date by the mean value. The resultant value is an implied monthly growth rate index used for creating a new time adjusted dependent variable. The time-adjusted sale price (TASP) is arrived at by multiplying the index value (time adjusted factor) with the sale price or value within the sample.

**Table 2.** Descriptive statistics of variables

| Variable name | Mean | Medium | SD | Minimum | Maximum |
|---|---|---|---|---|---|
| TASP | 5378814 | 4320000 | 3741556 | 988800 | 45600000 |
| Beds | 3.56 | 3.00 | 0.991 | 1.00 | 10.0 |
| Condition | 3.51 | 3.00 | 0.627 | 1.00 | 5.00 |
| Pool | 14.0 | 0.00 | 18.4 | 0.00 | 154.0 |
| Property Size | 178.0 | 168.0 | 78.7 | 40.0 | 599.0 |
| Quality | 3.49 | 3.00 | 0.616 | 1.00 | 6.0 |
| Storey | 1.52 | 1.00 | 0.553 | 1.00 | 3.0 |
| Style | 3.03 | 3.00 | 0.430 | 1.00 | 7.0 |
| View | 3.58 | 4.00 | 0.963 | 1.00 | 6.0 |

Table 2 summarises the descriptive statistics of the property data. This illustrates the variability within the data. The mean TASP in the sample is R5,378,814.00 with a range of R988,800.00 to R45,600,000.00. The average property size is 177 square meters with a range of 40 square metres for the smallest and 599 square metres for the largest property assessed in the sample. The mean number of a storey in the sample is two and the average property has four bedrooms in the study area.

## 3.2 Methods and specification

### 3.2.1 Hedonic Regression Models

According to McCluskey *et al*. (2012); McCluskey *et al.* (2013) the hedonic OLS estimator is widely accepted as the most extensively applied techniques within the *ad valorem* valuation process. The models are used to reveal the marginal influence of diverse property attributes to price. Janssen and Söderberg (1999), report that the models provide the structure for the assessment of “differentiated goods like housing units whose individual features do not have observable market prices”. The framework provides for the individual property assessed/market values to be regressed on measures of their characteristics. McCluskey *et al*. (2012) noted that there are three significant elements which must be addressed for the model to be effective. These are (i) careful selection of response and explanatory variables; (ii) choice of functional form and lastly, (iii) the statistical relevance and contribution of the explanatory/independent variables to the model. Following this, the current study carefully utilises the TASP as the dependent variable and selected structural, temporal and locational attributes as the independent variables. On the choice of functional form, the economic theory does not give specification on the particular form this must take; thus the linear (OLS) form is used. The linear model assumes that property prices are defined by a vector of continuous and dummy variables (Raymond and Peter, 2000) given in equation 1

$$TASP = b_0 + b_1(beds) + b_2(condition) + b_3(pool) + b_4(size) + b_5(quality) + \\ b_6(storey) + b_7(style) + b_8(view) + b_9(submkt) + \varepsilon \tag{1}$$

where $b_0$ represents the regression constant, $\varepsilon$ is the error term or arbitrary component that reflects the unnoticed variation in values and $b_1$, $b_2$, …, $b_9$ are regression coefficients or values per unit assigned by the model to the independent variables. The TASP, swimming pool and property size are continuous variables while all other variables are dummy variables (depicting "1" if the categorical condition is met and "0" if otherwise) excluding the most occurring category in order to avoid the dummy variable trap that could trigger the problem of multicollinearity (Greene, 2003 and Borst, 2007). Again, apart from the locational dummy variable, all other specifications are the same for the spatial models (SEM and SLM). The study of Gloudemans (2002) observes that OLS in the linear additive form is commonly used in the appraisal of single-family dwellings because it is simple to calibrate and also reveals the contribution of each variable in monetary terms. This is what gives the model wide acceptability for used within the property tax environment because of the explicit and transparent manner it reveals the entire appraisal process and predictive accuracy.

However, as noted earlier, the presence of spatial dependence or autocorrelation and spatial heterogeneity or non-stationarity inherent in the property markets are twin glitches that inhibit the effectiveness of the OLS. A simple way of dealing with the spatial effects is to introduce indicator variables into the model but the difficulty arises when the number of indicators is large (Pace, Barry, Clapp and Rodriquez, 1998a). A large number of input indicator variables might affect the aspiration for parsimony and strength of the model resulting in bias coefficients (Pace, Barry and Sirmans 1998b; Mueller and Loomis, 2008), hence, the OLS do not sufficiently cater for spatial effects (Páez *et al*., 2001; Bitter *et al*., 2007). The study of Pace *et al*. (1998b) suggests that a quick way-out of using multiple indicator input variable is to create a two-dimensional smooth surface using the polynomial expansion approach in both the east-west & north-south coordinates of every observation's location. Though, this seems an easy way-out it is also beset with several shortcomings which Pace *et al*. (1998b) noted to include changing global fit based on local errors; imposing more smoothness than desired and addition of more polynomials could lead to multicollinearity. Some of the shortcomings could, however, be remediated through the use of splines that has the capability of modelling local errors without altering the global fit and controlling the desired level of smoothness by the user. Nonetheless, the current study will utilise other improved versions of the OLS within the HRM framework to incorporate spatial effects.

Fotheringham *et al*. (2015) reported that two versions, the global and local HRMs are used to account for spatial effects (spatial heterogeneity and spatial dependence) in property price prediction. Examples of the local models are the moving window regression, GWR and multilevel modelling (Fotheringham *et al*., 2015) while examples of the global models are spatial autoregressive and spatial moving average (Anselin, 2003). The techniques of interest in this study are GWR and simultaneous autoregressive models.

### 3.2.2 Geographically Weighted Regression

According to Brunsdon, Fotheringham and Charlton (1996), the GWR is a technique that has its root embedded in the framework of HRMs, which permit its estimates to vary on the specific location. The standard formulation is as follows

$$y_i = b_0(u_i, v_i) + \sum_c b_c(u_i, v_i)\, x_{ic} + \varepsilon_i, \qquad i = 1,2,\ldots,n, \tag{2}$$

where, $(u_i, v_i)$ are the coordinates of the *i*th point in space; $b_c(u_i, v_i)$ depicts a set of values of parameters at the *i*th point, $y_i$ is the dependent variable and $b_0(u_i, v_i)$ denotes the intercept value. This technique is more flexible than the OLS because of its non-reliance on the numerous underlying assumptions and rigidity of having a single model that capture relationships in the entire property market. GWR creates a distinct regression model at each observation point which permits the estimation of coefficients at every location (Brunsdon *et al*., 1996; Bitter *et al*. (2007). The method uses data points to identify properties that are sold within the area of the subject property and measure the distance between them. Thus the farther the property is from the subject (regression point) the lower the weights assigned. The study of Borst (2012) illustrates how GWR operates in assigning weights that vary with distance from the regression point. The study utilised the peak of the surface as the regression point, thus showing that any point below the surface is assigned weights relative to its position (height) of the surface at that point. Accordingly, Huang *et al*. (2010) and McCluskey *et al*. (2013) report that an estimate is achieved from observation relative to the distance between points. The parameter estimate $b_c(u_i, v_i)$ for this is given as

$$b(u_i, v_i) = [X^T W(u_i, v_i) X]^{-1} X^T W(u_i, v_i) Y \tag{3}$$

where the spatial weighting matrix is specified as $W(u_i, v_i)$. Huang *et al*. (2010) noted that two weighting regimes are used; these are fixed kernel and adaptive kernel. The fixed kernel has a varied number of nearest neighbours but the distance is constant. In the adaptive spatial kernel, it is the reverse phenomenon where there is variation in the distance but the number of nearest neighbour remains constant. The adaptive spatial kernel that permits the variation of bandwidth on the density of property sales around each regression point is used in this study. Bitter *et al*. (2007) report that adaptive kernel captures different segments of the property market including a smaller area with rich data and a larger area with sparse data. It also guarantees a proportionate sharing of non-zero weighting values among observations relative to the regression points.

There are two accepted adaptive weighting specifications; these are Gaussian function and bi-square (Griffith, 2008). A preliminary analysis undertaken to select an optimal kernel in this study reveals that the Gaussian function as used by Bitter *et al*. (2007), Griffith (2008) and McCluskey *et al*. (2013) provides an optimal result with the Cape Town property data. The Gaussian function is used to specify the Euclidian distance *d* between the regression and the observation points and *h* denoting the scalar quantity identified as the “bandwidth” given by

$$W_i(u_i, v_i) = \exp\left(\tfrac{-d}{h}\right)^2 \tag{4}$$

The estimated results of GWR are sensitive to the type of bandwidth used. Therefore care must be taken in its selection as weighting procedures that stipulate a wide bandwidth with minimal distance decay will yield results that are similar to a global model. Equally, a narrower bandwidth will result in high variances in the estimators because the only point close will be measured (Bitter, 2007). The procedures used in this study follow the suggestion in Nakaya *et al.* (2016). Three different selection procedures are used to search for an ideal bandwidth size including golden search, interval size and single bandwidth. The golden search is used to automatically search for the best bandwidth size. It is flexible and has the assurance of reaching optimal position without having to use different search values. The golden search also has an option for the user to limit search around a predefined range but this might sometimes terminate without reaching optimum. The interval search criterion allows user intuition and is more robust because of its regular interval within a predefined range. The last bandwidth selection search routine is the single bandwidth where a specific number is provided. This study utilised the golden search method to fix the optimal bandwidth size of 60 as a suitable value for the GWR model. In modelling property prices with GWR, the variable used were left in their continuous and categorical states. Similarly, apart from the *x, y* coordinates that reflect the location of properties in space and calculate distances for usage in the kernel function, all other variables including locational dummy are used in assessment with ANNs and the hybrid system. This is because ANNs and GWR are limited in handling excessive dummy variables (McCluskey *et al*., 2013 and Feng and Jones, 2015).

### 3.2.3 Simultaneous autoregressive models (SAR)

According to Dormann *et al*. (2007), SAR can take different forms depending on where the spatial autoregressive process is believed to occur (Cliff and Ord, 1981; Anselin, 1988; Haining, 2003). There are three areas by which the spatial autoregressive process is said to happen including, on the response variable (spatial lag), on both the explanatory and response variables (spatial mix) and the error term (spatial error).

#### 3.2.3.1 Spatially lagged response model

The first SAR assumes that the autoregressive process is believed to occur only in the response variable (i.e. spatially lagged response model) and thus take the term $(\rho W)$ for the autocorrelation in the response variable $Y$, but also the standard term for the independent variables and errors $(X\beta + \varepsilon)$ as used in the OLS models. In the spatially lagged model (SLM), the response variable is related to itself in a particular way (Borst, 2006). The SLM allows for the observed sale prices of nearby properties $j$ to influence the sale price (dependent variable) of property $i$ in the model. This kind of spatial interaction is what Can (1992) referred to as "adjacency effect" in which property valuers/agents used the price history of nearby properties and other characteristics to determine a price for the subject property. Therefore, Krause and Bitter (2012) reports that the spatially lagged model attempt to capture the spatial dependence in the property market or account for the influence of sales of properties in nearby locations on current property prices. The basic spatially lagged response model $(SAR_{lag}\ or\ SLM)$ is given in equation 5 as

$$Y = \rho WY + X\beta + \varepsilon \tag{5}$$

#### 3.2.3.2 Spatially mixed model

The second SAR assumed that spatial autocorrelation can affect both predictor (independent) variables and response variable leading to the addition of another term $(WX\gamma)$ into the model that describes the autoregression coefficients $(\gamma)$ of the spatially lagged independent (explanatory) variables $(WX)$. Brasington and Hite (2005), reports that the term $WX\gamma$ in equation 6, permits the structural characteristics (independent variables) of nearby properties to impact the price of each property. Consequently, any small or large-sized properties within the vicinity might be negatively impacted leading to a discounted sale occasioned by the market reaction. Again, reflecting on the study of Glower, Haurin and Hendershott (1998) that the degree to which a property is nonconforming influences its time on the market and sale price, Brasington and Hite (2005) suggest that it might be important to incorporate the structural characteristics of nearby properties into the model. The spatially mixed $(SAR_{mix})$ version is given in equation 6 as follows

$$Y = \rho WY + X\beta + WX\gamma + \varepsilon \tag{6}$$

#### 3.2.3.3 Spatial error model

Another approach to SAR modelling is the spatial error model $(SAR_{err}\ or\ SEM)$ which can be applied only if there is significant spatial autocorrelation. In the spatial error model, the autoregressive process is believed to occur in the error term and neither in the response variable nor in the explanatory (independent) variables. The study of Kissling and Carl (2008) noted that this is usually the case if autocorrelation is not explained by the explanatory variables or if the autocorrelation is an integral property of the response variable itself. For the $SAR_{err}$, the normal OLS model is added by a term ($\lambda Wu$) which denotes the spatial structure ($\lambda W$) in the spatial response error term ($u$). The $SAR_{err}$ or SEM takes the form of

$$Y = X\beta + \lambda Wu + \varepsilon \qquad (7)$$

The major weaknesses of SAR and its variants CAR as Wall (2004) observed are in the area of the definition of the weighting schemes structure.

3.2.3.4 Specification of the spatial weight matrix

Weights specification is an integral and crucial part of spatial hedonic regression models but there exists little scientific guidance on the processes. Elhorst (2010) report that one major bane of spatial econometric modelling is that it is difficult to estimate spatial weights matrix *W*, which require *a priori* specification. The lack of empirical guidance results in authors utilising arbitrary approach relative to some established benchmark tests to specify weight matrix which Anselin (2002) observe is the reason for its limitation. Because of this, it has become necessary to establish the robustness of the results relative to the specified weight (*W*). The spatial weight is an *N* x *N* matrix that models the relationship of neighbours for every observation within the sample as nonzero features. Accordingly, Feng and Humphreys (2008) report that every spatial weights matrix has a row, indexed by $i$ containing $W_{ij}$ elements and nonzero features with a define column *j* as a neighbour of $i$. Suppose $i$ and *j* are properties in a neighbourhood of which property *i* is a neighbour to property *j*, then $Wij$ = 1, conversely $Wij$ = 0 if properties *i* and *j* are not neighbours. Again, since property cannot be a neighbour to itself such that $W_{ii} = 0$, the weight matrix elements are conventionally set to zero. The weights matrix is usually row standardised for simplicity of comparison so that each row will have weights that sum to 1. Anselin (1988) observes that though there is no mathematical basis for this, it facilitates interpretation of the coefficients

In the *GeoDa* software designed by Anselin (2005), the methods of specifying weights matrix are border contiguity, distance-based matrix and *k*-nearest neighbours. Two weighting regimes namely contiguity and distance-based specification are the most extensively used matrices. In creating weight matrix using border contiguity, *GeoDa* provides the platform to select first-order weights and higher-order weights based on rook contiguity (common boundaries) and queen contiguity (both boundaries and vertices) (Anselin, 2002 and Feng and Humphreys, 2008). The distance weight matrix specification is used to define weights relative to distance by ensuring that neighbouring properties are jointly considered within the least possible cut-off distance, or the *k*-nearest neighbours (*k*-NN) to every specified observation. The advantage of using this weight matrix approach is that properties without neighbours are eliminated.

The specification of weights is directly related to the nature and features of property data. Consequently, if there are details relating to borders and lot sizes within the data the border contiguity is appropriate for use in weights specification. If the properties are scattered like in most rural areas and some urban areas, a threshold distance based spatial weights matrix is appropriate to ensure that at least all observations have a neighbour. The Cape Town property data used in this analysis being more contiguous (in each submarket) is feasible to border contiguity but since some properties are scattered in the study area, the distance-based spatial weights matrices are also appropriate. Therefore to select an appropriate weighting scheme between the two spatial weights based specification approaches a preliminary test was done. In the border contiguity weights based matrix, both the first-order rook contiguity and queen contiguity frameworks are used while in the distance-based weights matrix specification a threshold distance value of 650 metres was used. This is the minimum distance at which every observation has at least a neighbour. The preliminary results reveal the distance-based spatial weights matrix to be the most preferred as it generates better results than the border contiguity based weights matrix.

To select between alternative specifications, the Lagrange Multiplier (LM) test statistics that have a proven history of effectiveness was used. However, the LM tests must essentially be undertaken with the Moran's *I* test to check for model misspecification and spatial autocorrelation (dependence) in the property data. There are four alternative LM statistic tests including LM-error and robust LM-error (for spatial error model); LM-lag and robust LM-lag (for spatial lag model). The decision rule advanced by Anselin (2005) for LM tests was used to identify the best autoregressive model. Thus, if both LM tests for SEM and SLM are significant, the robust LM test should be used, but if the robust LM tests reveal both SEM and SLM to be significant the model with the largest value should be selected. Accordingly, Table 3 provides the summary of Moran's *I*, LM and robust LM diagnostic tests.

**Table 3.** Diagnostics for spatial dependence (autocorrelation)

| Test | mi/df | *z*–value | *p*–value |
|---|---|---|---|
| Moran's *I* (error) | 0.3138 | 72.8274 | 0.00000 |
| Lagrange Multiplier (lag) | 1.0000 | 2214.60 | 0.00000 |
| Robust LM (lag) | 1.0000 | 143.271 | 0.00000 |
| Lagrange Multiplier (error) | 1.0000 | 3922.23 | 0.00000 |
| Robust LM (error) | 1.0000 | 1850.90 | 0.00000 |

The Moran's *I* statistic with a z–value of 72.8274 is highly significant which suggests a spatial autocorrelation problem. Additionally, the Moran scatter plot in Figure 1 reveals the spatial patterns in the residuals of the linear (OLS) model. Furthermore, to select between SEM and SLM, the LM and robust LM tests show a high level of significance between the two models making it difficult to select one of the two models. There may be the presence of other misspecification glitches that undermine the asymptotic results arising from the robust LM test statistics (Anselin, 2005). This study, therefore, utilises the two spatial dependence modelling techniques, and thus applied the goodness of fit criteria ($R^2$, log-likelihood and Akaike Information Criteria (AIC)) and other benchmark tests acceptable to the International Association of Assessing Officers (IAAO) to select the best spatial dependence model.

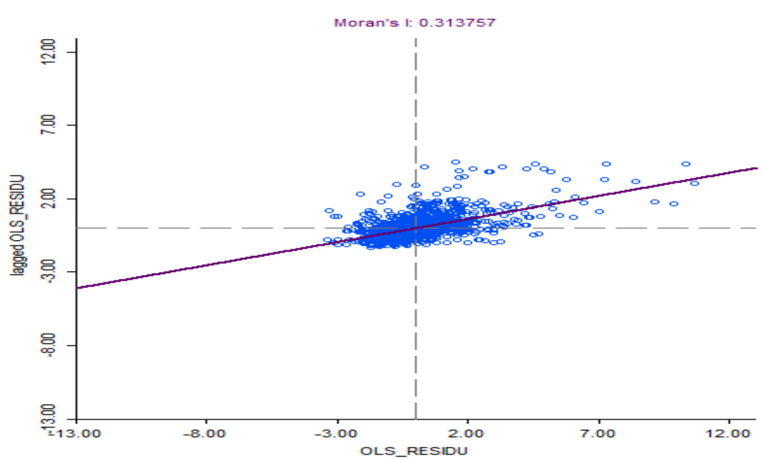


**Figure 1.** Moran scatter plot for residuals

Moreover, in choosing the best model for the South African property market, other models are not spatial but are devoid of the parametric restrictions of the OLS. These models will also be tested relative to the spatial models to select the technique that performs well in the modeling of property prices. The remainder of this section is devoted to discussion of the underlying philosophies of these models.

3.2.4 Artificial Neural Network Models

The ANNs is a parallel adaptive machine designed to mimic the operations of neurons in the human brain. The human biological neural systems are nerve cell sometimes referred to as neurons consisting of the cell body, dendrite and an axon that are tightly connected. The neuron takes a numerical input value from other neurons, summed them together with a bias node before effecting a simple transformation to generate an output value from the neuron (Collins and Evans, 1994). The multilayer perceptron (MLP) is the most extensively used ANNs architecture, consisting of three layers – the input layer, the output layer and at least a hidden layer for processing the nonlinear elements. According to a suggestion by Masters (1993), one hidden layer should be the initial choice for any practical ANNs design. Similarly, Lin and Mohan (2011) report that one single hidden layer is sufficient for the model to achieve accuracy in any complex nonlinear approximation (Hornik, 1991). In the current analysis, a hidden layer was used to build the architecture. The number of the input layer is determined by the configuration of the input data. Also, the output layer is effusively connected to the neurons of the input layer and this sequence of connection follow through the whole units of the network.

However, in the ANNs architecture, the total number of neurons in the hidden layers is a matter of user discretion which should be achieved through a trial and error process until an optimal result is found. Kwok and Yeung (1997) proposed that the method of finding an optimal number of hidden neurons is, to begin with, a smaller number of neurons, then increase the number until the desired result is established. This procedure was used by Lin and Mohan (2011) and McCluskey *et al*. (2013) and it is adopted in this study. The transfer function is used to measure the relationship between inputs and output (target) of the neuron and its network. This study used the tan-sigmoid in the neurons of the hidden layer and the linear transfer function in the neurons of the output layers. Furthermore, the network output is compared with the time adjusted assessed values to determine its accuracy by the error measure in equation 8

$$E(x,w) = \frac{1}{2}\sum_{c}^{C}\sum_{b}^{B}\left(d_{bc} - o_{bc}\right)^2 \tag{8}$$

where $x$ and $w$, are input and weight vectors of the network, $c$ is the index of patterns, from 1 to $C$, in which $C$ depicts the total number of training patterns; $b$ is the index of outputs, from 1 to $B$, in which $B$ is the total number of outputs and $d_{bc}$ *and* $o_{bc}$ are desired and actual values of the $c^{th}$ output and the $b^{th}$ patterns. The backpropagation (BP) training algorithm is the learning scheme used in this study. The training cycle was set at 1000 epochs, while the learning rate and momentum term were set at 0.3 and 0.2 respectively. The BP utilises the first-order derivative of the total error function to find the minimum in error space.

### 3.2.5 Support vector machines (SVMs)

The SVMs are techniques developed by Vapnik to solve classification and regression problems (Vapnik, 1999 and Zurada *et al*., 2011). The models became prominent due to their attractive features and successful application that traversed several fields. Specifically, the SVMs are used to detect frontal human face in images (Osuna, Freund and Girosi, 1997), financial time series (Tay and Cao, 2001), land cover classification (Dabike, Velickov, Solomantine and Abbott, 2001) and predict bankruptcy (Shin, Lee and Kim, 2005). SVMs uses input data and classify it into one of two groups or classes. SVMs first utilise a set of training input datasets, map them into multidimensional space, and use regression to discover a hyperplane[3] that is suitable for separating the two-class inputs. In selecting the best hyperplane for data classification, the one that characterises the largest separation or margin[4] between the two classes is the most preferred. Generally, the bigger the margin, the smaller the classifier's generalisation error will ensue.

For linearly separable data, SVMs use linear machines to train and separate data into the maximum distance (devoid of errors) between the hyperplane and the adjoining training points (Shin, Lee and Kim, 2005). The training points that adjoin the optimal separating hyperplane are referred to as support vectors while other training points are not important for defining the binary class borders. If the data is nonlinearly separated, SVMs will utilise nonlinear machines to train and discover a hyperplane that lessens the number of errors. Given the definition of a labelled training examples ($x_r$, $y_r$), an input vector $x_r \in R^n$, and a class value $y_r \in \{-1,1\}$, $r = 1, 2, ..., d$ (Shin *et al*., 2005).

For a linear separated instance, the decision rules defined by an optimal hyperplane separating the binary decision classes is given in equation (9) in terms of the support vectors (Shin, Lee and Kim, 2005).

$$Y = sign\left(\sum_{r+1}^{N} y_r \alpha_r (x \cdot x_r) + c\right) \tag{9}$$

The vector $x = (x_1, x_2, x_3, \ldots, x_n)$ corresponds to an input and the vectors $x_r$, where $r = 1, 2, \ldots N$, are support vectors, $Y$ denotes the outcome, $y_r$ denotes the class value of the training example $x_r$, $c$ and $\alpha_r$ are factors that determine the hyperplane, while the $\cdot$ symbolise the inner product. However, for a non-linearly separable instance, a kernel function $K$ is introduced in equation (9) to generate the inner products which construct machines that have different types of nonlinear decision surfaces in the input space as

$$Y = sign\left(\sum_{r+1}^{N} y_r \alpha_r K(x \cdot x_r) + c\right) \tag{10}$$

There have been many support vector machines that are used in practice. The study of Shin, Lee and Kim, (2005) reported three of the many different types of SVMs used in constructing the decision rules. These are: (1) polynomial kernel function, (2) a radial basis kernel function, and (3) a two-layer neural network machine with a kernel function. To achieve high accuracy, the parameters of the kernel function must be properly tuned. Two parameters namely the $C$ bound and $\gamma$ kernel parameter must be determined, however, the parameters are varied to select the optimal values for the best performance (Lam *et al*., 2009). The study of Tay and Cao (2001) reports on the sensitivity of SVMs to the parameters setting, of which, a chosen values could either lead to over-fit or under-fit of the training data. The implication of this is that the optimal result could only be achieved after several runs. According to Tay & Cao, the $C$ bound has a range of between 1 and 100. In this study, the limiting value of $C$ was set at the minimum value of 1.0, while the batch size was left at 100 to minimise error.

SVMs initially suffer a setback with its training, particularly when large datasets are used for quadratic programming (QP) solver to train. With this limitation, Osuna, Freund and Girosi (1997) noted that since the problem with QP is the requirement for enormous large scale data, the breaking down of the problem into a series of smaller sized QP problems will better optimise SVMs. In keeping with Osuna *et al's* idea, examples must be added and subtracted to keep the matrix size constant. However, Platt (1998) observes this process as ineffective and introduced sequential minimal optimisation (SMO) to swiftly tackle the SVMs QP problem without adding additional matrix space and numerical QP optimisation phases at all. To guarantee convergence, SMO disintegrates all QP problems into sub-problems without using QP solver to provide a solution. Thus, the process of finding new optimal value requires that SMO selects two combined Lagrange Multipliers (LM) to find optimal values for SVM updates. Thus the benefits of using SMO for SVMs training is that it avoids the numerical QP optimisation to solve for two LMs analytically.

#### 3.2.5.1 Neural networks Support vector machines (NNSVM)

The hybrid system is designed in such a way that the output of the combined ANNs and SVMs (NNSVM) is supplied by SVMs which take its input from a small central feature layer. The feature layer is successively the output of the ANNs trained with the backpropagation of the derivatives of the dual objectives of the SVMs concerning the feature node values (Wiering *et al*. (2013). To overcome the limitations of the standalone ANNs and SVMs, the NNSVM adds more layers to the SVMs making it "deeper", thereby enabling the ANNs to learn many features and increase the flexibility of the kernel functions. The hybrid systems used in this study is consistent with the work of Wiering *et al.* (2013) which also modified the approaches of Vincent and Bengio (2000), and Ghanty, Paul and Pal (2009). The NNSVM has an input layer comprising of $Q$ nodes representing property attributes; a central feature layer $z$ which consist of $q$ nodes; a MLP with three layers of neural networks $N$. When a training pattern $p$ of dimension $Q$ is presented to the NNSVM, this is propagated through the ANNs for values of the feature layer to be determined. The equation that follows represents how NNSVM calculate its output

$$f(p)=\sum_{r=1}^{N}\left(\alpha_r^{*}-\alpha_r\right)K(\Phi\left(\left(p_r/\theta\right),\Phi\left(p/\theta\right)\right)+b \tag{11}$$

in which K (·,·) is the chosen kernel function of SVM. The architecture of the NNSVM regression estimator consisting of three ANNs is given in Figure 2.

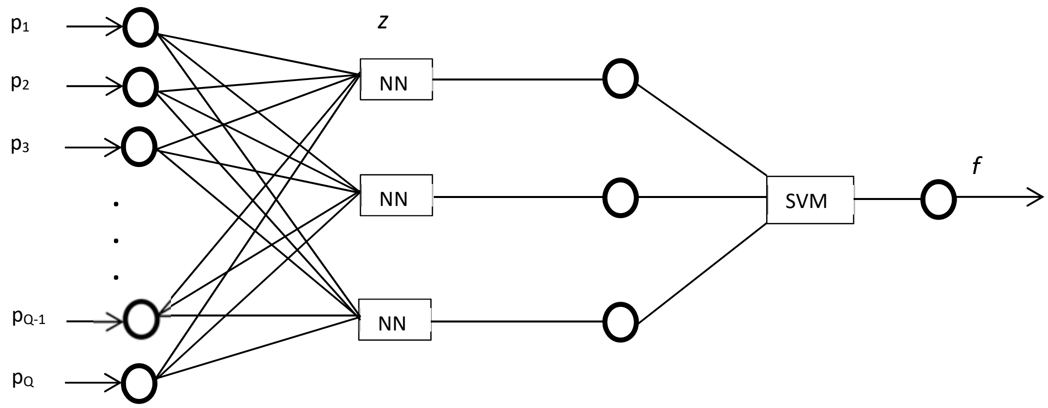


**Figure 2**. An architecture of NNSVM Regression Estimator (Adopted from (Wiering *et al*. 2013p. 249)).

The architecture in Figure 2 is the novel idea of Wiering *et al*. (2013) which is extended using "average" combination rule in this study for mass appraisal of properties. However, while their study ensemble SVMs and ANNs via stacking technique, this study ensemble SVMs and ANNs using the voting technique. The hybrid algorithm utilises neural networks for providing features that are used by support vector machines to determine final architectural output. This is achieved by allowing MLP's association with SVM in which SVM replaces each neuron of the network. There exist a sparse number of studies that utilise this type of combination to improve the performance of ANNs and SVMs. Specifically, Vincent and Bengio (2000) utilised the hybrid system in several experiments on classification dataset and found the performance to be comparable to standalone SVMs. Also, Ghanty *et al*. (2009) used the NNSVM on twelve benchmark data sets and found it to consistently perform better than the ANNs and SVMs.

## 4. Modelling results and discussion

This section contains the empirical part of the study. First, the section provides analysis of the goodness of fit measurement of all models; secondly, the detailed results of the regression-based (GWR, OLS, SEM and SLM) and artificial intelligence-based (ANNs, SVMs and NNSVMs) techniques is given in line with the benchmark test recommended by the IAAO. Thirdly the regression coefficients for the regression-based models are provided. Fourthly, the detailed result of the GWR revealing price variation across the Cape Town property market is provided. Lastly, a comparative analysis of all models to select an optimal model for mass appraisal of properties in the Cape Town property market context is provided.

### 4.1 Predictive accuracy and performance of mass appraisal models

The performance of the linear OLS (used as a benchmark for comparison with other models) represented by the $R^2$ reveals an explanatory power of (0.5887) 59 per cent of the variability in property prices (Table 3). The goodness of fit measurement of GWR, SEM and SLM represents a robust explanation of the variability in property prices relative to the OLS. In comparison, the SEM and SLM have similar results with those of GWR as revealed in their $R^2$ (0.7456, 0.7348 and 0.7309) with autoregressive models slightly outperforming the GWR. This shows that 75 per cent, 74 per cent and 73 per cent of the variability in the property prices are explained by the autoregressive (SEM and SLM) and non-stationarity (GWR) models. The high $R^2$ of SEM, SLM and GWR in comparison to the

OLS reflect the relative importance of spatial models in remediating the spatial bias limitations of the OLS hedonic regression models. Additionally, the study of Wilhelmsson (2002) noted that SEM and SLM can correct the shortcoming of a locational dummy variable in the OLS. The lagged variable in SLM is what enhances the $R^2$ relative to the OLS model. Again the Akaike Information Criterion (AIC) statistic which illustrates that estimates of $\beta_i$ are based on the least-squares and the maximum likelihood estimates are identical (McCluskey, *et al*., 2013) reveals similarities among the four models but the OLS (104003) produced an AIC goodness of fit that is slightly below SEM (102656), SLM (102710) and (GWR (103212). In terms of the log-likelihood, the SEM (-51294) had a better fit than SLM (-51319), GWR (-51540) and OLS (-51953).

**Table 3.** Goodness-of-fit measurements for all models

| Model | | $R^2$ | AIC | Log likelihood |
|---|---|---|---|---|
| OLS | | 0.5887 | 104003 | -51953 |
| SEM | | 0.7456 | 102656 | -51294 |
| SLM | | 0.7348 | 102710 | -51319 |
| GWR | | 0.7309 | 103212 | -51540 |
| ANNs | All data | 0.6463 | - | - |
| | Train data | 0.6657 | - | - |
| | Test data | 0.7935 | - | - |
| SVMs | All data | 0.5467 | - | - |
| | Train data | 0.5588 | - | - |
| | Test data | 0.5222 | - | - |
| NNSVMs | All data | 0.6519 | - | - |
| | Train data | 0.7712 | - | - |
| | Test data | 0.6569 | - | - |

The ANNs, SVMs and the hybrid system are constructed differently from the regression-based techniques. To avoid over-fitting and excessive training in the ANNs, 70 per cent and 30 per cent split of the dataset was applied. However, to compare variability in property prices between the artificial intelligence and the regression-based models the 100 per cent data was used. It is necessary to note that artificial intelligence-based techniques do not produce AIC and log-likelihood goodness of fit. The results reveal a good interaction between the sale prices and the appraised values for all data, training and testing datasets, respectively. The result shows that 65 per cent, 67 per cent and 79 per cent variability in property prices is explained by the ANNs architecture, for all (100 per cent) data, training data and testing data (Table 3). The variability in the price of properties is also explained by the SVMs as revealed by the goodness of fit. Specifically, in the all (100 per cent) data about 55 per cent, 56 per cent (training data) and 52 per cent(testing data) of the variation in property values is explained by the SVMs. Furthermore, for the hybrid system, the $R^2$ reveal 65 per cent (all 100 per cent data), 77 per cent (train data) and 66 per cent (test data) variation in property prices are explained by the NNSVMs. In comparison to standalone ANNs and SVMs, the hybrid system performed marginally below the ANNs in the test data, but higher than the SVMs, thus revealing its superiority over the standalone SVMs in all stratified datasets, but superior to ANNs only in the 100 per cent and train datasets. In all the spatial models have superior explanatory powers in comparison to the OLS and the artificial intelligence-based (100 per cent data) techniques in this assessment. The next phase of the analysis reveals the detail performance of models (Table 4).

**Table 4.** Performance comparison of models

| Model | Performance measures | | | | | |
|---|---|---|---|---|---|---|
| | Median ratio | Mean ratio | PRD | COD (%) | MAE | RMSE |
| OLS | 1.05 | 1.08 | 1.06 | 28 | 1482664 | 2399107 |
| SEM | 1.00 | 1.03 | 1.03 | 17 | 1087444 | 1822085 |
| SLM | 1.02 | 1.03 | 1.05 | 19 | 1165532 | 1860829 |
| GWR | 1.03 | 1.06 | 1.07 | 20 | 1109322 | 1940656 |
| ANNs (all data) | 1.03 | 1.05 | 1.08 | 26 | 1527818 | 2495936 |
| (train data) | 0.79 | 0.83 | 1.06 | 29 | 1406662 | 2167361 |
| (test data) | 1.04 | 1.09 | 1.03 | 20 | 1947283 | 2314584 |
| SVMs (all data) | 1.00 | 1.02 | 1.12 | 24 | 1420005 | 2665473 |
| (train data) | 1.00 | 1.02 | 1.12 | 22 | 1412577 | 2589821 |
| (test data) | 1.00 | 1.01 | 1.13 | 23 | 1416924 | 2821858 |
| NNSVMs (all data) | 1.01 | 1.03 | 1.11 | 21 | 1349933 | 2436083 |
| (train data) | 0.98 | 1.00 | 1.12 | 21 | 1301761 | 2237006 |
| (test data) | 1.01 | 1.02 | 1.12 | 22 | 1457704 | 2045910 |

Table 4 reveals the results of the regression-based models in terms of model accuracy. The root mean squared error (RMSE) test reveals the SEM (1822085) to perform better in comparison to SLM (1860829), GWR (1940656) and OLS (2399107). The results show a slight outperformance of the SLM over the GWR. According to Limsombunchai, *et al*. (2004), a model with the lowest RMSE is considered the best in terms of prediction accuracy. The mean absolute error (MAE) value reveals the SEM (1087444) to predict prices that are closer to the sale price than the SLM (1165532), GWR (1109322) and OLS (1482664) models. Apart from the results of ANNs for train dataset, all models performed well in terms of mean and median ratios. The quality assurance benchmark test is fundamental to property assessors because of the feedback they offer on the overall accuracy of predicted values. The standard stipulates that a median ratio of between 0.90 and 1.10 meets the required standard. Similarly, in terms of their coefficient of dispersion (COD) and price related differentials (PRD), the SEM performed better than all other models in this analysis. The standard benchmark required for a model to estimate acceptable prices or values for residential properties is within the range from 5.0 to 15.0 and 0.98 to 1.03, respectively.

The lower part of Table 4 reveals the performance of the artificial intelligence-based models relative to the results of the test model. Consequently, the RMSE accuracy-test reveals that the hybrid system (2045910) performs better in comparison to the ANNs (2314584) and SVMs (2821858), clearly demonstrating that in terms of prediction accuracy it is superior. Also, in term of prediction of property prices closer to the sale price, the MAE reveals the SVMs (1416924) to predict prices that are closer to sale prices than the ANNs (1947283) and the hybrid system (1457704). Though all artificial intelligence-based techniques demonstrate regressivity because of high PRD and COD, the ANNs demonstrate COD and PRD that is better with vertical equity and uniformity across the total sales compared to the SVMs and NNVMs in the test sample. In all assessment, the results suggest that the SEM in terms of vertical equity, uniformity and horizontal dispersion is the best model for the Cape Town property market.

The next analysis reveals the regression coefficients, t-statistics, and significant level for OLS, SEM and SLM are reported in Table 5. The results reveal lag coefficients (Lambda (SEM) and Rho (SLM)) for the autoregressive models. In terms of the overall lag coefficients performance, the SEM (0.919892) outperform the SLM (0.827061). However, both are highly significant ($P < 0.0000$) and usual for observations that exceed 1000 which also reflect the asymptotic nature of the analytical expression used for the variance (Anselin, 2005).

**Table 5.** Regression based model coefficients

| Variable | OLS | | SEM | | SLM | |
|---|---|---|---|---|---|---|
| | Coefficients | *T* | Coefficients | *t* | Coefficients | *t* |
| W_TASP | | | | | 0.827061 | 83.9876* |
| Constant | 2253740 | 13.0185* | 1825150 | 4.58261* | -2124370 | -18.2099* |
| Bed_1 | -114204 | -0.217935 | -460830 | -1.13622 | -355264 | -0.852534 |
| Bed_2 | 32755.5 | 0.19818 | -163868 | -1.29086 | -29125.9 | -0.223191 |
| Bed_4 | 30476.2 | 0.289311 | 179722 | 2.2241* | 162942 | 1.9606* |
| Bed_5 | 350253 | 2.16015* | 564891 | 4.5392* | 536532 | 4.17958* |
| Bed_6 | 745329 | 2.77931* | 735542 | 3.56032* | 857644 | 4.04568* |
| Bed_7 | 761550 | 1.32682 | 717524 | 1.628 | 882186 | 1.93251* |
| Bed_8 | 5960390 | 6.46576* | 5058950 | 7.1176* | 5060670 | 6.89354* |
| Bed_9 | 96600.4 | 0.0561522 | -22273.2 | -0.016787 | 4956.98 | 0.003617 |
| Bed_10 | 4411440 | 1.8121 | 2927450 | 1.58562 | 2246400 | 1.15851 |
| Quality_poor | -332301 | -0.223268 | -431231 | -0.380218 | -884402 | -0.745923 |
| Quality_fair | -60844.5 | -0.117857 | 132826 | 0.336369 | -129684 | -0.316295 |
| Quality_good | 536926 | 3.94148* | 534073 | 5.03926* | 389507 | 3.61127* |
| Quality_v/good | 3170770 | 7.08062* | 2712160 | 7.8963* | 2671190 | 7.48881* |
| Quality_excellent | 3306300 | 7.9586* | 2062280 | 6.41534* | 2164420 | 6.55603* |
| Condition_poor | 452650 | 0.359788 | -569131 | -0.594358 | 244423 | 0.244027 |
| Condition_fair | -60174.3 | -0.147467 | -258533 | -0.823619 | -137250 | -0.422474 |
| Condition_good | 299979 | 2.20331* | 412130 | 3.92275* | 247995 | 2.31729* |
| Condition_excel. | -242585 | -1.01948 | 244683 | 1.29842 | -25527.1 | -0.138077 |
| Storey_2 | 735701 | 7.35126* | 837208 | 10.5127* | 769443 | 9.79951* |
| Storey_3 | 2868780 | 9.84054* | 2085690 | 9.15798* | 2346360 | 10.1916* |
| Style_s/economic | 355148 | 0.252222 | -3342.35 | -0.003089 | -17965.7 | -0.016038 |
| Style_Unconven. | 2525130 | 8.74038* | 1361700 | 5.96028* | 1641270 | 7.17997 |
| Style_G/victor | 674370 | 1.74641 | 4685.01 | 0.015467 | 497944 | 1.62271* |
| Style_C/Dutch | -382395 | -0.564484 | 9327.06 | 0.017893 | -103824 | -0.192614 |
| Style_Maisonette | 158156 | 0.387754 | -254658 | -0.801553 | 97907.1 | 0.302235 |
| Style_Mediterr. | -37888.8 | -0.0344716 | -428808 | -0.480365 | -387918 | -0.448612 |
| View_p/obstructed | 232662 | 1.1098 | 229120 | 1.35195 | 40992.5 | 0.245812 |
| View_b/average | -135474 | -0.241458 | 30027.4 | 0.069096 | 53962.6 | 0.120849 |
| View_a/ average | 579379 | 5.73054 | 460771 | 5.095* | 339444 | 4.25587* |
| View_Panoramic | 1608060 | 11.3489* | 1276210 | 9.50211* | 1202380 | 10.9944* |
| View_Excellent | 2408660 | 6.9387* | 1901970 | 6.77239* | 1730320 | 6.30929* |
| Property size | 11082.1 | 15.3483* | 11643 | 20.3755* | 9605.37 | 17.2909* |
| Swimming pool | 8743.85 | 3.3141* | 8111.74 | 3.91482* | 8331.61 | 4.03514* |
| Lambda | | | 0.919892 | 110.152* | | |
| Dependent variable: TASP | | | | | | |

* depicts significant at 95%

Locational dummies for OLS are presented at the appendix

The property size and swimming pool variables are significant in OLS, SEM and SLM. There relative contribution to property price reveals similarities, appropriate signs and all have a strong and positive influence on property prices. Again the variables storey_2 and storey_3 are significant and have a positive influence on the price in the three regression-based models. The base storey building in the study area is one storey depicting that two and three storey buildings are worth more than one storey building. Also good, very good and excellent quality properties are significant and have a positive relationship to property price in Cape Town as revealed by results of the three models. This was expected as good quality property attracts more home buyers and thus increases their willingness to pay while the poor and fair quality property has a negative influence on property prices (Table 5). Similarly, while properties that are of poor and fair qualities should ordinarily attract less additive and significant contribution to property prices, the same could not be said of properties in excellent condition (condition_excel) in OLS and SLM which should primarily be additive towards value but at a depreciating rate. Importantly, the spatial lag term in Table 5 is positive and significant indicating that property prices in Cape Town are strongly influenced by the prices of nearby properties in line with Tobler's (1979) first law of geography to wit nearby properties are more related than properties that are far apart.

The results also show that properties with more than three bedrooms contribute more to property prices than properties with less than three bedrooms in Cape Town. In all, there is a high degree of similarities in the number of significant variables in SEM, SLM and the OLS. The negative values in some of the variables including bedroom, quality and condition of properties, and building style is contrary to the *a priori* expectation, but might be due to the law of diminishing marginal utility, i.e. an increase to the number of bedrooms from one to two should ordinarily be an additive to property price but at a depreciating rate. If combined with other variables, although within limits, multicollinearity might influence the results and provide incorrect absolute values which can only be assessed at the individual attribute level. This would normally occur where insufficient data points exist that represent a particular variable.

The GWR results reflect the importance of localised spatial influences within the Cape Town property market (Table 6). The GWR parameter estimates vary at each of the 3225 observation points is revealed in their minimum, maximum, median, interquartile, lower and upper quartile ranges. There is variability at each of the observation points in all the variables as shown in the interquartile range greater than zero. Consequently, the parameter estimates reveal a variation over space for all the independent variables. Specifically, the coefficient of the variable property size ($m^2$) reveals that property within a particular precinct of the study area commands a price of R1805.97 ($m^2$), while a property in another area can command a price of R23657.54 ($m^2$).

**Table 6.** GWR model coefficients

| Variable | Minimum | Lower quartile | Median | Interquartile | Upper quartile | Maximum |
|---|---|---|---|---|---|---|
| Intercept | -13386996.1 | -5391446.65 | -2867981.29 | 4074114.235 | -1317332.42 | 9492400.19 |
| Beds | -474789.524 | 78061.397 | 196950.550 | 272839.9989 | 350901.396 | 1901310.70 |
| Quality | -572477.466 | 467445.755 | 703019.935 | 571590.1544 | 1039035.91 | 2520509.13 |
| Condition | -1681390.34 | -91215.7038 | 207551.18 | 501777.3591 | 410561.655 | 1374501.96 |
| Storey | -1045020.84 | 665911.745 | 1088328.17 | 759953.1534 | 1425864.89 | 2794285.85 |
| Style | -4475244.66 | -821664.132 | -145811.93 | 895466.5544 | 73802.4223 | 1459943.55 |
| View | -39664.4375 | 180541.307 | 319954.579 | 353097.0535 | 533638.36 | 2984422.65 |
| Size | 1805.97378 | 10144.1165 | 13353.5358 | 5717.240045 | 15861.3565 | 23657.5412 |
| Pool | -45036.3109 | 854.34007 | 7934.54583 | 15577.74802 | 16432.0881 | 106514.939 |

The negative sign on the lower end of the property price in some of the attributes is somewhat counterintuitive. This kind of scenario was observed in the studies of Bitter *et al*. (2007), and McCluskey *et al*. (2013). This might be the effect of low quality and out-dated building style/design which might require modernisation. However, details about the state of the property can only be verified from the property agents and tax assessors. Specifically, the building style (style) has results which range from R-4475244.66 to R1459943.55 which is linked to an increase in the value of a property with a good building style in one area while property in another is low with a negative sign. Again, the parameter estimates for the storey shows that "*ceteris paribus*" it sold from a range of as little as R-1045020.84 at one location and R2794285.85 more at another location of Cape Town. The negative estimates suggest that the value of a storey building in Cape Town is highly dependent on the location. The result shows the benefit of using a non-stationary model over the stationary–coefficient model as it reveals prices of properties from different neighbourhoods/locations within a city.

## 4.2 Optimal selection of mass appraisal models for the Cape Town property market

In all, seven models are used from which the best technique for mass appraisal of properties in the Cape Town property market is to be objectively selected. The tests used for comparison are the model performance, explainability and reliability ranking order used in McCluskey *et al*. (2013) and adopted in Yacim and Boshoff (2018b) in line with the suggestions of McCluskey (1997); Kryvobokov (2004), d' Amato Bitter *et al*. (2007), and Kauko (2008). This study utilised these tests to give a fair assessment of models because the decision is not only based on predictive accuracies but also explainability and usage within the mass appraisal environment. It is not enough for a model to perform optimal relative to price prediction; its ability to receive acceptability among practitioners is sine-qua-non. The internal working of a model is what appraisers are particularly interested in knowing so that estimates can easily be defended should there be any appeal from the dissatisfied party(ies). The preceding results reveal that most models do not meet the IAAO guidelines in terms of the accuracy of the model outputs, but rather serve as a baseline of comparison for the methods described in this study (Borst, 2006). This might be the consequences of model misspecification and outliers. The first test in this section is the performance and reliability ranking order test relative to what the different platforms of analysis provided. The coefficient of determination ($R^2$), median ratio, mean ratio, COD, PRD, MAE and RMSE accuracy statistics were used. The analysis is carried out on the results produced by the 100% data used for the artificial intelligence-based models. This is to bring all results to be in line with the 100% data used for the regression-based models. Furthermore, the model with the overall lowest rank is preferred as the best. The results are summarised in Table 7.

**Table 7**. Performance accuracy and reliability ranking order

| Accuracy measures | OLS Rank | SEM Rank | SLM Rank | GWR Rank | ANNs Rank | SVMs Rank | NNSVMs Rank |
|---|---|---|---|---|---|---|---|
| $R^2$ | 6 | 1 | 2 | 3 | 5 | 7 | 4 |
| Median ratio | 7 | 1 | 4 | 5 | 5 | 1 | 3 |
| Mean ratio | 7 | 2 | 2 | 6 | 5 | 1 | 2 |
| PRD | 3 | 1 | 2 | 4 | 5 | 7 | 6 |
| COD | 7 | 1 | 2 | 3 | 6 | 5 | 4 |
| MAE | 6 | 1 | 2 | 3 | 7 | 5 | 4 |
| RMSE | 4 | 1 | 2 | 3 | 6 | 7 | 5 |
| Mean | 5.71 | 1.14 | 2.29 | 3.86 | 4.71 | 4.71 | 4.00 |
| Position | 7 | 1 | 2 | 3 | 6 | 6 | 4 |

The comparative analysis reveals the SEM to perform best than all other models relative to all performance measures with the exclusion of mean ratio. Interestingly, all spatial models performed better than other models ranking first, second and third, respectively. The ability to adequately capture explicit location and deal with the problems of spatial dependence and spatial heterogeneity gave the SEM, SLM and GWR an edge over the OLS and the artificial intelligence-based models utilised in this analysis. Additionally, in the SLM, the inclusion of a lag term help to curb potential bias resulting from the omission of this variable (Bitter *et al*., 2007). Similarly, the third-place earned by the GWR stems from its ability to spatially mapped and visualised patterns in the parameter estimates. For instance, the GWR vary the property prices in the city of Cape Town, thus capturing different prices relative to the location characteristics. The general position in the valuation field that no two properties are exactly alike is adequately taken care of by the spatial models, particularly the GWR. The SVMs performed analogously to the ANN models in this analysis as suggested by the result in Table 7. The NNSVM performed better than the standalone ANNs and SVMs, and the OLS suggesting that building a hybrid system of the duo has enhanced its performance. Although the marginal performance of the intelligent-based techniques in comparison to the OLS supports previous studies that they are superior, their performance in this study, in comparison to spatial models is contrary to *a priori* expectation. This is because of their high computing and pattern recognition properties.

Furthermore, the test in Table 8 is fundamental to the selection of a model, this is because as McCluskey *et al*. (2012) observe, the predictive accuracy should not be the only measure of selecting a model. The technique must sufficiently demonstrate that it can be used within the mass appraisal environment, can effectively capture the location of the property and compare it based on the relative distance of the nearest neighbour. Five qualitative measures including simplicity, consistency, transparency, locational and applicability within the mass appraisal environment were used in testing a model.

**Table 8.** Model ranking based on the degree of explainability and acceptability

| Qualitative measures | OLS | SEM | SLM | GWR | ANNs | SVMs | NNSVMs |
|---|---|---|---|---|---|---|---|
| Simplicity | 1 | 2 | 2 | 3 | 4 | 4 | 5 |
| Consistency | 1 | 1 | 1 | 2 | 3 | 3 | 4 |
| Transparency | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| Locational | 3 | 2 | 2 | 1 | 4 | 4 | 4 |
| Applicability | 1 | 1 | 1 | 1 | 2 | 2 | 2 |
| Mean ranking | 1.40 | 1.40 | 1.40 | 1.60 | 3.00 | 3.00 | 3.40 |
| Position | 1 | 1 | 1 | 4 | 5 | 5 | 7 |

The results demonstrate that the OLS, SEM and SLM are preferred relative to other techniques utilised. Thus in terms of simplicity, the OLS is relatively simple to use and its estimate are consistent and transparent as it reveals the contribution of every variable to property prices but became affected by location. The OLS uses indicator dummy variables to capture the influence of location on property prices which to a reasonable extent remediate the problem but as noted earlier spatial effects are not completely tackled (see Moran's *I* result in Table 3). The SEM and SLM have these features, though not as simple as the OLS but surpass the OLS in terms of the manner they explicitly capture the location of properties and correcting the problem of spatial dependence. The ANNs, SVMs and NNSVMs also utilised the locational dummies to account for property location. It was noted in McCluskey *et al*. (2013) that explicit location could be used in the artificial intelligence-based techniques but this was not done in this study. The ANNs, SVMs and NNSVMs though also easy to use but are not simple nor consistent as each time an adjustment is made, a completely different result is achieved. Additionally, their estimates are not transparent because of the black-box nature of these techniques. Transparency is important to the appraisal process, because should there be an appeal from dis-satisfied parties the estimate is what the appraiser would use to defend the entire process. Much has been said on the GWR in the preceding paragraphs hence its performance in this test was *a priori* expected. Having regards to the performance of the four regression-based techniques (SEM, SLM, GWR and OLS), and the strength of the result in Table 7, the SEM is the most preferred model in this analysis and thus should be used for pricing of properties in the Cape Town property market.

The high COD of the SEM (most preferred), notwithstanding, it can safely be used within the South African property market for levying property taxes and mortgage underwriting. The city valuation office, Cape Town and probably other valuation offices within the South African market are familiar with the OLS. Thus, applying the SEM into the property market should provide fewer challenges to these bodies. Should expertise be needed in handling spatial modelling tools, experts from countries that have successfully implemented the techniques could be invited to provide requisite services. The SEM and other spatial models could be applied in other property markets within the sub-Saharan African countries provided a test for suitability is first undertaken. Furthermore, jurisdictions with data limitation or non-availability should make it a priority to secure quality and usable data. Moreover, academics would find the use of these modelling tools very useful not only in research, but giving of instructions to students on mass appraisal of properties.

## 5. Conclusion

The hedonic OLS models have traditionally been utilised for the appraisal of properties but have over the years been associated with several limitations including functionality, nonlinearity and poor handling of spatial effects. Several models including the ANNs, GWR, SEM, SLM, SVMs amongst others have been introduced to deal with the obvious weaknesses of the OLS. However, previous studies undertaken to compare the performance of different modelling approaches to ascertain the superiority have not yielded the needed results. Primarily the major concerns are related to the data quality and the geographical contexts where the study is been undertaken. The artificial intelligence and spatial regression models offer good modelling methodology that is not affected by the many limitations of the OLS. The ANNs and SVMs have powerful pattern recognition properties that can efficaciously recognise complex value effects in the property price analysis. The spatial models

(GWR, SEM and SLM) have abilities to tackle spatial heterogeneity and spatial dependence thus predicting property prices devoid of inefficient and inconsistent parameter estimates.

The recent developments in the artificial intelligence field have tended towards the use of the hybrid system to remediate the limitations of the individual model to increase its capability and predictive performance. Whilst it is true that ANNs and SVMs are flexible in handling nonlinearities and developing price models without having to create binaries or linearise variables through transformation; again whilst it is equally true that ANNs have relatively overcome training limitations with the use of nature-inspired and other improved algorithms, they do however suffer from overfitting (with ANNs) and choice of kernel function (with SVMs). These limitations have since been identified and resolved through the building of a hybrid system of ANNs and SVMs in other fields but till date, this has not been utilised in the mass appraisal industry. Therefore, this study extends the use of the hybrid system into the mass appraisal field. Additionally, this research provides a thorough and complete comparison of various high-level appraisal methods, which provide testing on a single database to compare all on an equal base. To the knowledge of the researchers, this has not been conducted in the past. In addition to this, the research provides an introduction of these models for use in South Africa, where even the OLS is used to a very limited extent.

Several evaluation criteria were used to assess the level of their performance. The $R^2$ used only reveal the explainability of variance in property prices of the models but other accuracy-test statistics such as MAE, RMSE and those acceptable to the IAAO including the PRD and COD revealed much detail of the performance of models. Within the mass appraisal environment, a model is adjudged worthy of use if it is simple, consistent and transparently provide details that can aid appraisers defence before a tribunal. Unfortunately, despite the relative performance of the artificial intelligence-based models in this study, these requirements are virtually absent in the ANNs, SVMs and the hybrid system but the GWR, OLS, SEM and SLM provides the details. The spatial models (GWR, SEM and SLM) have all proved to be better in price estimation because of their abilities to tackle the problems of spatial dependence and spatial heterogeneity with the SEM surpassing them. Therefore the SEM is favoured for mass appraisal of properties in the Cape Town property market. The ANNs and SVMs, however, still have a good history of ease of use, cost-effectiveness and capability of handling nonlinear data which placed them at an advantaged position of being used as a check on the estimates produced by the regression-based models. Similarly, the NNSVMs built in this study can also be used as a check on estimate predicted by the OLS, SEM, SLM and GWR because in this analysis it has proved to be a superior model to the other artificial intelligence based techniques.

Following the preceding discussion, future research can be directed towards the possible improvement of the methodologies. Possible way of such improvements could be to (i) use a semi-log form of the response variable in the spatial models; (ii) revisit the Cape Town property data in order to see what causes the problem of higher COD and PRD in the techniques used, i.e. accuracy of data collected, model specification errors caused by uncaptured attribute information, or market inefficiencies; (iii) data from other provinces within South Africa would be useful to strengthen the findings; and (iv) improve the degree of explainability of ANNs and SVMs through a study that encourage the visibility of the black-box such that parameter estimates can be observed clearly. Once this requirement is met, the artificial intelligence-based models would be regarded as substitutes to the regression-based techniques and not as supporting or complementing tools. It might also be of interest to consider building a hybrid system of the OLS and the SVMs so that results from SVMs could be used to improve the output of the OLS. This will guarantee transparency and explainability of estimates by appraisers.

Notes

1 Kernel functions are algorithms used to change the input data into a form necessary for use in the SVMs. This class of algorithms perform the task of pattern analysis within a data. There are several of kernel functions including the RBF, polynomial, Gaussian, and sigmoid among others

2 Dimensionality reduction is process of reducing many input variables in a model. If the number of input variables is high it would affect analysis and reduce the strength of the model. The step-wise regression approach can be used to eliminate all redundant variables and preserve a set of principal variables. The SVMs has such functionality that enables it to reduce redundant variables.

3 Hyperplane is used to separates an input data into classes without having to overlap such cases. In effect, it is like a straight line that separates data into different classes. The property data sometimes exhibit chaotic behaviour between linear and non-linear elements, the hyperplane can safely classify the data into several classes.

4 Margin is needed in SVMs to ensure accurate classification which sometimes poses a challenge during separation by hyperplane. Thus during classification, the SVMs look for the right hyperplane that ensure margin maximisation and reduction or minimisation of misclassification.

**References**

Anselin, L. and Bera, A.K. (1998), "Spatial dependence in linear regression models with an introduction to spatial econometrics", Handbook of Applied economic statistics (eds. Ullah A. & Giles, D.E.A.), pp. 237-289, Marcel Dekker, New York.

Anselin, L. (1988), *Spatial Econometrics: Methods and Models*, Dordrecht: Kluwer Academic Publisher.

Anselin, L. (2003), "Spatial Econometrics", in *A Companion to Theoretical Econometrics*, Ed. Baltagi, B. H., Blackwell Publishing Limited.

Anselin, L. (2005), *Exploring Spatial Data with GeoDaTM: A Workbook*. http://www.csiss.org/clearinghouse/GeoDa/geodaworkbook.pdf

Bitter, C., Mulligan, G.F. and Dall'erba, S. (2007), "Incorporating spatial variation in housing attribute prices: a comparison of geographically weighted regression and the spatial expansion method", *Journal of Geographical Systems,* Vol. 9 No. 1, pp. 7-27.

Borst, R.A. (1995), "Artificial neural networks in mass appraisal", *Journal of Property Tax Assessment & Administration*, Vol. 1 No. 2, pp. 5-15.

Borst, R.A. (2006), "The comparable sales method as the basis for a property tax valuation system and its relationship and comparison to geostatistical valuation models", *paper in the international congress on advances in mass appraisal methods*, October, 30th, Delft university of technology, Netherland, pp. 9-19.

Borst, R.A. (2007), Discovering and applying locational influence pattern in the mass valuation of domestic real property, PhD thesis submitted to the faculty of engineering, University of Ulster, UK.

Borst, R.A. (2012), "A space-time model for computer assisted mass appraisal", *XLI Incontro di Studio del Ce.S.E.T.*, pp. 535-545.

Bourassa, S.C., Cantoni, E. and Hoesli, M. (2007), "Spatial dependence, housing submarkets and house prediction", *Journal of Real Estate Finance & Economics*, Vol. 35 No. 2, pp. 143-160.

Bourassa, S.C., Cantoni, E. and Hoesli, M. (2010), "Predicting house prices with spatial dependence: a comparison of methods", *Journal of Real Estate Research,* Vol. 32 No. 2, pp. 143-160.

Brunsdon, C., Fotheringham, A.S. and Charlton, M.E. (1996), "Geographically weighted regression: a method for exploring spatial nonstationarity", *Geographical Analysis*, Vol. 28 No. 4, pp. 281-298.

Casetti, E. (1972), "Generating models by the expansion method: applications to geographical research", *Geographical Analysis*, Vol. 4 No. 1, pp. 81-91.

Casetti, E. (1997), "The expansion method, mathematical modelling and spatial econometrics", *International Regional Science Review*, Vol. 20 No. 1-2, pp. 9-32.

Cliff, A.D. and Ord, J.K. (1981), *Spatial Processes: Models and Applications*, Pion Limited, London.

Collins, A. and Evans, A. (1994), "Aircraft noise and residential property values: an artificial neural network approach", *Journal of Transport Economics and Policy*, Vol. 28 No. 2, pp. 175-197.

Court, A.T. (1939), "Hedonic price indexes with automobile examples, In General Motors Corporation", *The Dynamics of Automobile Demand*, New York, pp. 99-177.

Crooper, L.M, Deck, L.B. and McConnell, K.E. (1988), "On the choice of functional form for hedonic functions", *The Review of Economics and Statistics*, Vol. 70 No. 4, pp. 668-678.

Cui, D. and Curry, D. (2005), "Prediction in marketing using the support vector machine", *Marketing Science,* Vol. 24 No. 4, pp. 595-615.

Dabike, Y. B., Velickov, S., Solomatine, D. and Abbott, M. B. (2001), "Model induction with support vector machines: introduction and applications", *ASCE Journal of Computing in Civil Engineering*, Vol. 15 No. 3, pp. 208-216.

d' Amato, M. and Kauko, T. (2008), "Property market classification and mass appraisal methodology", In Kauko, T. and d'Amato, M. (Eds) *Mass Appraisal Methods : An International Perspective for Property Valuers*, pp. 280-303. Oxford: Wiley-Blackwell

de Graaff, T., Florax, R.J.G.M., Nijkamp, P. and Reggiani, A. (2001), "A General Misspecification Test for Spatial Regression Models: Dependence, Heterogeneity, and Nonlinearity", *Journal of Regional Science*, Vol. 41 No. 2, pp. 255-276.

Des Rosier, F. and Thériault, M. (2008), "Mass appraisal, hedonic price modelling and urban externalities: understanding property value shaping processes", in Kauko, T and d'Amato, M (Eds) *Mass Appraisal Methods: An International Perspective for Property Valuers*, pp. 1-24. Oxford: Wiley – Blackwell.

Do, A.Q. and Grudnitski, G. (1992), "A neural network approach to residential property appraisal", *The Real Estate Appraiser,* Vol. 58 No. 3, pp. 38-45.

Dormann, C.F. *et al*. (2007), "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review", *Ecography*, Vol. 30, pp. 609-628.

Dubin, R., Pace, K.R. and Thibodeau, T.G. (1999), "Spatial autoregression techniques for real estate data", *Journal of Real Estate Literature*, Vol. 7 No. 1, pp. 79-95.

Elhorst, J.P. (2010), "Applied spatial econometrics: raising the bar", *Spatial Economic Analysis*, Vol. 5 No. 1, pp. 9-28.

Evans, A., James, H. and Collins, A. (1992), "Artificial neural networks: an application to residential valuation in the UK", *Journal of Property Valuation and Investment,* Vol. 11 No. 2, pp. 195-204.

Farber, S. and Yeates, M. (2006), "A comparison of localised regression models in a hedonic house price context", *Canadian Journal of Regional Science*, Vol. 29 No. 3, pp. 405-420.

Feng, Y. and Jones, K. (2015), “Comparing multilevel modelling and artificial neural networks in house price prediction” in *2nd IEEE International Conference on Spatial Data Mining and Geographical Knowledge Services (ICSDM 2015): Proceedings of a meeting held 8-10 July, Fuzhou, China.* Institute of Electrical and Electronics Engineers (IEEE), Vol. 1 No. 15, pp. 108-114.

Fortin, M.J. and Dale, M.R.T. (2005) *Spatial Analysis – A Guide for Analysis*, Cambridge University Press, Cambridge.

Fotheringham, A.S., Crespo, R. and Yao, J. (2015), “Exploring, modelling and predicting spatiotemporal variations in house prices”, *The Annals of Regional Science*, Vol. 54 No. 2, pp. 417-426.

Ghanty, P., Paul, S. and Pal, N.R. (2009), “NEUROSVM: an architecture to reduce the effect of the choice of kernel on the performance of SVM”, *Journal of Machine Learning Research*, Vol. 10, pp. 591-622.

Gloudemans, R.J. (1990), “Adjusting for time in computer assisted mass appraisal”, *Property Tax Journal*, Vol. 9, pp. 83-99.

Gloudemans, R.J. (1999), *Mass Appraisal of Real Property*, Chicago, IL. International Association of Assessing Officers.

Gloudemans, R.J. (2002), “Comparison of three residential regression models: additive, multiplicative, and nonlinear”, *Assessment Journal*, Vol. 9 No. 4, pp. 25-36.

Gonzalez, S.A.M. and Formoso, C.T. (2006), “Mass appraisal with genetic fuzzy rule Based systems”, *Property Management*, Vol. 24 No. 1, pp. 20-30.

Goodman, A.C. (1978), “Hedonic prices, price indices and housing markets”, *Journal of Urban Economics*, Vol. 5 No. 4, pp. 471-484.

Greene, H. W. (2003), *Econometric Analysis*, Fifth Edition, Prentice Hall, Upper Saddle River, New Jersey.

Griffith, D.A. (2008), “Spatial-filtering-based contributions to a critique of geographically weighted regression (GWR)”, *Environment and planning A*, Vol. 40 No. 11, pp. 2751-2769.

Guan, J., Zurada, J. and Levitan, A.S. (2008), “An adaptive neuro-fuzzy inference system based approach to real estate property assessment”, *Journal of Real Estate Research*, Vol. 30 No. 4, pp. 395-420.

Haining, R. (2003), *Spatial Data Analysis- Theory and Practice*, Cambridge, Cambridge University Press.

Hayrullahoğlu, G., Aliefedioğlu, Y., Tanrivermis, H. and Hayrullahoğlu, A.C. (2018), “Estimation of the hedonic valuation model in housing markets: the case of Cukurambar region in Cakanya of district of Ankara”, *Ecoforum*, Vol. 7 No. 1, pp.14

Henry, M.K.M., Patrick, P.K.C. and Yiu-Sun, C. (1995), “A hedonic price model for private properties in Hong Kong”, *Journal of Real Estate Finance and Economics*, Vol. 10 No. 1, pp. 37-48.

Hornik, K. (1991), “Approximation capabilities of multilayer feed-forward networks”, *Neural Networks*, Vol. 4 No. 2, pp. 251-257.

Huang, B., Wu, B. and Barry, M. (2010), "Geographically and temporally weighted regression for modelling spatio-temporal variation in house prices", *International Journal of Geographical Information Science*, Vol. 24 No. 3, pp. 383-401.

Janssen, C. and Söderberg, B. (1999), "Estimating market prices and assessed values for income properties", *Urban Studies*, Vol. 36 No. 2, pp. 359-396.

Kauko, T. (2003), "On current neural network application involving spatial modelling of property prices", *Journal of Housing and the Built Environment*, Vol. 18 No. 2, pp. 159-181.

Kestens, Y., Thériault, M. and Des Rosier, F. (2006), "Heterogeneity in hedonic modelling of house prices: looking at buyers' household profiles", *Journal of Geographical Systems*, Vol. 8 No. 1, pp. 61-96.

Kilpatrick, J. (2011), "Expert systems and mass appraisal", *Journal of Property Investment and Finance,* Vol. 29 Nos. 4/5, pp. 529-550.

Kissling, W.D. and Karl, G. (2008), "Spatial autocorrelation and the selection of simultaneous autoregressive models", *Global Ecology and Biogeography,* Vol. 17 No. 1, pp. 59-71.

Krause, A.L. and Bitter, C. (2012), "Spatial econometrics, land values and sustainability: trends in real estate valuation research", *Cities*, Vol. 29, pp. S19-S25.

Kryvobokov, M. (2004), "Urban land zoning for taxation purposes in Ukraine: possible methods under an immature land market", *Property Management*, Vol. 22 Nos. 3/4, pp. 214-229

Kryvobokov, M. and Wilhelmsson, M. (2007), "Analysis location attributes with the hedonic model for apartment prices in Donetsk, Ukraine", *International Journal of Strategic Property Management*, Vol. 11 No. 3, pp. 157-178.

Kwok, T.Y. and Yeung, D.Y. (1997), "Constructive algorithms for structure learning in feedforward neural networks for regression problems", *IEEE Transactions on Neural Networks*, Vol. 8 No. 3, pp. 630-645.

Lam, K.C., Yu, C.Y. and Lam, K.Y. (2009), "Support vector machine and entropy based decision support system for property valuation", *Journal of Property Research,* Vol. 26 No. 3, pp. 213-233.

Lancaster, K. (1966), "A new approach to consumer theory", *The Journal of Political Economy,* Vol. 74 No. 2, pp. 23-29.

Lenk K., Worzala, E.M. and Silva, A. (1997), "High-tech valuation: should artificial neural networks bypass the human valuer?" *Journal of Property Valuation and Investment*, Vol. 15 No. 1, pp. 8-26.

Limsombunchai, V., Gan, C. and Lee, M. (2004), "House price prediction: hedonic price model vs artificial neural network", *American Journal of Applied Sciences,* Vol. 1 No. 3, pp. 193-201.

Lin, C.C. and Mohan, S.B. (2011), "Effectiveness comparison of the residential property mass appraisal methodologies in the USA", *International Journal of Housing Markets and Analysis*, Vol. 4 No. 3, pp. 224-243.

Masters, T. (1993), *Practical Neural Network Recipes in C++*, Academic Press, Boston, MA. McGraw Hill.

McCluskey, W.J. (2016), "Real property taxation in the republic of Kazakhstan", *Land Tenure Journal,* Vol. 2, pp. 119-138.

McCluskey, W. and Anand, S. (1999), "The application of intelligent hybrid techniques for the mass appraisal of residential properties", *Journal of Property Investment and Finance,* Vol. 17 No. 3, pp. 218-238.

McCluskey, W.J. and Borst, R.A. (2011), "Detecting and validating residential housing submarkets: a geostatistical approach for use in mass appraisal", *International Journal of Housing Market and Analysis*, Vol. 4 No. 3, pp. 290-318.

McCluskey, W., Davis, P., Haran, M., McCord, M. and McIlhatton, D. (2012), "The potential of artificial neural networks in mass appraisal: the case revisited", *Journal of Financial Management of Property and Construction*, Vol. 17 No. 3, pp. 274-292.

McCluskey, W.J., McCord, M., Davis, P.T., Haran, M. and McIlhatton, D. (2013), "Prediction accuracy in mass appraisal: a comparison of modern approaches", *Journal of Property Research,* Vol. 30 No. 4, pp. 239-265.

McCluskey, W.J. (1997), "A critical review of computer assisted mass appraisal techniques", in McCluskey, W.J. and Adair, A.S. (Eds.) Computer assisted mass appraisal: An international review, Ashgate, London, pp. 1-25.

McGreal, S., Adair, A., McBurney, D. and Patterson, D. (1998), "Neural networks: the prediction of residential values", *Journal of Property Valuation and Investment,* Vol. 10 No. 1, pp. 57-70.

Militino, A.F., Ugarte, M.D. and Garcia-Reinados, L. (2004), "Alternative models for describing spatial dependence among dwelling selling prices", *Journal of Real estate Finance and Economics*, Vol. 29 No. 2, pp. 193-209

Mueller, J.M. and Loomis, J.B. (2008), "Spatial dependence in hedonic property models: do different corrections for spatial dependence results in economically significant difference in estimated implicit prices?", *Journal of Agricultural and Resource Economics,* Vol. 33 No. 2, pp. 212-231.

Nakaya, T., *et al.* (2016), GWR.09 user manual. Access online via: *https://raw.githubusercontent.com/gwrtools/gwr4/master/GWR4manual_409.pdf*

Nguyen, N. and Cripps, A. (2001), "Predicting housing values: a comparison of multiple regression analysis and artificial neural networks", *Journal of Real Estate Research*, Vol. 22 No. 3, pp. 313-336.

Orford, S. (2000), "Modelling spatial structures in local housing market dynamics: a multilevel perspective", *Urban Studies,* Vol. 37 No. 9, pp. 1643-1671.

Osland, L. (2010), "An application of spatial econometrics in relation to hedonic house price modelling", *Journal of Real Estate Research*, Vol. 32 No. 2, pp. 289-320.

Osuna, E., Freund, R. and Girosi, F. (1997), "Training support vector machines: an application to face detection", *Proceedings of Computer Vision and Pattern Recognition*, June 17–19 Puerto Rico, pp. 130-136.

Pace, R.K. and Gilley, O.W. (1997), "Using the spatial configuration of the data to improve estimation", *Journal of Real Estate Finance and Economics*, Vol. 14 No. 3, pp. 333-340.

Pace, R.K., Barry, R., Clapp, J.M. and Rodriquez, M. (1998a), "Spatiotemporal autoregressive models of neighborhood effects", *Journal of Real Estate Finance and Economics*, Vol. 17 No. 1, pp. 15-33.

Pace, R.K., Barry, R. and Sirmans, C.F. (1998b), "Spatial statistics and real estate", *Journal of Real Estate Finance and Economics*, Vol. 17 No. 1, pp. 5-13

Páez, A., Uchida, T. and Miyamoto, K. (2001), “Spatial association and heterogeneity issues in land price models”, *Urban Studies*, Vol. 38 No. 9, pp. 1493-1508.

Páez, A., Long, F. and Farber, S. (2008), “Moving windows approaches for hedonic price estimation: an empirical comparison of moving techniques”, *Urban Studies*, Vol. 45 No. 8, pp. 1565–1581.

Peterson, S. and Flanagan, A.B. (2009), “Neural network hedonic pricing models in mass real estate appraisal”, *Journal of Real Estate Research*, Vol. 31 No. 2, pp. 147-164.

Platt, J.C. (1998), “Fast training of support vector machines using sequential minimal optimization”, in *B Scolkopf, C Burges and A Simola (eds), Advances in Kernel Methods: Support Vector Machines*, Cambridge, MA: MIT Press.

Raymond, T.Y.C. and Peter, E.D.L. (2000), “Measuring residential property values in Hong Kong”, *Property Management*, Vol. 18 No. 5, pp. 366-374.

Rosen, S. (1974), “Hedonic prices and implicit markets, product differentiation in pure competition”, *Journal of Political Economy*, Vol. 82, pp. 218-233.

Shimizu, C. (2014), “Estimation of hedonic single-family house price function considering neighborhood effect variables”, *Sustainability*, Vol. 6, pp. 2946-2960.

Shin, K.Y., Lee, T.S. and Kim, H.J. (2005), “An application of support vector machines in bankruptcy prediction model”, *Expert Systems with Applications*, Vol. 28 No. 1, pp. 127-135.

Stevenson, S. (2004), “New empirical evidence on heteroscedasticity in hedonic housing models”, *Journal of Housing Economics*, Vol. 13, pp. 136-153.

Tay, F.E.H. and Cao, L. (2001), “Application of support vector machines in financial time series forecasting”, *Omega*, Vol. 29, pp. 309-317.

Tay, D.P.H. and Ho, D.K.K. (1992), “Artificial intelligence and the mass appraisal of residential apartment”, *Journal of Property Valuation and Investment,* Vol. 10, pp. 525-540.

Theriault, M., Des Rosier, F., Villeneuve, P. and Kestens, Y. (2003), “Modelling interactions of location with specific value of housing attributes”, *Property Management*, Vol. 21 No. 1, pp. 25-48.

Tobler, W. (1979), “Celluar Geography”, in S. Gale and G. Olsson (eds.), *Philosophy in Geography*. Dordrecht.

Valente, J., Wu, S., Gelfand, A. and Sirmans, C.F. (2005), “Apartment rent prediction using spatial modelling”, *Journal of Real Estate Research*, Vol. 27 No. 1, pp. 105-136.

Vapnik, V.N. (1999), “An overview of statistical learning theory”, *IEEE Transaction on Neural Networks*, Vol. 10 No. 5, pp. 988-999.

Vincent, P. and Bengio, Y. (2000), “A neural support vector network architecture with adaptive kernels”, *in proceedings of IJCNN*, Vol. 5, pp. 187–192.

Wall, M.M. (2004), “A close look at the spatial structure implied by the CAR and SAR models”, *Journal of Statistical Planning & Inference*, Vol. 121, pp. 311-324.

Wheeler D.C. (2007), “Diagnostic tools and a remedial method for collinearity in geographically weighted regression”, *Environment and Planning A*, Vol. 39 No. 10, pp. 2464-2481.

Wilhelmsson, M. (2002), “Spatial Models in Real Estate Economics”, *Housing, Theory and Society*, Vol. 19, pp. 92-101.

Wiering, M.A., van der Ree, M.H., Embrechts, M.J, Stollenga, M.F., Meijster, A, Nolte., A. and Schomaker, L.R.B. (2013), "The neural support vector machine", in *Proceedings of the 25th Benelux Artificial Intelligence Conference (BNAIC)*, November 7- 8, pp. 247-254.

Worzala, E.M., Lenk, M.M. and Silva, A. (1995), "An exploration of neural networks and its application to real estate valuation", *Journal of Real Estate Research,* Vol. 10 No. 2, pp. 185-202.

Yacim, J.A. and Boshoff, D.G.B. (2018a), "Combining BP with PSO algorithms in weights optimisation and ANNs training for mass appraisal of properties", *International Journal of Housing Markets and Analysis*, Vol. 11 No. 2, pp. 290-314.

Yacim, J.A. and Boshoff, D.G.B. (2018b), "Impact of artificial neural networks training algorithms on accurate prediction of property values", *Journal of Real Estate research,* Vol. 40 No. 3, pp. 375-418.

Yacim, J.A. and Boshoff, D.G.B. (2019), "A Comparison of bandwidth and kernel function selection in geographically weighted regression for House Valuation", *International Journal of Technology*, Vol. 10 No. 1, pp. 58-68.

Yang, Z. (2001), "An application of the hedonic price model with uncertain attribute: the case of the People's Republic of China", *Property Management*, Vol. 19 No. 1, pp. 50-63.

Zurada, J., Levitan, A.S. and Guan, J.A. (2011), "Comparison of regression and artificial intelligence methods in a mass appraisal context", *Journal of Real Estate Research,* Vol. 33 No. 3, pp. 349-387.

**Appendix**

Locational dummies for the OLS

| Location | Coefficients | *t* |
|---|---|---|
| Submkt48 | -3000680 | -7.57322* |
| Submkt50 | -2521550 | -13.361* |
| Submkt52 | -1585220 | -11.1429* |
| Submkt53 | -1369620 | -8.93155* |
| Submkt55 | 4625470 | 23.8114* |
| Submkt56 | 1656350 | 8.93704* |
| Submkt64 | -2559500 | -3.62303* |
| Submkt65 | -2272260 | -1.32561 |
| Submkt66 | -999663 | -3.94929* |
| Submkt67 | -1060610 | -6.37898* |
| Submkt68 | -527519 | -2.66905* |
| Submkt69 | -2440050 | -5.91442* |
| Submkt70 | -2767070 | -8.63609* |
| Submkt73 | -2087930 | -1.48964 |

* depicts significant at 95%