ENHR Prague 2009,WS 1

Hybrid Models and AVM. An Application of LVRS models in Bari

Abstract

The theoretical part of the paper explore the relationship between automated valuation models and property valuation. In particular it is highlighted how this integration may improve the efficiency and the independence of property valuation in the credit crunch. The empirical part of the paper is focused on an application of an Automated Valuation Model based on a group of residential property transactions in an area of Bari. Property transactions have been modelled using location value response surface (O'Connor, 1982) a technique which has been applied in several contexts: (Eichenbaum, 1989 & 1995; Ward et al., 1999), in England (Gallimore et al., 1996), and Northern Ireland (McCluskey et at ., 2000)

Maurizio d'Amato
1st Faculty of Engineering - Technical University Politecnico di Bari
Via Calefati, 272 - 70122 Bari - Italy
Email: madamato@interfree.it
Telephone: +39 (0)80 9645267
Fax: +39 (0)80 0999777
Research Web: www.noaves.com
Didactic web: http://mdamato.altervista.org

Introduction

The paper has the finality to explore the relationship between real estate valuation and automated valuation processes. There is a progressive integration between automated valuation methods and property valuations through the so called hybrid models. The work explores the relationship among these methodologies from empirical and theoretical points of view. This integration may vary according to the nature of process. It may be Mass Appraising or in the field of the AVMs. While in the case of the Mass Appraisal property valuation have the finality to determine value for fiscal finality, in the AVMs the finality is related to the process of mortgage lending. The works from a theorical point of view highlights the relationship between property valuation and automated valuation methodology. This relation is analysed in the first paragraph. In the empirical part of the paper there is the second application of location value response surface modelling to the italian context in Bari. In particular the second paragraph gives a brief outline of this automated valuation modelling. The third paragraph shows an application of LVRS to an italian context in Bari. Final remarks will be offered at the end.

1. AVM, Hybrid models and the Institutional Context

Several documents (White Paper and Green Paper) EU address the central role of property valuation in mortgage lending and in the harmonization process of EU real estate and secondary mortgage markets. The white paper stated that there are different institutional contexts all over the EU. The institutional context is not only relevant in mass appraising modelling (Kauko and d'Amato, 2008) but also in Automated Valuation Methodologies. There is an urgent need to explore the relationship between automated valuation methodology and their relationship with the institutional context. Since the introduction of automated valuation methodology in the United States in 1995 (Cassens,2006) there has been a growing use of automated valuation modelling in the mortgage lending industry and in the secondary mortgage market. In US there has been[1] a debate on journals addressing the superiority of AVM results compared to the professional valuation. In any case there are several possible forms of integration between automated valuation modelling and mass appraisal that has been classified in a previous work (Cassens,2006). These models can be defined as AVM Hybrid models. They explore the relation between property valuation activity and automated valuation results, integrating both. In general term it is possible to see the Hybrid models as an evolution addressed in the following figure number 1:

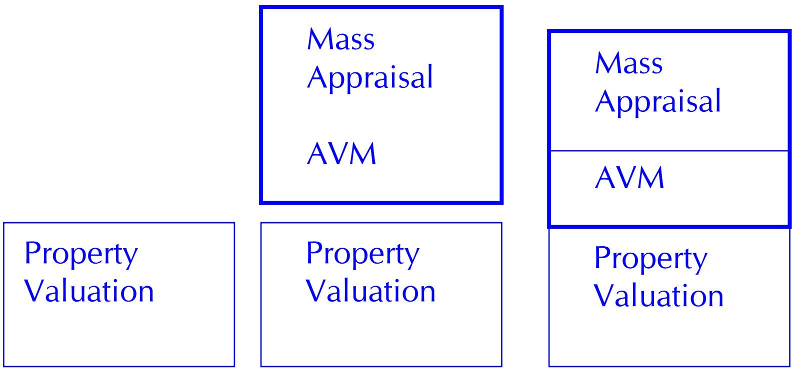

Fig.1 – Evolution starting from property valuation arriving at integration between mass appraisal-AVM models and property valuation

In the first phase there is a development of property valuation models. In the second phase the growth of computer computation capability caused a growing interest in Computer Assisted Mass Appraisal together with the born of organized ad reliable real estate property data bank. In the third phase there is a strict cooperation between Mass Appraisal and Property Valuation. The relationship is not always defined. It depends on the institutional context. For example US property market seems to be entered in the third phase while Italian market can be allocated in the second one. There is a growing interest in Italy on property databank after the introduction of the third version of Italian Property

[1] Susan E. Woodwards (2008), Rescued by Fannie Mae, October 14 on The Washington Post wrote that automated valuation methodologies may replace the valuation carried out by professional valuers. The president of Appraisal Institute R. Wayne Pugh answered Woodwards defending the importance of property valuation carried out by professional valuers.

Valuation Standard (Tecnoborsa,2005). Information is available and reliable in many countries but there are also countries which is still experiencing the born of property data bank. It is possible to say that also AVM is strictly related to the institutional context. In order to explore the power of AVM in the Italian context three AVM models have been applied to an urban context of Bari in south east of Italy.

2. Location Value Response Surface Models: a brief outline

There is a growing interest in Italy for Mass Appraisal methodologies (Simonotti,2005;Simonotti,2006), although the application of models and property databank are still under development. This work represents one of the first application of Location Value Response Surface modelling to Italian context. The first work introducing Location Value Response Surface (LVRS) Models (O'Connor, 1982) has used an innovative method to appraise single family houses in Lucas County (USA) which did not refer to fixed neighbourhoods or composite submarkets analysis. The application of this method is based on spatial interpolation of property prices or spatial interpolation of error term . At the moment has been applied in the U.S. (Eichenbaum, 1989; Eichenbaum, 1995; Ward et al., 1999), in England (Gallimore et al., 1996), Northern Ireland (McCluskey et al ., 2000), Italy (d'Amato,2009). In this models the location is analysed using Geographical Information Systems (GIS). There are several possible approaches to LVRS among them it is possible to classify the following three approaches. The first one (McCluskey W.J., et al. ;2000) calculate a location adjustment factor referred to the spatial distribution of the selling prices.The process starts dividing the actual price by the gross floor area of the dwelling in order to obtain a price per square meter. Therefore a contour plot overlying the area map is originated. This map portrays the peaks and troughs of property values which are also called value influence centres (VICs). The Value Influence Centers are point(s), line(s) or area(s) in a map where it is possible to observe a relative maximum (positive) or a minimum (negative) location values or rent (errors) . A Value Influence Center is an are or points able to influence the values (or rents) of near properties. Therefore the physical distance from each VIC indicated in the contour map is calculated for each property of the sample. The selling price per square meter is regressed both on the coordinates and on the distance of each property to each VIC. The predicted price is then divided by the average estimated price. As a consequence will be obtained a local adjustment factor having a mean of 1. In particular better locations will have a factor greater than 1, while poorer locations will have a factor less than 1. This local adjustment factor will vary from -1 and 1 measuring the impact of location in the final regression model. A second approach to LVRS consists in measuring the variance between actual prices and predicted prices using a MRA location blind model. There will be higher value of forecasting error in some areas and lower value of forecasting error in other areas generating a contour plot of errors instead of value. Using the error ratio related to under valuation or over valuation and the coordinates of each observation it will be possible determine a coefficient similar to the previous model . 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). Response surface model is strongly depending on the VIC positions and the adopted distance measure. In the third approach an interpolation grid is modelled to reflect the influence on each property of the location ratio factors within its proximity. These methods has not been applied to residential flats. This work represents the second application to Italian real estate market. A prerequisite of LVRS application is having sufficient amount of data in each zone of the area

considered in order to produce the spatial interpolation. There are not a minimum number of observations but real estate market, especially in the Italian context presents a scarcity of data.
Location Adjustment factor does not indicate the value of a certain location, but only the relative location values for the property analysed. Spatial interpolation requires the surface of the z variable (selling price or error term) to be continuous and the data value at any location can be estimated if sufficient information about the surface is given using the sample. The z variable (selling price or error term) must be spatially dependent therefore the value at any specific location is related to the values of surrounding locations.

3. Applying LVRS method in Bari,Italy. Data and Methodology

The sample is composed by 105 observations located in an urban area called Carrassi near the dowtown of Bari. The data are referred to residential aparment in condominium. Statistics on the selected observations are indicated in table n.1

Minimum	€	50.850,00
Maximum	€	320.000,00
Mean	€	150.684,67
Standard Deviation	€	61.873,99

Table 1 Statistic of the sample

The observations included in this cross-section analysis are related to a temporal range included in an interval between 20/05/1991 and 01/04/2004. 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 below:

DATE	Date of Sales	month
SQM	Square Meters of Flat	sqm
SQM_BAL	Square Meters of Balcony	sqm

PARK	Presence of Parking	dichotomic
PRICE	Price of the Property	euro

Table 2 – Description of dependent and independent variables

Among several forms was selected a linear multiple regression model

The linear MRA model location blind (or insensitive) is the following

$$PRICE = 100357{,}91 + 1371{,}19 SQM + 482 SQM\_BAL + 36584{,}14 PARK - 1122{,}07 DATE \quad (1)$$

The output of regression is reported in the paragraph 1 of Appendix.The regression was carried out using SPSS vers. 16.0.

Therefore according to spatial distribution of observations, two neighbours groups were detected. The clusters were selected observing both the geographic coordinates and the Market Basket Value. A new multiple regression models location with two fixed neighborhoods is indicated in the formula 2 below:

$$PRZ = 87.460 + 1.432{,}833 \bullet SQM + 531{,}643 \bullet BALCONY - DATE \bullet 1.057{,}37 + PARK \bullet 40.242{,}731 + \quad (2)$$
$$- NG1 \bullet 17.815{,}410 + NG2 \bullet 41.710{,}42 + \varepsilon$$

The constant term is diminshed comparing to formula n.1 the t-test are verified for all the independen variables except for BALCONY and NG1. The R2 improved. The model indicated in the formula 1 has a R^2 of 0,78 while the model indicated in the formula 2 shows a R^2 of 0,804. The output of regression can be read in the paragraph 2 of Appendix. The third model is a Location Value Surface Model applied to the same context. The model location blind indicated in the formula 1 will be integrated with a location adjustment factor. This location adjustment factor will be based on a contour map devloped on a market basket value. The observations have been spatially located using latitude and longitude. Spatial autocorrelation was indivituated using Moran I test. Market basket value represents the ratio between the property price and the square meters. The fig.2 below indicate the contour map originated by a linear variogram in the context. It is possible to observe a value influence center located having the coordinates LAT 41,1030000 LONG 16,8800000. This value is originated by the presence of a urban park. This confirms one the cause of spatial autocorrelation indicated (Gillen, Thibodeau et. al)

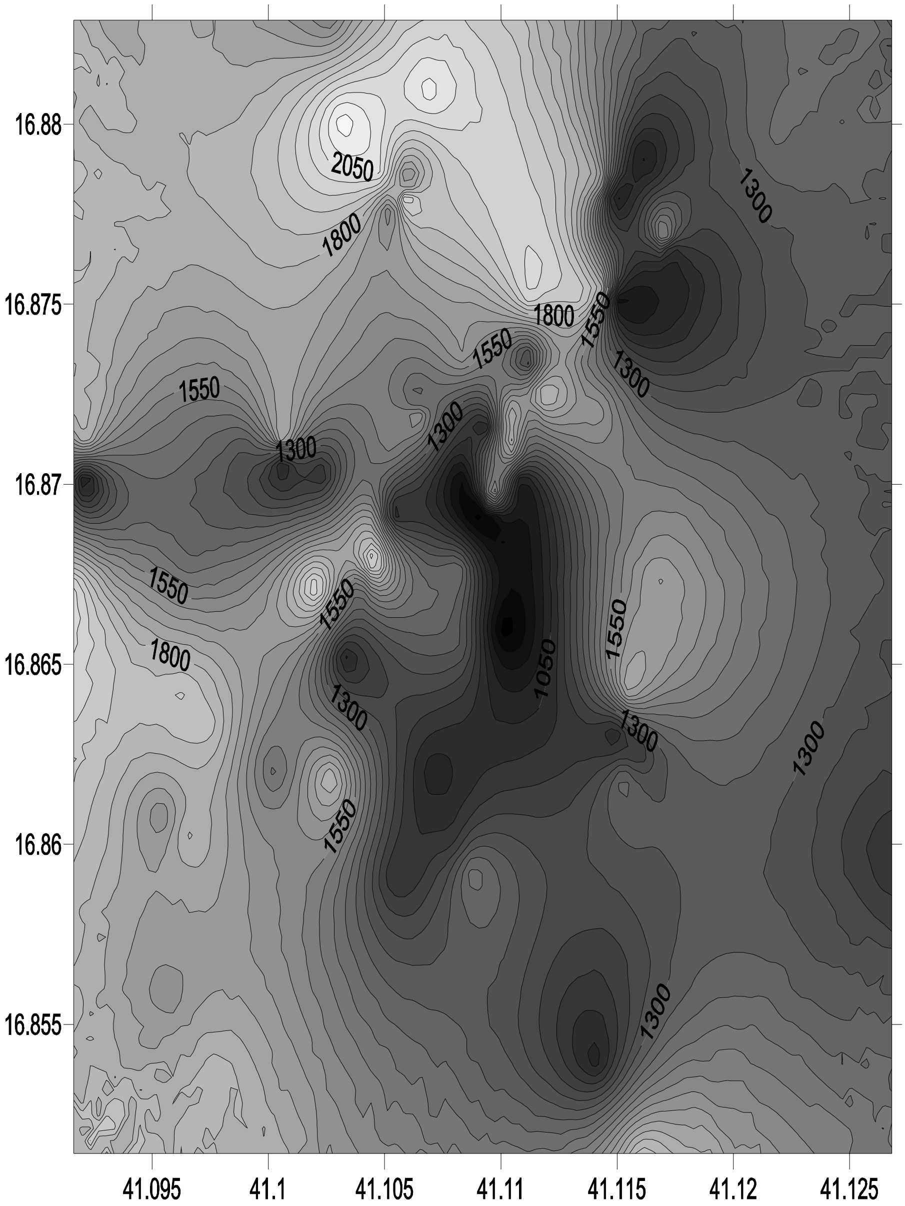

Fig.2 Contour Plot of Area Carrassi San Pasquale Using a Linear variogram

A three dimensions figure indicated with the number 2 highlights the dynamic of property values in the selected areas.

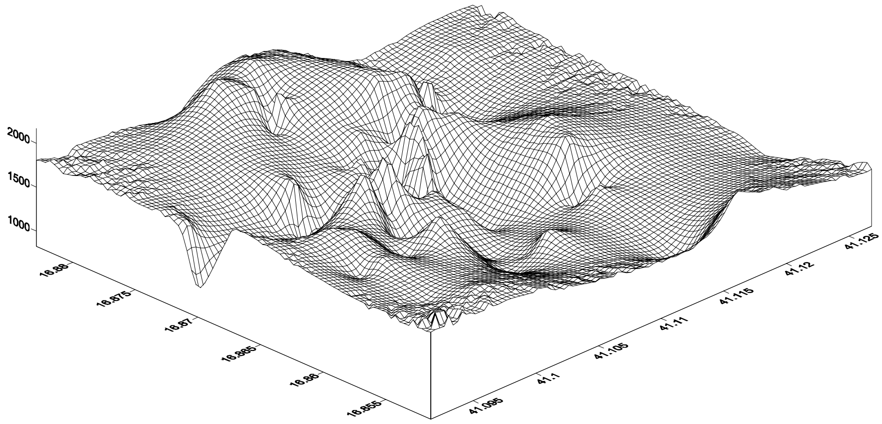

Fig.2 Three dimensions representation of market basket value

A location adjustment factor was calculated regressing the market basket value as dependent variable on the coordinates of observations and the distance between the value influence center and each observations. The distance was calculated computing the physical distance between the coordinates (Latitude and Longitude) on earth between two points individuate. The result of the regression is therefore divided by the average estimated price having a location factor greater than one in presence of significant location less than one otherwise. The model including the location adjustment factor is indicated in the formula 3 below:

$$PRZ = -20.918{,}024 + 1.484 \bullet SQM + 284{,}203 \bullet BALCONY - DATE \bullet 1.119{,}72 + \\ + PARK \bullet 25.948{,}91 + LAF \bullet 114.477{,}46 + \varepsilon \qquad (3)$$

The constant term is diminished compared to formula n.1 and fomula n.2. The R^2 is improved growing to 0,81 and the t-student test are signficant for all the variable except for Balcony. There is also an improvent in term of mean absolute percentage error. The table 3 below compare the mean aboslute percentage errors of the three models:

MAPE		
Model 1	Model 2	Model 3
17,5	17	15

Table 3- Comparing three models MAPE

The model 1 is the location blind indicated in the formula 1, the model number 2 is the fixed neighbour model indicated in the formula 2 while the third model is the application of LVRS model indicated in the formula n.3

Conclusions and Future Directions of Research

The application, which is the second application of this AVM modelling to italian context, demonstrates that LVRS may be an interesting option for Computer Assisted Mass Appraisal and automated valuation modelling in Italian context. The integration between AVM and professional valuation should be based on property data bank which are growing also in Italy like in other countries. More empirical studies may be required to analyse the difference between property valuation carried out by valuers and automated valuation results. The valuation variation between AVM and professional valuation may give interesting contributions to understand how the relationship (evolution) described in the figure 1.

Bibliography

Cassen Zilloux V. (2006), Automated valuation Models: Automation vs. Hybrid Blending Technology and Human Intervention, Paper presented at European Real Estate Society meeting, Weimar, Germany, June

d'Amato M. (2009) A Location Value Response Surface Model for Mass Appraising. An "Iterative" Location Adjustment Factor in Bari, Italy, Paper Submitted to Special Issue on Property Valuation of International Journal of Strategic Property Management

Eichenbaum, J. (1989), Incorporating Location into Computer-Assisted Valuation, Property Tax Journal Vol.8:2 p.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 Vol.1:3 May 1995 p.46-60

Gallimore, P., Fletcher, M. and Carter, M. (1996), Modeling the Influence of Location on Value, Journal of Property Valuation and Investment, Vol. 14 No. 1, pp. 6-19.

Gillen K, Thibodeau T. Wachter S. (2001) Anisotropic Autocorrelation in House Prices, Journal of Real Estate Finance and Economics, 23:1, pp.5-30

Green Paper on Mortgage Credit in EU http://ec.europa.eu/internal_market/finservices-retail/home-loans/integration_en.htm#greenpaper - last access 21.06.2009

Kauko T. d'Amato M. (2008), Mass Appraising. An International Perspective for Property Valuers, Wiley Blackwell, London

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, Vol.18 n.2 pp.162-176, MCB University Press

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.

Simonotti M. (2006), Metodi di Stima Immobiliare, Ed. Flaccovio, Palermo, Italy

Simonotti M. (2005), Manuale delle Stime Immobiliari, Geoval, Roma, Italy

TECNOBORSA (2005) , Codice delle Valutazioni Immobiliari. Italian Property Valuation Standard,ed. Tecnoborsa, Roma, Italy

Ward R.D., Weaver J.R., German J. C. (1999) Improving Models using Geographic Information Systems/Response Surface Analysis Location Factors. Assessment Journal Vol. 6 : 1 Jan/Feb 1999 p.30-38

White Paper on the Integration of EU Mortgage Credit Markets http://ec.europa.eu/internal_market/finservices-retail/home-loans/integration_en.htm#greenpaper - last access 21.06.2009

Appendix

1. MULTIPLE LINEAR REGRESSION MODEL 105 OBSERVATIONS IN CARRASSI AREA

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,884[a]	,782	,773	29481,49935

a. Predictors: (Constant), PARK, DATE, SQM, SQM_BAL

b. Dependent Variable: PRICE

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3,112E11	4	7,781E10	89,522	,000[a]
	Residual	8,692E10	100	8,692E8		
	Total	3,982E11	104			

a. Predictors: (Constant), PARK, DATE, SQM, SQM_BAL

b. Dependent Variable: PRICE

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.
1	(Constant)	100357,913	18785,571		5,342	,000
	SQM	1371,197	122,992	,601	11,149	,000
	SQM_BAL	482,613	471,585	,058	1,023	,309
	DATE	-1122,073	153,559	-,354	-7,307	,000
	PARK	36584,147	8338,881	,233	4,387	,000

a. Dependent Variable: PRICE

2. MULTIPLE LINEAR REGRESSION MODEL ON 105 OBSERVATIONS IN CARRASSI AREA USING TWO FIXED NEIGHBOURHOOD

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,897[a]	,804	,793	28183,600

a. Predictors: (Constant), NG2, NG1, PARK, DATE, SQM, SQM_BAL

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3,203E11	6	5,338E10	67,209	,000[a]
	Residual	7,784E10	98	7,943E8		
	Total	3,982E11	104			

a. Predictors: (Constant), NG2, NG1, PARK, DATE, SQM, SQM_BAL

b. Dependent Variable: PRICE

Coefficients[a]

Model		Unstandardized Coefficients B	Unstandardized Coefficients Std. Error	Standardized Coefficients Beta	t	Sig.	Collinearity Statistics Tolerance	Collinearity Statistics VIF
1	(Constant)	87460,838	18874,869		4,634	,000		
	SQM	1432,833	118,774	,628	12,064	,000	,737	1,357
	DATE	-1057,375	150,087	-,333	-7,045	,000	,892	1,121
	SQM_BAL	531,643	448,595	,064	1,185	,239	,684	1,463
	PARK	40242,731	8327,121	,241	4,833	,000	,804	1,244
	NG1	-17815,410	14471,882	-,055	-1,231	,221	,986	1,014
	NG2	41710,427	15076,974	,130	2,766	,007	,908	1,101

a. Dependent Variable: PRICE

3. MULTIPLE LINEAR REGRESSION MODEL ON 105 OBSERVATIONS IN CARRASSI AREA USING LOCATION ADJUSTMENT FACTOR

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	,900[a]	,810	,800	2,767798E4

a. Predictors: (Constant), LAF, Date, Balcony, Sq_meter, Park

b. Dependent Variable: Price

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3,223E11	5	6,446E10	84,147	,000[a]
	Residual	7,584E10	99	7,661E8		
	Total	3,982E11	104			

a. Predictors: (Constant), LAF, Date, Balcony, Sq_meter, Park

b. Dependent Variable: Price

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Collinearity Statistics	
		B	Std. Error	Beta			Tolerance	VIF
1	(Constant)	-20918,024	36447,335		-,574	,567		
	Sq_meter	1484,108	119,225	,650	12,448	,000	,705	1,418
	Balcony	284,203	445,800	,034	,638	,525	,668	1,498
	Date	-1119,721	144,166	-,353	-7,767	,000	,932	1,073
	Park	25948,910	8313,438	,165	3,121	,002	,685	1,461
	LAF	114477,466	30108,155	,184	3,802	,000	,824	1,214

a. Dependent Variable: Price