Location Value Response Surface Model as Automated Valuation Methodology a Case in Bari

Maurizio d'Amato

Abstract The paper is focused on the third application in the Italian context of Location Value Response Surface Modelling for Automated Valuation Modelling. LVRS (Connor in Locational Valuation Derived Directly from the Real Estate Market with the Assistance of Response Surface Techniques 1982) modelling is a procedure normally applied for automated valuation method purposes. In this context has been tested for a group of property prices exploring its use as automatic valuation methodology. The results showed that this method may have a potential role in those automated valuation model dealing with spatial autocorrelation.

Keywords Spatial models · Location value response surface · Housing prices

1 Introduction

Location value response surface modelling is an automated valuation method to appraise the estimate value of a property. The application of this method provides an integration between the tradition Multiple Regression Analysis with a location adjustment factor defined in this work as LAF. These factors can be used to include the distance as a measure of proximity effect of a specific place increasing or decreasing the property value in a specific urban context. The article shows a successful application of this kind of method to the Italian urban context. The article is organized as follows: the next paragraph will offer a brief literature review of Location Value Response Surface models, the second paragraph will show a second application of this model to an italian sample in Bari an italian south eastern city. Final remark and future direction of research will be offered at the end.

M. d'Amato (✉)
DICATECh Technical University,
Politecnico di Bari, Bari, Italy
e-mail: madamato@fastwebnet.it

M. d'Amato and T. Kauko (eds.), *Advances in Automated Valuation Modeling*,
Studies in Systems, Decision and Control 86, DOI 10.1007/978-3-319-49746-4_10

2 Location Value Response Surface Models: A Literature Review

Among different AVM modeling, LVRS model has been recently applied to Italian real estate market (d'Amato 2010). This works represents the third application of this automated valuation model to Italian real estate market. Location Value Response Surface (LVRS) Models (O'Connor 1982) were introduced to appraise single family houses in Lucas County (USA) without referring to fixed neighbourhoods or composite submarkets analysis. The method has been applied in the U.S. (Eichenbaum 1989, 1995; Ward et al. 1999), in England (Gallimore et al. 1996), Northern Ireland (McCluskey et al. 2000) and in Italy (d'Amato 2010, 2011). There are three approaches to LVRS. The first one (McCluskey et al. 2000) calculate a location adjustment factor referred to the spatial distribution of the selling prices. In this case a contour plot based on the ratio between prices and square meters will be originated overlying the area. The map will show the area and the point with higher and lower value of property values which are also called value influence centres (VICs). The distance from each VIC indicated in the contour map is calculated for each property of the sample. The impact of each VIC on any property is determined using different possible measures of the distance from the property to the VIC (Eckert 1990; Eckert et al. 1993). In fact these models are strongly dependent on the VIC positions and the adopted distance measure and transformation. The local adjustment factor will vary from −1 to +1 measuring the impact of location in the final regression model. Location adjustment factor does not indicate the value of a certain location, but only the relative location values for the property analysed. A further approach to LVRS consists in measuring the variance between actual prices and predicted prices using a MRA location blind model. Considering the error ratio related to under valuation or over valuation together with the coordinates of each observation an error map will be generated. The coefficient is included in the original location insensitive regression model. The third approach starts from an interpolation grid modelled to reflect the influence on each property of the location ratio factors within its proximity. In this work it is proposed an application to residential flats in the Italian real estate market. Spatial interpolation needs that the surface of the z variable (selling price or error term) would be continuous and the data value at any location can be estimated if sufficient information about the surface is given using the sample. The variable (selling price or error term) must be spatially correlated and the value at any specific location should be related to the values of surrounding locations.

3 LVRS Modelling

The sample is composed by 114 observations residential properties in an urban area called Carrassi near the downtown of Bari a city located in the south east of Italy. The sample of properties has been partially used for a previous application. The data are referred to residential flats in condominium in a temporal range between 20/05/1992 and 01/04/1997. Statistics on the selected observations are indicated in Table 1.

The dependent variable is the PRICE while DATE; SQM; SQM_BAL and PARK are the selected independent variables whose explanation is indicated in the Table 2.

Hedonic modelling have been used as a method to appraise the market value of a house for several decades (Palmquist 1980; Rosen 1974; Ciuna et al. 2014a, b). Selecting an appropriate functional form has been a recurring problem in the literature (Halvorsen and Pollakowski 1981). Five outliers has been detected. The problem is caused by the absence of theoretical justification for the appropriate functional relationship between housing price and its attributes (d'Amato and Siniak 2008; d'Amato 2008; d'Amato and Kauko 2012; d'Amato 2015). The functional form which better fit of the data in term of mean absolute percentage error and Box Cox test (for box cox test equal to 1), is the linear multiple regression model. The model location blind (or location insensitive) is indicated in the following Formula (1).

$$\begin{aligned} PRICE &= 101397.87 + 1391.762\,SQM + 482.614\,BAL \\ &\quad + 38654.23\,PARK - 1132.63\,DATE \end{aligned} \tag{1}$$

The output of this regression model is indicated in the Table 3.

The variables are not correlated and there is not multicollinearity as the VIF (Variance Inflation Factor) index is always less than the threshold. In the model are

Table 1 Descriptive statistics of the sample

Minimum	€50,850.00
Maximum	€293,000.00
Mean	€150,698.14
Standard deviation	€60,066.02

Table 2 Description of dependent and independent variables

Variables	Explication	Type of variable	Measure
DATE	Date of sales	Cardinal	Month
SQM	Square meters of flat	Cardinal	SQM
SQM_BAL	Square meters of balcony	Cardinal	SQM
PARK	Presence of parking	Dummy	Dichotomic
PRICE	Price of the property	Cardinal	Euro

Table 3 Linear MRA model n. 1 location blind

Variable	Acronym	Coefficient	T-stat	F-ratio	Adj. R2	MAPE	N. Obs
Location	LOC	101397	5.343	91.524	78.32	16.71	109
Square meters	SQM	1391.767	12.149				
Balcony	BAL	482.614	1.022				
Date	DAT	−1132.63	−6.9452				
Parking	PARK	38654.23	4.297				

included observation belonging to three fixed neighbourhood (Carrassi, Poggiofranco, S. Pasquale). All of these "neighbourhood" belongs to the same residential area called Picone. The new linear multiple regression model was applied dividing the observation between two fixed neighborhoods as indicated in the Formula (2) avoiding the dummy variable trap (Green 2003). The new model will be

$$\begin{aligned} \text{PRZ} \; &= 84.570 + 1,410.22 \cdot \text{SQM} \; + 511.252 \cdot \text{BALCONY} \\ &- \; \text{DATE} \cdot 982.63 + \; \text{PARK} \cdot 42,243.32 + \\ &- \text{NG1},16849.324 + \; \text{NG2} \cdot 43725.623 \end{aligned} \quad (2)$$

These neighborhoods are coincident with two different zones in the area, the third one is included in the constant term. The output of regression can be read in the Table 4.

The variables are not correlated and there is not multicollinearity as the VIF (Variance Inflation Factor) index is always less than the threshold. The role of spatially modeled variable in hedonic pricing has been highlighted in several different works (Des Rosiers et al. 2003, 2005). In order to analyse the opportunity of the application of spatial related model, it is necessary to consider the degree of spatial autocorrelation of a fundamental variable as the ratio between price and square meters. The most commonly used and robust indicator was proposed by the statistician Moran (1948, 1950) and it is normally indicated as Moran I test indicated in the Formula (3).

Table 4 Linear MRA model n. 2 using 3 fixed neighbour groups

Variable	Acronym	Coefficient	T-stat	F-ratio	Adj. R2	MAPE	N. Obs
Location	LOC	84570	5.0012	77.422	79.57	16.19	109
Square meters	SQM	1410.22	12.4272				
Balcony	BAL	511.252	−7.0752				
Date	DAT	−982.63	1.284				
Parking	PARK	42243.32	4.924				
Neighbour group 1	NG_1	−16849.324	−1.254				
Neighbour group 2	NG_2	43725.623	2.822				

$$I = \frac{N \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})^2} \quad (3)$$

where x is the variable (, and *wij* represents the set of neighbours j for observation i). In this case the inverse squared distance among the observations has been considered according to previous works (Des Rosiers et al. 1999; Shiller 1993; Wyatt 1997; Kauko and d'Amato 2008a; Kauko and d'Amato 2008b; Kauko and d'Amato 2008c; Kauko and d'Amato 2011; Kaklauskas et al. 2012). The Moran's I ranges from −1 to +1 and each observation is only compared with its relevant neighbourhood as a consequence positive Moran's I means positive autocorrelation. In this case high values for x value should be located near other high values while lower market basket values should be located near other lower market basket values. In this work the Moran I test showed a high positive autocorrelation assuming a value of 0.7274. There is a theoretical premise to integrate traditional multiple regression analysis with spatial analysis. A location adjustment factor will be added to the model location blind indicated in the Formula (1). This location

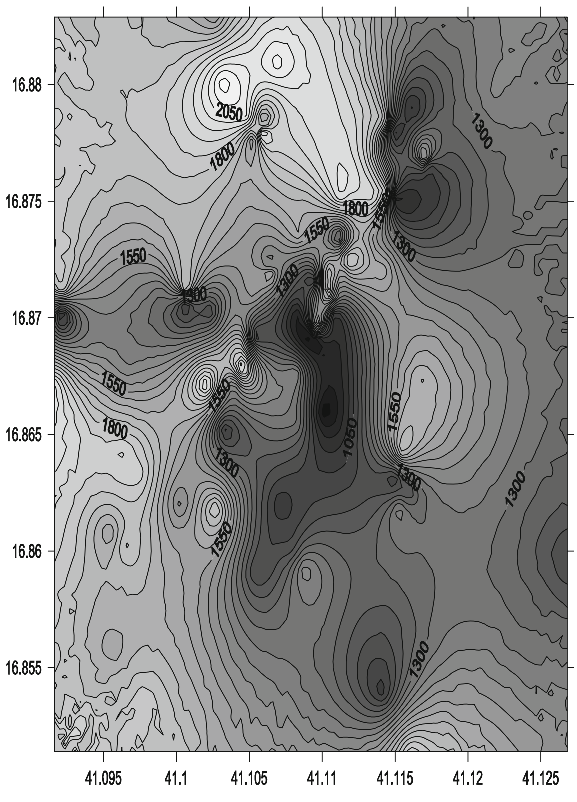

Fig. 1 Contour plot of market basket value carrassi san pasquale area using a linear variogram

adjustment factor will be based on a contour map developed on the ratio price per square meters. The Fig. 1 indicate the contour map originated by a linear variogram in the context calculated using SURFER 8.

In the Fig. 1 there are 12 value influence centers located having the coordinates indicated in the Table 5.

In the Table 5 each VIC is individuated with the decimal coordinates (longitude and latitude). The final column of Table 6 indicates the nature of the value influence center, H means higher values while the L indicates lower values. It is worth to notice that the value influence centers n. 2; 4; 7; 10; 11; 12 are closer to an important urban park in the area called *Parco Largo 2 Giugno* therefore higher values can be observed. This confirms the cause of spatial autocorrelation indicated in previous scientific work (Gillen and Thibodeau et al. 2001). A location adjustment factor was calculated regressing the market basket value as dependent variable on the location and the distance between each point and the value influence center. Including VIC in the regression model require a method to transform physical distance such as half Gaussian or gravity models. They transform the physical measure in a proximity variable. In this case the transformation occurred using a half Gaussian function normally indicated in literature (O'Connor 1982). The model including the location adjustment factor is indicated in the Formula (4).

$$\begin{aligned} \mathrm{PRZ} = & -15{,}194.30 + 1{,}384.12 \cdot \mathrm{SQM} + 307.91 \cdot \mathrm{BALCONY} \\ & - \mathrm{DATE} \cdot 857.93 + \mathrm{PARK} \cdot 19{,}243.84 + \mathrm{LAF} \cdot 107{,}249.80 + \varepsilon \end{aligned} \tag{4}$$

The last term indicated as LAF is the location adjustment factors. The constant term of the Formula (4) is diminished compared to the other models, the adjusted R^2 of the model n. 3 is improved growing to 0.836 and the t-student test are significant for all the variable except for Balcony. There is also an improvement in term of mean absolute percentage error. Table 6 indicates the output of regression analysis.

Table 5 Value influence centers individuated in the contour map of Fig. 1

Value influence center	Latitude	Longitude	Nature VIC
VIC1	41.1100000	16.8693216	L
VIC2	41.1038779	16.8800000	H
VIC3	41.1153174	16.8543407	L
VIC4	41.0999831	16.8752060	H
VIC5	41.1177992	16.8782044	L
VIC6	41.1156999	16.8786358	L
VIC7	41.1111999	16.8771798	H
VIC8	41.1029999	16.8716999	L
VIC9	41.1087545	16.8713759	L
VIC10	41.1132996	16.8578977	H
VIC11	41.1191000	16.8526769	H
VIC12	41.0936229	16.8810161	H

Table 6 Location value response surface model

Variable	Acronym	Coefficient	T-stat	F-ratio	Adj. R2	MAPE	N. Obs
Location	LOC	−15,194.30	1.9342	116.29	83.6	13.21	109
Square meters	SQM	1384.12	12.3729				
Balcony	BAL	307.91	1.5329				
Date	DAT	−857.93	−4.7983				
Parking	PARK	19,243.84	4.924				
Location adjust factor	LAF	107,249.80	7.0002				

Table 7 Comparing the three model's main findings

	Model 1	Model 2	Model 3
	Location blind	*Model with neighbour groups*	*Location value response surface*
MAPE	16.71	16.19	13.21
Adj R2	78.32	79.57	83.6

The variables are not correlated and there is not multicollinearity as the VIF (Variance Inflation Factor) index is always less than the threshold. Finally the Table 7 compares the mean absolute percentage errors and the adjusted R^2 of the four models presented.

The model 1 is the location blind indicated in the Formula (1), the model number 2 is the fixed neighbor model indicated in the Formula (2) while the third is an application of LVRS model presented in the previous paragraph.

4 Conclusions and Future Directions of Research

Empirical findings demonstrates that LVRS may be an interesting option for automated valuation methodologies, too. This application in Italian context and residential segment showed an evident increase of the quality of the model. Empirical studies are required to analyse the difference between property valuation carried out by valuers, mass appraisal and price. This would be helpful to understand the relation between the price and the valuation (Renigier-Biłozor 2014a, b) and how spatial context affect both. In particular further works may be required testing also the error correction model in the Italian context. The application showed an evident vagueness of boundaries among different VIC, a typical situation determined by the nature of Italian context with several different urban areas. As one can see the number of VIC used in italian application is much higher than in other experience because the complex nature of Italian urban context.

References

Box, G., & Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society, 26*, 211–252.

Ciuna, M., Salvo, F., & De Ruggiero, M. (2014a). Property prices index numbers and derived indices. *Property management, 32*(2), 139–153. DOI(Permanent URL):10.1108/PM-03-2013-0021.

Ciuna, M., Salvo, F., & Simonotti, M. (2014b). Multilevel methodology approach for the construction of real estate monthly index numbers. *Journal of Real Estate Literature, 22*(2), 281–302.

d'Amato, M. (2015). Income approach and property market cycle. *International Journal of Strategic Property Management, 29*(3), 207–219.

d'Amato, M. & Kauko, T. (2012). Sustainability and risk premium estimation in property valuation and assessment of worth. *Building Research and Information, 40*(2), 174–185 (March–April 2012).

d'Amato, M. (2008). Rough set theory as property valuation methodology: The whole story (Chap. 11). In T. Kauko & M. d'Amato (Eds.), *Mass appraisal an international perspective for property valuers* (pp. 220–258). Wiley Blackwell.

d'Amato, M. (2010). A location value response surface model for mass appraising. An "iterative" location adjustment factor in Bari, Italy. *International Journal of Strategic Property Management, 14*, 231–244.

d'Amato, M. & Kauko, T. (2008). Property market classification and mass appraisal methodology (Chap. 13). In T. Kauko & M. d'Amato (Eds.), *Mass appraisal an international perspective for property valuers* (pp. 280–303). Wiley Blackwell.

d'Amato, M. & Siniak, N. (2008). Using fuzzy numbers in mass appraisal: The case of belarusian property market (Chap. 5). In T. Kauko & M. d'Amato (Eds.), *Mass appraisal an international perspective for property valuers* (pp. 91–107). Wiley Blackwell.

Des, Rosiers F., Theriault, M., & Menetrier, L. (2005). Spatial versus non-spatial determinants of shopping center rents: Modeling location and neighborhood-related factors. *Journal of Real Estate Research, 27*(3), 293–319.

Des Rosiers, F., Theriault, M., Villeneuve, P., & Kestens, Y. (1999). *House price and spatial dependence: Towards an integrated procedure to model neighborhood dynamics*. Paper presented at 1999 AREUEA annual meeting.

Des, Rosiers F., Theriault, M., Villeneuve, P., & Kestens, Y. (2003). Modelling interactions of location with specific value of housing attributes. *Property Management, 21*(1), 25–62.

Eckert, J., Ed. (1990). Property appraisal and assessment administration. *International Association of Assessing Officers*.

Eckert, J., O'Connor, P., & Chamberlain, C. (1993). Computer-assisted real estate appraisal: a california savings and loan case study. *The Appraisal Journal*, *LXI*(4), 524–532.

Eichenbaum, J. (1989). Incorporating location into computer-assisted valuation. *Property Tax Journal, 8*(2), 151–169.

Eichenbaum, J. (1995). The location variable in world class cities: Lessons from CAMA valuation in New York City. *Journal of Property Tax Assessment & Administration, 1*(3), 46–60.

Gallimore, P., Fletcher, M., & Carter, M. (1996). Modeling the influence of location on value. *Journal of Property Valuation and Investment, 14*(1), 6–19.

Gillen, K., Thibodeau, T., & Wachter, S. (2001). Anisotropic autocorrelation in house prices. *Journal of Real Estate Finance and Economics, 23*(1), 5–30.

Green, W. H. (2003). *Econometric analysis*. New Jersey: Prentice Hall.

Halvorsen, R., & Pollakowski, H. (1981). Choice of functional form for hedonic price equations. *Journal of Urban Economics, 10*, 37–49.

Kaklauskas, A., Daniūnas, A., Dilanthi, A., Vilius, U., Lill, I., Gudauskas, R., et al. (2012). Life cycle process model of a market-oriented and student centered higher education. *International Journal of Strategic Property Management, 16* (4), 414–430.

Kauko, T., & d'Amato, M. (2011). *Neighbourhood effect, international encyclopedia of housing and home*. Edited by Elsevier Publisher.

Kauko, T., & d'Amato, M. (2008a). *Mass appraising. An international perspective for property valuers.* London: Wiley Blackwell.

Kauko, T., & d'Amato, M. (2008b). Introduction: Suitability issues in mass appraisal methodology. In T. Kauko & M. d'Amato (Eds.), *Mass appraisal an international perspective for property valuers* (pp. 1–24). Wiley Blackwell.

Kauko, T., & d'Amato, M. (2008c). Preface. In T. Kauko & M. d'Amato. (2008). (Eds.), *Mass appraisal an international perspective for property valuers* (p. 1). Wiley Blackwell.

McCluskey, W. J., Deddis, W. G., Lamont, I. G., & Borst, R. A. (2000). The application of surface generated interpolation 3a models for the prediction of residential property values. *Journal of Property Investment and Finance, 18*(2), pp. 162–176. MCB University Press.

Moran, P. A. P. (1948). The interpretation of statistical maps. *Journal of the Royal Statistical Society B, 10*, 243–251.

Moran, P. A. P. (1950). Notes on continuous stochastic phenomena. *Biometrika, 37*, 17–23.

O'Connor, P. (1982). *Locational valuation derived directly from the real estate market with the assistance of response surface techniques*, Lincoln Institute of Land Policy.

Palmquist, R. (1980). Alternative techniques for developing real estate price indexes. *Review of Economics and Statistics, 62*, 442–448.

Rasmussen, D., & Zuehlke, T. (1990). On the choice of functional form for hedonic price functions. *Applied Economics, 22*, 431–438.

Renigier-Biłozor, M., Wiśniewski, R., Biłozor, A., & Kaklauskas, A. (2014a). Rating methodology for real estate markets—Poland case study. *Pub. International Journal of Strategic Property Management. 18*(2), 198–212. ISNN. 1648-715X.

Renigier-Biłozor, M., Dawidowicz, A., & Radzewicz, A. (2014b). An algorithm for the purposes of determining the real estate markets efficiency in Land Administration System. *Pub. Survey Review, 46*(336), 189–204.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*, 34–55.

Shiller, R. (1993). Measuring asset value for cash settlement in derivative markets: Hedonic repeated measures indices and perpetual futures. *Journal of Finance, 48*, 911–931.

Ward, R. D., Weaver, J. R., & German, J. C. (1999). Improving models using geographic information systems/response surface analysis location factors. *Assessment Journal, 6*(1), 30–38 (Jan/Feb 1999).

Wyatt, P. J. (1997). The development of a GIS-based property information system for real estate valuation. *International Journal of Geographical Information Science, 11*(5), 435–450.