EFFICACY IN MODELLING LOCATION WITHIN THE MASS APPRAISAL PROCESS

TONY LOCKWOOD and PETER ROSSINI
University of South Australia

ABSTRACT

Although Geographic Information Systems (GIS) has long been recognised as a natural partner in the computer assisted mass appraisal (CAMA) process, it has not always been clear how CAMA management (practitioners) may be able to utilise this partnership to produce regular re-assessments at an acceptable level of accuracy and cost. The objective of this study is to help demonstrate how this may occur by examining the accuracy generated by various simple, transparent and cost effective approaches traditionally used to model location as part of the CAMA process and compare the accuracy of the predicted result to that generated by using an integrated GIS environment to model location. Models were constructed to account for 'location' in three ways. They are firstly, in an a priori fashion based on established suburb and post code administrative boundaries. Secondly, by utilising the GIS to generate location factors based on the residuals of location 'blind' global hedonic models and creating an interpolated location factor surface that can be applied to global hedonic models to give a predicted value. Finally, by using hedonic Geographically Weighted Regression (GWR) that allows the regression coefficients to vary across geographic space in response to local variation. These last two approaches take advantage of the parcel's spatial coordinates to model location within a GIS environment. All three approaches are used to generate values using available secondary data normally collected as part of the determination of capital market value integrated within the spatial framework of a digital cadastre. The results indicate an acceptable degree of accuracy can be achieved when using basic hedonic GWR models that account for location in an intuitively simple way, thus providing transparency and efficiency to the mass appraisal process. GWR accounts for location giving comparatively more accurate results at little or no extra cost.

Keywords: Mass appraisal, geographic information systems, location, geographically weighted regression

INTRODUCTION

Mass appraisal is a unique valuation process that, of its very nature, must account for the locational aspects of property value in a way that preserves the relativity of property value, ensuring the resulting property tax is distributed in a fair and equitable manner. This must be done across an entire jurisdiction (often hundreds of thousands of properties) at a single point in time and be capable of being repeated on a regular cycle (optimally just prior to the levying of the tax - often annually).

The constraints on jurisdictions responsible for maintaining a property taxation base are broadly the costs associated with the collection and maintenance of appropriate data; an environment in which these data, potentially from a variety of data custodians, can be integrated and the professional skill sets needed to analysis and interpret the data. Although the use of integrated GIS environments is not new and has long been recognised as being able to make a logical contribution to the mass appraisal process (O'Connor and Eichenbaum 1988; Ward et al. 1999; Ward 2006), the ability to produce a re-assessment, in a cost effective manner, across an entire jurisdiction is now becoming a reality. Ward et al (1999) concluded there were a number of stages through which GIS may progress in making a contribution to the CAMA process. From the basic exploratory analysis through to sophisticated spatial analysis estimating the value due to location, different jurisdictions will be different stages along this evolutionary path. All jurisdictions in Australia recognise the importance of the spatial enablement of the CAMA process, but to date have used it to a limited extent and mainly in the area of thematic mapping of value changes as a result of the reappraisal process to identify potential problem areas. The advances made in the spatial analytical powers of GIS now provide practitioners with powerful tools enhancing the long established partnership between CAMA and GIS. This paper examines one aspect of this; namely the ability of GWR to generate a mass appraisal.

The 'locational' aspects of property value have historically been accounted for by selecting comparable sales that minimise the difference between the market and the subject by selecting only sales within the same location as the subject property. Where such differences are minimal, then very little adjustment may be necessary to infer value of the subject from the market. However, this is not always possible, and even in circumstances where it is, some important information may be lost by doing so. The challenge therefore becomes one of understanding the effect of location on the real estate market.

There are a wide range of mass appraisal techniques used by many jurisdictions around the world. These have been well discussed (O'Connor and Eichenbaum 1988; Ward, Weaver et al. 1999; Ward 2006). Often differences reflect the differing context in which a particular jurisdiction operates and clearly the notion that 'one size fits all' is not appropriate in this field. By their very nature, mass appraisal systems can be data hungry as they try to account for the wide variety of variables that have been found to contribute to property value; some are jurisdiction specific while others not (Kauko and d'Amato 2008). There has also been a wide range of modelling techniques examined in the literature to account for location. Some are discussed later, but it is the practical, operational question of how to deal with the complexity of location in a simple and transparent manner capable of being easily understood by practitioners and explained to taxpayers that may be answered by using the integrated GIS environment and the comparatively new geographically weighted regression (GWR) technique that can now be readily incorporated within it.

The objective of this study is to compare the accuracy of various basic models each accounting for 'location' in a simple but contrasting manner in order to understand the level of complexity and data requirement for an accurate model capable of confidently providing a fair and equitable property tax base.

In this study, accounting for 'location' is examined using three approaches, namely:

1. assessing properties with separate models within existing 'a priori' spatial boundaries (post codes and suburbs) effectively holding location constant within small submarkets.

2. establishing a value surface by creating Location Factors (LF). This surface is created from an hedonic global OLS model using independent variables that may be considered 'blind to location' and using the residual as a proxy for location.

3. creating hedonic geographically weighted regression models at each data point allowing regression coefficients to spatially vary across geographic space, thus reflecting local variation.

Each adopts a simple transparent methodology capable of being utilised by jurisdictions with basic property characteristic data within a spatial framework, such as a Digital Cadastral Data Base (DCDB) allowing the spatial analysis to be conducted within a GIS environment.

The models are then tested using traditional quality assurance methods including a number of IAAO sales ratio statistics; mean and median AS ratio, the COV and COD sales ratio statistics and PRD (Adair, Berry et al. 2000; Chhetri, Stimson et al. 2006), as well as the more broadly adopted FSD and 'hit rates'. These are calculated for each of the models as well as for the actual assessed values used by the local valuation authority and compared to benchmark standards suggested by Eckert (1990) and applied these to both the sales data themselves (in-sample testing) and to a holdout sample.

The study area was the metropolitan area of Adelaide in South Australia containing approximately 440,000 residential properties as shown in Figure 1. Adelaide is a geographically isolated but active market with over 20,000 residential transactions each year.

LITERATURE REVIEW

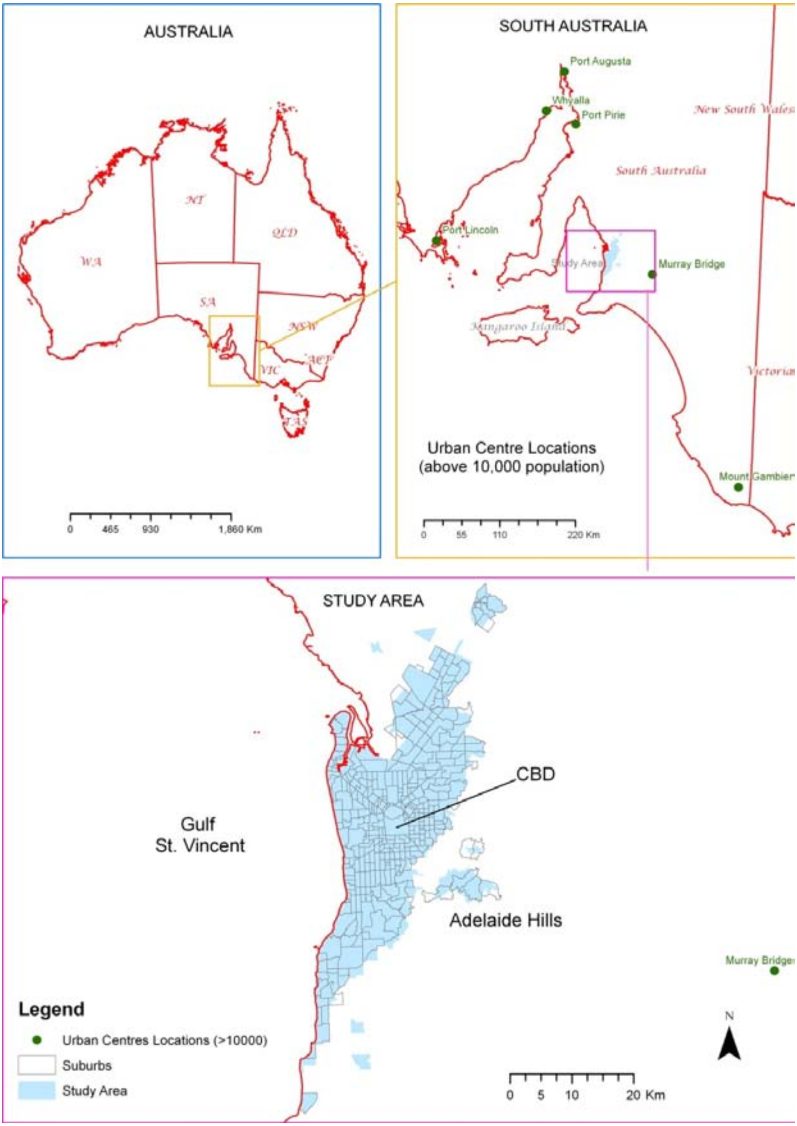
The unique nature of property as a heterogeneous, immovable commodity gives location a special significance in the market place (Rossini and Kershaw 2008). Although it is recognised that the property market is made up of a series of interrelated sub markets, there is no consensus as to the definition or delineation of these sub markets (Galster 1996). It may be as Bourassa et al.(2007) suggest that the size of the sub market is rather dependent upon the application it is to be used for. For mass appraisal purposes, smaller spatial units are likely to be more homogeneous housing markets than larger ones and hence likely to require less complex models in order to produce accurate results. The question as to how sub market segregation should be defined so as to reflect the importance of location in mass appraisal modelling has been approached in several different ways.

Firstly, perhaps the most cost effective, simplest and quickest approach is to use existing *a priori* administrative boundaries such as suburbs or postcodes. Although this is not an attempt to find optimal sub markets as much as it was to confirm that smaller areas did exist that were each more homogeneous than the global market and within which a separate global model can be calibrated using only sales transaction data from that smaller geographical area.

A second approach is to include the 'location' as part of the hedonic modelling process giving each property a unique 'location factor', thus removing the need to be specifically aware of sub market boundaries(Adair et al. 1996; Goodman and Thibodeau 2003; Ming and Siu 2003). This concept is attractive in that it recognises each property as its own submarket and together portrays location as a continuous geographic surface.

Clapp (2003) suggests surfaces generated from low order polynomial expansions value surfaces may not be sufficient to capture an 'arbitrarily flexible value surface' and to use sufficiently high order expansion may be problematic in terms of collinearity and loss of degrees of freedom.

Figure 1: Study area



An earlier study by Gallimore et al.(1996) took the approach of using the residual of an hedonic global model created to be deliberately ‘blind’ to location as a proxy for a location factor. This is a simple approach that examines ‘location’ through the residuals, allowing a quick method of viewing the global spatial distribution of the effect on value that may be due to location. Calculating the ratio of the actual price to the estimated value from such global models, a location factor may be determined for each of the sale properties which when interpolated across the study area can give a value surface that can be applied to any property in the study area. By not relying on a trend surface, it can more accurately reflect the actual value surface as shown by that particular sales sample.

Also included in this group is the use of geographically weighted regression (GWR) that allows the hedonic regression coefficients to vary depending on location. There is no assumed stationarity in this model. It is intuitively sensible and has been used to investigate the effect of location in real estate markets (Gallimore et al. 1996). The advantage in the mass appraisal process is that the local variation is captured where it exists and as with other approaches in this group is not reliant upon non market related definitions of sub market boundaries.

The objective of this study is to therefore compare the accuracy of GWR with the *a priori* approach and the location factor derived from the residual of a global model as representing three transparent and cost effective methods of accounting for location in the mass appraisal process.

METHODOLOGY

Data

In adopting the three modelling approaches, the data used was taken from the 2009 valuation list data of the South Australian Valuer General. The data used in this study is summarised as Table 1.

Table 1: Summary of data used in study

Variable	Type	Description
Suburb and postcode	nominal	Suburb name and 4 digit postcode used to split models into these a-priori groups
Coordinate values	continuous	X,Y coordinate - false eastings and northings using GDA_1994_MGA_Zone_54
Sale Price	continuous	Sale Price in dollars
Dwelling size (DS)	continuous	Equivalent main area in square metres
Dwelling age (DA)	continuous	Age in years
Dwelling land area (LA)	continuous	Area in square metres taken from the digital cadastra
Dwelling type DETACHED (D_1)	Dummy	1 if DETACHED else 0
Dwelling wall construction BRICK (D_2)	Dummy	1 if BRICK else 0
Dwelling wall construction STONE (D_3)	Dummy	1 if STONE else 0
Dwelling QUALITY (D_4)	Dummy	1 if HIGH else 0

Model calibration and testing used two sales data sets. The first was used for model calibration and contained approximately 12,000 sales transactions completed between 1/10/08 and 1/4/09. Model calibration and in-sample testing was carried out using this data set. A second sales data set containing approximately 4,000 sales that had occurred between 1/4/09 and 1/7/09 were used as a hold-out data set and used to undertake the out of sample accuracy testing. No adjustment to price due to date of sale has been made on either data set as the market was deemed stable over the respective time periods. In both data sets, sales were those that had occurred in the study area during the specified time periods but omitting only those not deemed to represent market value. There was deliberately no screening of sales purely on the basis of them being 'outliers' so as to make the accuracy testing as being unbiased as possible from a valuation perspective.

Model specification

Three main model specifications were used to allow for location and in each instance are specified with 3 or 7 independent variables; first with the 3 continuous independent variables and the second with these 3 plus 4 dummy variables.

The first set of models use linear multiple regression models at different levels of market segmentation. The second set uses a single global model with a land value surface and the third uses the GWR approach with models created at each data point. Although the assessment industry utilises non-linear modelling, this study compares only the efficacy of linear models as this is currently the only readily accessible GWR model accessible to the practitioner. In each case, these models are calibrated and then tested against both the in-sample and out-of-sample sales data sets. Their construction is described below and summarised in Table 2 below.

Global, suburb and post code models

The global models were constructed over the whole study area using the three and seven independent variables respectively making no allowance for location. It is expected that these models will not produce acceptable results in terms of the adopted standards as no allowance is made for location but provide a benchmark to observe any improvements in subsequent models accounting for location.

The *a priori* segregation of the study area into suburbs to create smaller geographic areas provides more homogeneous regions in which increased accuracy may be achieved with often small amounts of data typically found in many jurisdictions. There are 307 suburbs within the study area and separate linear models are calibrated for each. Similarly, *a priori* geographical segregation into postcode districts determined by Australia Post providing a more homogeneous grouping of residential properties than taking the study area as a whole. There are 206 postcodes within the study area. Postcodes are generally geographically larger than suburbs and thus contain larger sale samples than suburbs, yet still provide the level of homogeneity needed for the better performance of the less complex hedonic models. As in the case of the suburb modelling, the separate models are in a linear form based on previous studies in the city (Rossini 2006 and 2008) which showed that linear and log-linear forms produced very similar results.

Global, Suburb and Postcode - 3 variable equations (GL_3var, Sub_3var and PC_3var)

$$\text{SalePrice}_i = \beta_{0(l)} + \beta_{1(l)} DS_i + \beta_{2(l)} DA_i + \beta_{3(l)} LA_i + \varepsilon_l \qquad \textbf{(1)}$$

Global, Suburb and Postcode - 7 variable equations (GL_7var, Sub_7var and PC_7var)

$$\text{SalePrice}_i = \beta_{0(l)} + \beta_{1(l)} DS_i + \beta_{2(l)} DA_i + \beta_{3(l)} LA_i + \beta_{4(l)} D_{1i} ... \beta_{7(l)} D_{4i} + \varepsilon_l \quad \textbf{(2)}$$

where:

SalePrice$_i$ = the sale price of the ith property
β$_0$ to β$_7$ = the parameter estimates
(l) = the location indicator of the ith property – either global, one of 307 suburbs or one of 206 postcodes
DS$_i$ = the dwelling size for the ith property·
DA$_i$ = the dwelling age for the ith property·
LA$_i$ = the Land Area for the ith property·
D$_{1i}$ to D$_{4i}$ = dummy variables for Detached, Brick, Stone and Quality for the ith property·
ε$_{(l)}$ = stochastic error for location (l)

Location factor models

The location factor models use an interpolated location factor surface to account for the locational variation. In a two stage approach, location factors are first estimated from the 3 variable global hedonic model ((1) which is considered to be 'blind' to location. The value due to location is then deemed to be the residual when predicting back the sale price with the factor determined by dividing the sale price by the predicted value and may be represented as (3. These location factors are then smoothed using the Inverse Distance Weighted (IDW) interpolator. It is from this continuous location factor surface that a factor is assigned to each of the sale properties in both the in sales sample and in the hold out sample. The derived location factors are predicted and applied in the 3 variable location factors models ((5). This process is repeated using the 7 independent variables in stages 1 and 2 – see (4 and (6.

Stage One Location Factor Model - 3 variable equation

$$lf_i = SalePrice_i / \beta_0 + \beta_1 DS_i + \beta_2 DA_i + \beta_3 LA_i + \varepsilon \quad \textbf{(3)}$$

Stage One Location Factor Model - 7 variable equation

$$lf_i = SalePrice_i / \beta_0 + \beta_1 DS_i + \beta_2 DA_i + \beta_3 LA_i + \beta_4 D_{1i} ... \beta_7 D_{4i} + \varepsilon \quad \textbf{(4)}$$

Location Factor - 3 variable equation (LF_3var)

$$\text{SalePrice}_i = L\hat{F}_i \left(\beta_0 + \beta_1 DS_i + \beta_2 DA_i + \beta_3 LA_i + \varepsilon \right) \quad \textbf{(5)}$$

Location Factor – 7 variable equation (LF_7var)

$$SalePrice_i = L\hat{F}_i \left(\beta_0 + \beta_1 DS_i + \beta_2 DA_i + \beta_3 LA_i + \beta_4 D_{1i} ... \beta_7 D_{4i} + \varepsilon \right) \quad \textbf{(6)}$$

where:

$SalePrice_i$ = the sale price of the ith property
lf_i = the estimated location factor for the ith property
LF_i = the derived location factor for the ith property from the value surface
β_0 to β_7 = the parameter estimates
DS_i = the dwelling size for the ith property·
DA_i = the dwelling age for the ith property·
LA_i = the Land Area for the ith property·
D_{1i} to D_{1i} = dummy variables for Detached, Brick, Stone and Quality for the ith property·
ε = stochastic error at location

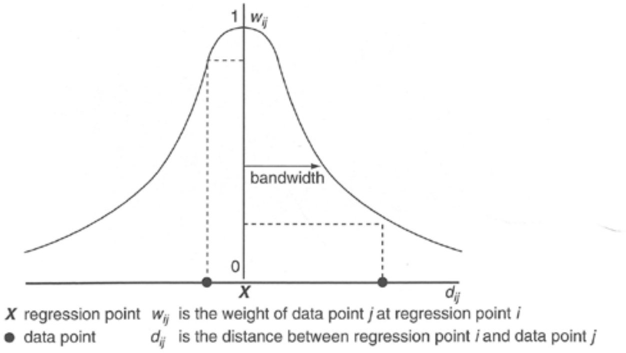
Interpretation suggests that when LF > 1.0, it may indicate that there is more variation in the sale price than indicated by the independent variables that described just the dwelling and therefore must be due to locational factors (and random error) in comparison to the average for the whole study area. Conversely, LF < 1.0 may indicate a negative contribution. The interpolated (smoothed) value surface can be used to derive a location factor for any point on the surface and is used to establish the relevant factor for both the in-sample and out-of-sample observations. This is then multiplied by the corresponding global model to generate the predicted values. This methodology recognises the residual error term comprises both random error and the error due to spatial autocorrelation. The extent to which the genuine random error is present may compromise this approach. It is the spatial autocorrelation that provides the relative expression of value change due to location and genuine *random* error should not affect this relativity.

Geographically weighted regression (GWR) models

This OLS hedonic model accounts for location by allowing the regression coefficients to vary across geographical space with the equation estimated at individual coordinates. It makes no assumption as to the non-stationarity or otherwise of the coefficients and is not limited by the artificial administrative boundaries of suburbs or postcodes as outlined in model 2.

Figure 2 shows conceptually how a weighting is applied using the GWR model. For a given property located at coordinate (u,v), a model is created where those sales closest to (u,v) are given the highest weight. The weight decreases the further away a sale is from the given regression point. This conforms with Tobler's first law (Borst and McCluskey 2008; Moore and Myers 2010) preserving the concept that events closer together are more likely to be similar than those further apart. At subsequent coordinates, the same method of weighting is applied meaning that even though different regression points may share some common sales data in their calibration process, they will be weighted differently at different points in space. It is in this way that GWR uniquely calibrates the local models as it moves across the geographical surface and can thus capture local variations not possible in global or segmentation models.

Figure 2: Concept of a spatial kernel adopted in GWR


1
w_{ij}
bandwidth
0
X
d_{ij}
X regression point w_{ij} is the weight of data point j at regression point i
● data point d_{ij} is the distance between regression point i and data point j

Source: (Fotheringham et al., 2002)

The bandwidth shown in Figure 2 can be either fixed or adaptive. In a GWR models with fixed kernels, the bandwidth does not vary as it crosses geographic space. The disadvantage of this is that in areas of low density, the local GWR model may be calibrated on the evidence of very few sales and in some cases may not have sufficient data to calibrate a model. To help overcome this problem, GWR kernels have the capability to adapt themselves in size to accommodate available data. In areas where data has low density, the bandwidth of the spatial GWR kernel increases so as to include sufficient data to calibrate the local model. Conversely, it can decrease the bandwidth where the data has high density, giving it the capability to calibrate a more local model. In this study, the GWR models used an adaptive kernel allowing the bandwidth to vary so as to incorporate enough sales data to calibrate the model (Tobler 1970).

This is achieved using an adaptive kernel with a “near-Gaussian” weighting function adapting its bandwidth to include an optimal number of nearest neighbours (n). The following bi-square decay function employed in the ESRI ArcGIS software (Fotheringham et al. 2002; Borst and McCluskey 2008) is used in this study.

$$w_{ij} = [1-(d_{ij}/b)^2]^2 \text{ if } d_{ij} < b \text{ otherwise } w_{ij}=0 \qquad \textbf{(7)}$$

where

d = the distance between
i = (model calibration point)
j = (the data point)
b = the distance beyond which w_{ij}=0

The optimal constant number of nearest neighbours to be included in the adaptive kernel in this study is determined by minimising the Akaike Information Criterion (AIC). This is one of many methods able to determine the best model structure.

$$\mathrm{AIC_c} = 2n\log_e(\ddot{\varpi})+\log_e(2\pi)+n\left\{\frac{n+tr(S)}{n-2-tr(S)}\right\} \qquad \textbf{(8)}$$

where:
n = the number of observations
$\ddot{\varpi}$ = the estimated standard deviation of the error term
tr(S) = the trace of the HAT matrix of the GWR.

Source: (Fotheringham et al. 2002).

The AIC$_c$ is a relative measure allowing comparison to be made between various models with the lower value being the preferred model.

GWR - 3 variable equation (GWR_3var)

$$SalePrice_i = \beta_{0(u,v)} + \beta_{1(u,v)}DS_i + \beta_{2(u,v)}DA_i + \beta_{3(u,v)}LA_i + \varepsilon_{(u,v)} \qquad \textbf{(9)}$$

GWR - 7 variable equation (GWR_7var)

$$SalePrice_i = \beta_{0(u,v)} + \beta_{1(u,v)}DS_i + \beta_{2(u,v)}DA_i + \beta_{3(u,v)}LA_i + \beta_{4(u,v)}D_{1i}...\beta_{7(u,v)}D_{4i} + \varepsilon_{(u,v)} \qquad \textbf{(10)}$$

where:
SalePrice$_i$ = the sale price of the ith property
(u,v) = the location coordinates of the ith property
β_0 to β_7 = the parameter estimates at location (u,v)
DS$_i$ = the dwelling size for the ith property·
DA$_i$ = the dwelling age for the ith property·
LA$_i$ = the Land Area for the ith property·
D$_{1i}$ to D$_{1i}$ = dummy variables for Detached, Brick, Stone and Quality for the ith property·
$\varepsilon_{(u,v)}$ = stochastic error at location (u,v)

Table 2: Model descriptions

Model name	Model type	Independent variables
CV2009	Assessed values - determined independently from this study using a computer assisted manual process with exception and objection adjustments. Derived by the Valuer General and used for comparison.	
GL_3var	Global hedonic OLS	Dwelling age; Dwelling equivalent main area; land area
GL_7var	Global hedonic OLS	Dwelling age; Dwelling equivalent main area; land area Dummy variables for Brick, Stone, Detached and quality.
Sub_3var	Hedonic OLS segregated by 307 Suburbs	Dwelling age; Dwelling equivalent main area; land area
Sub_7var	Hedonic OLS segregated by 307 Suburbs	Dwelling age; Dwelling equivalent main area; land area Dummy variables for Brick, Stone, Detached and quality
PC_3var	Hedonic OLS segregated by 206 Postcodes	Dwelling age; Dwelling equivalent main area; land area
PC_7var	Hedonic OLS segregated by 206 Postcodes	Dwelling age; Dwelling equivalent main area; land area Dummy variables for Brick, Stone, Detached and quality
LF_3var	Global OLS hedonic model multiplied by the'location factor' derived from a 'blind' 3 variable hedonic global hedonic OLS.	Dwelling age; Dwelling equivalent main area; land area
LF_7var	Global OLS hedonic model multiplied by the'location factor' derived from a 'blind' 7 variable hedonic global hedonic OLS.	Dwelling age; Dwelling equivalent main area; land area Dummy variables for Brick, Stone, Detached and quality
GWR_3var	Hedonic geographically weighted regression OLS over whole study area	Dwelling age; Dwelling equivalent main area; land area
GWR_7var	Hedonic geographically weighted regression OLS over whole study area	Dwelling age; Dwelling equivalent main area; land area Dummy variables for Brick, Stone, Detached and quality

Accuracy tests

The accuracy tests used in this paper to assess the various models are based upon broad accuracy tools used generally in model testing. These include the MAPE, RMSE and FSD or upon a series of accuracy indicators recommended by the IAAO in the area AVM accuracy standards as Eckert (1990) suggests. These have been used in previous studies and the tests selected are based on the findings of Rossini and Kershaw (2008) as to acceptable levels of accuracy against which comparison of results can be made. The accuracy statistics used in this study are calculated as shown in the appendix.

RESULTS AND DISCUSSION

The predicted values produced by the various calibrated models are compared with the in sample and out of sample sales using statistics determined using equations 9 to14 and presented in Table 3.

Table 3: Model evaluation results

IN SAMPLE (12,644 SALES 1/10/08 TO 1/4/09)

Evaluaton / Model namer	CV2009	GL_3 var	GL_7 var	Sub_3 var	Sub_7 var	PC_3 var	PC_7 var	GWR_3 var	GWR_7 var	LF_3 var	LF_7 var
MAPE	13.7%	25.5%	23.7%	11.3%	10.2%	12.6%	12.1%	11.4%	12.6%	12.0%	13.8%
RMSE	$ 93,077	$ 157,034	$ 148,877	$ 82,380	$ 74,186	$ 98,022	$ 93,244	$ 87,160	$ 98,903	$ 91,594	$ 91,899
Percentage within + or - 5%	17%	13%	16%	35%	38%	31%	32%	34%	31%	33%	27%
Percentage within + or - 10%	39%	26%	30%	60%	64%	55%	57%	60%	56%	58%	50%
Percentage within + or - 15%	62%	39%	44%	76%	79%	72%	73%	76%	72%	73%	66%
Percentage within + or - 20%	80%	50%	56%	85%	87%	82%	83%	85%	82%	84%	78%
Percentage within + or - 50%	99%	88%	90%	98%	99%	98%	98%	98%	98%	98%	98%
FSD	17.3%	38.7%	32.0%	20.6%	54.8%	18.9%	17.9%	20.1%	18.6%	15.4%	17.4%
Mean A/S	0.890	1.080	1.072	1.021	1.018	1.025	1.023	1.023	1.025	1.032	1.038
Median A/S	0.885	1.044	1.025	1.005	1.002	1.009	1.006	1.006	1.008	1.009	1.014
COV	14.058	30.671	30.392	16.237	14.848	17.864	17.179	16.419	17.954	17.241	18.624
COD	9.660	24.189	23.037	11.215	10.214	12.468	11.992	11.335	12.483	11.860	13.521
PRD	1.016	1.080	1.072	1.020	1.017	1.025	1.023	1.023	1.026	1.006	1.011
R - Squared	0.889	0.482	0.535	0.858	0.884	0.798	0.817	0.841	0.795	0.837	0.831

F Test (ANOVA)	101131	11779	14528	76102	96787	50033	56621	66720	48980	64749	62124

OUT of SAMPLE (4,076 SALES 1/4/09 TO 1/7/09)

Evaluaton / Model namer	CV2009	GL_3var	GL_7var	Sub_3var	Sub_7var	PC_3var	PC_7var	GWR_3var	GWR_7 var	LF_3var	LF_7var
MAPE	12.7%	24.7%	22.5%	12.0%	12.8%	11.4%	11.5%	10.4%	10.8%	15.5%	18.0%
RMSE	$ 57,036	$113,870	$ 107,479	$ 76,689	$ 89,477	$ 70,985	$ 73,718	$ 64,470	$ 65,265	$ 92,357	$ 98,432
Percentage within + or - 5%	17%	13%	16%	34%	34%	32%	33%	35%	34%	23%	19%
Percentage within + or - 10%	39%	27%	32%	60%	59%	59%	59%	63%	61%	44%	37%
Percentage within + or - 15%	63%	40%	46%	76%	74%	76%	76%	80%	78%	61%	53%
Percentage within + or - 20%	84%	52%	58%	84%	83%	86%	85%	88%	87%	74%	66%
Percentage within + or - 50%	100%	89%	91%	97%	97%	98%	98%	98%	98%	97%	96%
FSD	11.5%	27.9%	26.4%	26.4%	34.0%	19.7%	17.8%	14.5%	14.4%	19.8%	22.7%
Mean A/S	0.894	1.114	1.083	1.022	1.023	1.017	1.018	1.015	1.016	1.033	1.042
Median A/S	0.881	1.079	1.036	0.996	0.994	0.996	0.994	0.996	0.994	1.005	1.008
COV	11.757	28.008	28.314	18.114	19.954	16.349	16.801	15.253	15.525	20.902	23.139
COD	8.203	22.156	21.520	12.073	12.820	11.418	11.581	10.487	10.883	15.470	17.883
PRD	1.003	1.053	1.046	1.011	1.009	1.009	1.008	1.010	1.010	0.997	1.000
R - Squared	0.937	0.540	0.589	0.799	0.760	0.824	0.815	0.849	0.846	0.770	0.746
F Test (ANOVA)	60365	4791	5847	16196	12871	19063	17910	22966	22418	13620	11978

Generally in-sample testing will provide superior results as the models are calibrated and tested using the same data. This means that in more complex models it is possible to "overfit", meaning that the model works well for the particular data but does not generalise when applied to new data. If the models are robust, in-sample and out-of-sample testing will provide a similar outcome, but it is generally excepted that out-of-sample accuracy will still be slightly lower than that of in-sample testing. Large discrepancies between the two would suggest that the particular model attempted tends to “overfit” and will not create good overall assessments when applied more generally.

These can then be readily compared with the indicated standards of acceptability identified by Rossini and Kershaw (2008) and reproduced in Table 4.

Table 4: Accuracy standards

Absolute minimum benchmark	Reasonable level of acceptance
MAPE : 13%	MAPE 10%
50% of estimates within +- 10%	65% of estimates within +- 10%
65% of estimates within +- 15%	80% of estimates within +- 15%
80% of estimates within +- 20%	90% of estimates within +- 20%
FSD less than 19%	FSD less than 15%
COV of A/S ratios less than 17	COV of A/S ratios less than 13
COD less than 13	COD less than 10

Source: Rossini, P., Kershaw, P. Automated Valuation Model Accuracy: Some Empirical Testing. *14th Pacific Rim Real Estate Conference*. Kuala Lumpur, January 2008

The challenge of these models is to produce at least as good a result as does the current methodology shown by model ‘CV2009’ in Table 3. The two global models (GL_3var and GL_7var) were included as a benchmark to show the comparative standard achieved by models making no allowance for location in this study area and using the same sales transaction data. As expected, these global results proved non acceptable in terms of meeting the minimum accuracy standards. The MAPE and RMSE are both very large and nearly twice the level that is considered to meet even the minimum benchmark. The hit ranges are well outside of the minimum standard even in-sample with only around 50% (3 variable model) and 56% (seven variable model) being within + or - 20% of the actual sale price. The out-of-sample tests are similar, although slightly worse as expected and do at least suggest that the model generalises relatively well. A reasonable level of acceptance would have 90% of observations within the + or – 20% range. Considering the out-of-sample data, the mean AS ratio for the three variable global model shows that properties are over assessed by almost 14% on average although only 8% on median but these overestimates are improved by considering the additional variables (7 variable model) when the overestimates are only 8% on average and 4 % on median. These overestimates suggest that the vast majority of typical properties are over assessed in order for the model to deal with a small number of very expensive properties which would typically be located in central areas. This might be expected in a model that is “blind” to location. This situation is emphasised by the large price related differentials (PRD) figures. The PRD is used to determine if the assessment is generally regressive or progressive. The very high figure in this instance suggests that overall the assessments are extremely regressive, meaning that on average low priced properties are over assessed and higher priced properties are under assessed and this is symptomatic of a model which fails to deal adequately with location. The FSD, COV and COD all suggest that the distributions of assessment errors are wide and many of these are well outside even the ‘absolute minimum benchmark’ standards shown in Table 4.

The statistical analysis of the out-of-sample models shows that for the 3 variable model, 48% of the variation and for the 7 variable model 53% of the variation are explained by the respective models and that in each instance this will be statistically significant. The F test would be expected to improve with more complex models and increases in the F-test are a good indicator of improved explanatory power of the models.

The suburb models

In these models, each *a priori* group of properties is based on a simple suburb boundary that does not attempt to present optimal market segregation, but instead recognise that a smaller geographic area may well represent a more homogeneous market than the global market. In this study, several suburbs did not contain enough sales to satisfactorily calibrate a model. Amalgamation with geographic contiguous suburbs had to be undertaken to achieve this and this presents an administrative weakness in this form of simple arbitrary market segregation. Using such administrative boundaries can lead to large difference between the predicted values of properties that are close together differentiated only by being on opposite sides of such boundaries. This edge effect requires resources to smooth out and the problems that this creates are not dealt within this paper as any resulting models would clearly display this problem. The model results show that simple breaking the global model into a series of simple suburb models delivers significant improvements over the global model and that this will provide a level of locational influence in the final assessments. The in-sample accuracy tests are particularly good; however these degrade when we consider the out-of-sample results. This is to be expected because many of the models will be based on relatively small sample sizes and this is more likely to cause over-fitting. Considering the out-of-sample resulting the MAPE of 12 % for the 3 variable model and 12.8% for the 7 variable model is significantly reduced from the global model and is within the absolute acceptable level of 13% suggest by Rossini and Kershaw (2008). Notably the 7 variable results are worse than the 3 variable result which in the in-sample testing the reverse was the case where the additional variable produced superior results. This trend is also noticed in the FSD, COV and COD and suggests that the additional dummy variables are not sufficiently represented in many of the suburb models and because of this serious over fitting occurs. On this basis, only the 3 variable model will be considered further in the suburb models.

The hit ranges show that the 3 variable suburb models produce estimates that are within the absolute minimal benchmarks but not reasonable acceptable. The COD and COV figures are not quite within the minimum level. The increased r-squared and F values show significant improvements in the explanatory power of these models compared to the gobal model.

The postcode models

Conceptually, this *a priori* spatial segregation along administrative boundaries which are not necessarily aligned with market structures and is subject to similar strengths and weaknesses as the suburb. In this study area, the postcode is generally a larger geographic area than the suburb and therefore less likely to require amalgamation due to lack of market sales transaction data. The larger geographic areas meant that most models had slightly greater power and this should tend to provide more robust models that generalise better. The results reflect this with the slightly inferior in-sample tests compared to the suburb models, but superior out-of-sample results. When considering the out-of-sample tests, the 3 variable models are still superior to the 7 variable versions, again suggesting that the models may not have sufficient data to generalise with dummy variables but also supports the case for very simple models using land and building area and building age as independent variables. The models fall within the minimal standards, but not the reasonable level.

The location factor models

The two models using the derived Location Factor (LF) from both the 3 variable and the 7 variable ‘blind to location’ global models, while better predictors of value than the two ‘global’ models on their own, they are still outside the ‘absolute minimum’ acceptable benchmark limits and therefore discounted as a feasible, simple and acceptably accurate methodology for accounting for location in this study. The accuracy tests are generally inferior to the a-priori submarket models and have a serious degradation between the in-sample and out-of-sample tests. While the mean and median AS ratio and the PRD suggest no significant overall bias, the variation measures, in particular the FSD, COV and COD, show that the errors are too variable and not minimally acceptable and the simple measures such as the MAPE don’t suggest it is acceptable either. This single model uses all the data and should not suffer from the low sample size problems of the a-priori submarket models. The poor performance of the seven variable model, provides the best evidence that the simple 3 variable model is sufficient. In these models, the additional variables should never decrease performance (you cannot explain less variation by adding more variables); however in these models, the LF factors are derived separately using a IDW interpolation with its own errors associated with it.

The use of the surface will overcome the problem of boundary variation that is prevalent in the a-priori submarket models. However it appears that in this study, better overall accuracy is achieved by assuming a "flat" location structure within each submarket area even though this is likely to result in less accurate values near submarket boundaries and significant variation in value levels as we move from one submarket to the next submarket across the a-priori submarket boundary.

The GWR models

Based on a comparison of the calibration of the two GWR models, GWR_3var would be the preferred model. Table 5 contains model calibration output from the two GWR models with the two global models included for comparison. It shows there is less spatial autocorrelation in the residuals (Moran's I is much lower), the AIC$_c$ is lower indicating the better the model fits the data, and the optimal number of nearest neighbours required in the adaptive kernel is much less (only 194 as against 833 for the GWR_7var model). This makes the GWR_3var model much more of a local model likely able to capture more local variation where it exists thus optimising the objective of the study. The R-squared is about the same for both and is acceptable.

Table 5: Comparison of GWR models & two global models

Model	Global Moran's I	AIC$_c$	Adj R^2	Optimal number of nearest neighbours
GWR_3var	0.11	324,913	0.83	194
GWR_7var	0.20	327,370	0.85	833
GL_3var	0.38	338,443	0.48	
GL_7var	0.62	337,102	0.54	

Both the GWR models appear to be no weaker, in fact a little stronger, in the out of sample than in the in-sample testing. The level of the predicted values seem to be quite acceptable with the mean AS ratio at 1.02 in both models for the in-sample group and slightly better at 1.01 for the out of sample group (Table 3). This is reflected in the MAPE which in the out of sample group fell to a reasonable level of acceptance while in the in-sample group was just within the minimum level of acceptance. More particularly, the dispersion of the A/S ratio appears to be reasonable for both the COV and the COD statistic although the GWR_7var model has a COV of 17.9 for the in sample testing which is too high. However, again the statistics are slightly better in the out of sample group with the GWR_7var's COV reducing to an acceptable level of 15.5. The GWR models appear better than both the suburb post code and location factor models in terms of the comparability (expressed in the COV statistic) in the out of sample group being 15.2 to 15.5, whereas the others were above the absolute minimum benchmark. This is also reflected in the in-sample "hit rates" which show the GWR models with a higher percentage within all groups than do the suburb, postcode and location factor models.

The concept of GWR makes the results of such modelling more in tune with the market than those models using all sales, arbitrarily included in an administrative boundary and all given the same weight. GWR alleviates the edge effect of sudden jumps in predicted value crossing the artificial administrative boundary. In theoretical terms, the GWR models are more intuitively explained and transparent than the *a priori* models. A current weakness of these models is that they require a spatial framework in which to operate and in terms of low cost simple modelling, this may be problematic in some jurisdictions. However, there is an understanding amongst Australian jurisdictions that such a spatially enabled CAMA environment is not only desirable but essential.

Amongst the 10 models presented, the two GWR models exhibit higher levels of accuracy in terms of the conventional tests. Each have higher percentages in the various accuracy categories and lower coefficients of dispersion and variation making them preferred models. In addition, the quality assurance phase of the CAMA cycle can provide understanding of the GWR models behaviour in terms of potential strengths and weakness across geographic space, by providing plots of the local variation in R-squared and the local significance of chosen independent variables that can be displayed, giving all stakeholders more confidence in the outcome.

The accuracy of the various models tested would be improved in reality as normal exception reporting procedures and the ratepayer objection process would enhance model accuracy. This has not been done in this study. An overall summary of the better models in this study is provided in Table 6.

Table 6: Model summary

MODEL (out of sample)	**Summary comment**
Suburb	Lack of sales in some suburbs – arbitrary amalgamation Edge effect
Post Code	Edge effect – larger geographic areas not necessarily market orientated
Location Factor models	Accuracy results generally not as good as the suburb or postcode *a priori* models nor GWR models Overcomes the edge effect
GWR models	Acceptable results Best model using only 3 simple property characteristics Overcomes the edge effect The 3 variable model is the most local of models providing the most parsimonious, accurate and cost effective solution.

CONCLUSION

The objective of this study was to ascertain if the complexity of location could be satisfactorily accounted for as part of the CAMA process using cost effective, easily understood and transparent modelling. This study has demonstrated that it may be achieved using hedonic GWR modelling, which simply constructed can include the complexity of the effect of 'location' on value to acceptable standards as part of the mass appraisal process. This exploits the natural and long established partnership between GIS and CAMA by utilising the coordinates of the land parcels as location variables.

This study examined three approaches that may account for location in a CAMA process. It found that using *a priori* spatial segmentation (suburbs and postcodes) could achieve minimal standards without the benefit of checking outliers. The advantage of this approach is that they are simple to establish as they pre-exist in the community and rely on smaller geographic areas being more homogeneous than larger areas. It makes no claim that they represent optimal submarkets and the extent to which they do not may compromise the model. The disadvantage lies in the assumption that there is no local variation across the suburb or postcode. This can lead to sudden differences in value between two similar properties adjacent to each other, but with the administrative boundary between them. This is often termed the 'boundary problem' or 'edge effect'. Another drawback lies in the number of available sales for model calibration. If there are insufficient sales in one area, it may require amalgamation of adjacent areas. This leads to larger and hence more global predictive models compromising the accuracy of the predicted values. The advantage is the simplicity in the identification of smaller more homogenous in which simplier models may be specified.

This study found using the location factor models were unsatisfactory in both the in-sample and out of sample testing; particularily when measuring the dispersion of the A/S ratio and as this the most critical aspect of ensuring equitable tax distribution, these models could not be considered in this study. The GWR model on the other hand does. Both the COV and the COD were acceptable, especially in the out of sample testing. One advantage of the GWR models lies in the continuous nature of the predicted value surface that results from the calibrated model. This overcomes the 'boundary problem' or edge effect generated by the *a priori* suburb and post code models. The difference between the two GWR models demonstrates the more local GWR model (3 variable model) being the one with variables exhibiting the strongest concentration of variability in the variables used in calibration. The three variable model using dwelling area, age and land area required a much lower number of nearest neighbours to be included than did the seven variable model which incorporated data with not so much variation. This allowed the GWR 3 variable model to capture more local variation to be included in the resulting predicted value and hence more closely optimise the effect due to location; an objective of the study. The limitation in these models, as with any model, is in the availability of suitable data, but this study shows acceptable accuracy may be obtained from relatively standard data and achieved at minimal additional cost.

This study concludes that the GWR 3 variable model satisfactorily accounts for local variation in price in a parsimonious and cost effective manner, providing the practitioner with an easily understood model capable of transparently demonstrating to stakeholders how the CAMA modelling works. This is subject to a caveat that GWR modelling is relatively new to the mass appraisal industry and must by regarded as somewhat experimental until accepted by tribunals and/or courts in defence of resulting assessments.

REFERENCES

Adair, A., J. Berry, et al. (1996). "Hedonic Modelling, housing submarkets and residential valuation." Journal of Property Research **13**(1): 67-83.

Adair, A. S., J. N. Berry, et al. (2000). "The Local Housing System in Craigavon, N. Ireland: Ethno-religious Residential Segregation. Socio-tenurial Polarisation and Sub-markets." Urban Studies **37**(7): 1079-1092.

Borst, R. and W. McCluskey (2008). "Using Geographically Weighted Regression to Detect Housing Submarkets: Modelling Large-Scale Spatial Variations in Value." Journal of Property Tax Assessment & Administration **5**(1).

Bourassa, S., E. Cantoni, et al. (2007). "Spatial Dependence, Housing Submarkets, and House Price Prediction." Journal of Real Estate Finance and Economics **35**: 143-160.

Chhetri, P., R. Stimson, et al. (2006). "Modelling the Factors of Neighbourhood Attractiveness Reflected in Residential Location Decision Choices." Chiikigaku Kenkyu (Studies in Regional Science) **36**(2): 393-417.

Clapp, J. (2003), "A Semiparametric Method for Valuing Residential Locations: Application to Automated Valuation". Journal of Real Estate Finance and Economics 27:303-320.

Eckert, J., Ed. (1990). Property Appraisal and Assessment Administration, International Association of Assessing Officers.

Fotheringham, A., C. Brunsdon, et al. (2002). Geographically Weighted Regression the analysis of spatially varying relationships, Wiley.

Gallimore, P., M. Fletcher, et al. (1996). "Modelling the Influence of Location on Value." Journal of Property Valuation and Investment **14**(1): 6-19.

Galster, G. (1996). "William Grigsby and the Analysis of Housing Sub-markets and Filtering." Urban Studies **33**(10): 1797-1805.

Goodman, A. and T. Thibodeau (2003). "Housing market segmentation and hedonic prediction accuracy." Journal of Housing Economics **12**: 181-201.

Kauko, T. and M. d'Amato, Eds. (2008). Mass Appraisal Methods An international perspective for property valuers. Real Estate Issues. Oxford, Wiley-Blackwell.

Ming, Y. and K. Siu (2003). "Refining the Effects of Location on Computer Assisted Rating Valutaion." Pacific Rim Property Research Journal **9**(3): 224-247.

Moore, W. and J. Myers (2010). "Using Geographic-attribute Weighted Regression for CAMA Modeling." Journal of Property Tax Assessment & Administration **7**(3): 5-28.

O'Connor and Eichenbaum (1988). "Location Value Response Surfaces: The Geometry of Advanced Mass Appraisal." Property Tax Journal **7**(3): 277-296.

Rossini, P. and P. Kershaw (2008). Automated Valuation Model Accuracy: Some Empirical Testing. 14th Pacific Rim Real Estate Conference. Kuala Lumpur.

Tobler, W. (1970). "A Computer Movie Simulating Urban Growth in the Detroit Region." Economic Geography **46**(Supplement): 234-240.

Ward, R. (2006). "Developing Location Effects Using Cluster Analysis with Response Surface Analysis." Journal of Property Tax Assessment & Administration **3**(2): 5-17.

Ward, R. D., J. R. Weaver, et al. (1999). "Improving CAMA Models Using Geographic Information Systems/Response Surface Analysis Location Factors." Assessment Journal **6**(1): 30-39.

Appendix

The mean absolute percentage error (MAPE)

$$MAPE = \frac{\Sigma\left|(S_i - A_i)/S_i\right|}{n}$$

The Root-mean-squared error (RMSE)

$$RMSE = \sqrt{\frac{\Sigma(S_i - A_i)^2}{n}}$$

Coefficient of variation (COV)

$$COV = \frac{100}{\overline{AS}}\sqrt{\left(\frac{\sum_{i=1}^{n}\left(A_i/S_i - \overline{AS}\right)^2}{n-1}\right)}$$

gives a measure of dispersion about the mean in percentage terms. The higher the COV the more disperse the A/S ratios.

Coefficient of Dispersion (COD)

$$COD = \frac{100}{\widetilde{AS}}\left(\frac{\sum_{i=1}^{n}\left|A_i/S_i - \widetilde{AS}\right|}{n-1}\right)$$

gives a measure of the average absolute deviation from the median expressed as a percentage. The higher the COD the more disperse the A/S ratios.

Price related differential

$$PRD = \frac{\overline{AS}}{\sum_{i=1}^{n} A_i \Big/ \sum_{i=1}^{n} S_i}$$

A PRD value larger than 1 is referred to as *regressive* assessment indicating that the high valued properties are being under assessed (weighted mean is less than the mean) with respect to low valued properties, and conversely, less than 1, referred to as *progressive* assessment, indicating the high valued properties are being over assessed (weighted mean above the mean).

Forecast standard deviation

The Forecast Standard Deviation (FSD) is defined as the standard deviation of percentage forecast errors.

$$FSD = \sqrt{\left(\frac{\sum_{i=1}^{n} \left(\left(\frac{S_i - A_i}{A_i} \right) - \overline{\left(\frac{S_i - A_i}{A_i} \right)} \right)^2}{n-1} \right)}$$

where

$\widetilde{AS}$	=	median assessment
$\overline{AS}$	=	mean assessment
n	=	number of ratios
A_i	=	assessment for property i
S_i	=	Sale price for property i

This represents a one standard deviation (68%) probability that the predicted value falls within the percentage range given by the FSD. The lower the FSD, the smaller the error associated with the prediction. It is useful to consider here as it is quoted in relation to some commercially produced AVMs.

Email contact: tony.lockwood@unisa.edu.au