DEFINING A GEOGRAPHICALLY WEIGHTED REGRESSION MODEL OF URBAN EVOLUTION. APPLICATION TO THE CITY OF VOLOS, GREECE.

Milaka Kyratso[1], Photis Yiorgos[2]

[1]PhD Candidate, Department of Urban Planning and Regional Development, Laboratory for Spatial Analysis, GIS and Thematic Planning, University of Thessaly, Pedion Areos, Volos, Greece, tel: +302421074426, fax: +302421074382

[2] Assistant Professor, Department of Urban Planning and Regional Development, Laboratory for Spatial Analysis, GIS and Thematic Planning, University of Thessaly, Pedion Areos, Volos, Greece, tel: +302421074418, fax: +302421074382

Abstract

The main objective of this paper is the multivariate analysis of urban space and specifically with the use of data that refer to the level of city block. Part of the analysis has been the comparative assessment of multiple linear regression and geographically weighted regression (GWR) analysis as well as the application of the aforementioned methods in the study of the central district of the Volos metropolitan area.

The city of Volos is an urban conglomeration of approximately 110.000 inhabitants, located at the middle-east of Greece and is considered to be in the upper extreme in the cities' urban hierarchy in Greece. The results provide a response to a question raised by spatial scientists during the last decades: *is there a way that regression analysis can reveal spatial variations of results and with respect to scale fluctuation?*

The use of classical multiple regression analysis provides a single result – equation for the entire area. On the other hand, geographically weighted regression analysis stems from the fact that the above result is inadequate to reflect the different relational levels among selected variables characterizing the entire area. New estimations with the use of GWR declare the existence of various sub-areas – divisions of the initial territory – formulating a set of equations that reveal the spatial variations of variable relations.

The results of the application have well proved the dominance of the analysis in the local level towards the analysis in the global level, highlighting the existence of intense spatial differentiations of variables that "interpret" the rate of land values in the city. Moreover, the distinct spatial patterns that emerge throughout the entire area, establish an alternative approach urban spatial phenomena interpretation and a new explanatory basis for the clarification of obscure relations.

1. INTRODUCTION

The description of phenomena that affect and determine land values present significant interest in the urban space analysis. In the current article this analysis has a dual objective: a) the determination of variables that define substantially the evolutional process of land market and b) the investigation of a potential spatial inconstancy in the way the selected variables affect land values. The first part of the analysis has been subject of research during the last decades (e.g Alonso, Ricardo, Tiebout etc) whereas the second part constitutes a rather modern approach of the aforementioned phenomena.

In what concerns the urban level, regression analysis is more frequently used in the determination of models of multi-variant analysis. If this is the case, land values are defined by an entity of variables that are characteristics of size, location etc. The contribution of bibliography has been great for the acquisition of experience. (Azar et al, 1994)

Modern methods of multivariate analysis on the other hand draw attention to scale, redirecting the target of the analysis from the verification of spatial homogeneity to the observation of local spatial variation. Geographical regression analysis constitutes a practical example of implementing regression analysis in local level (Fotheringham & Charlton, 1998). In fact, it is a method about the ascertainment of spatial inconstancy generated by the correlation of phenomena. A wide range of literature proves that the use of GWR is prevalent in the comprehension of the process related with the interaction amongst variables. (Huang & Leung, 2002 ; Li &Rainwater, 2003)

In the chapters that follow our interest is focused in the description of the comparative application of both methods (multiple linear regression and geographically weighted regression) in a predefined case study. The area of interest in the particular analysis is a section of the urban conglomeration of the city of Volos. Field reference of the analysis is the urban block. One can easily detect the "gaps" among the results of multiple linear regression analysis that can be filled with the use of geographically weighted regression as well as the fact that observations of great importance reveal the specificities of space.

2. CONCEPTUAL PRINCIPLES – METHODS

Urban land market analysis is a complicated procedure that requires essential knowledge of the theory that is related with the processes of market which formulates urban space. Former theories concerning the configuration of land values and bibliography for the knowledge of acquired experience contribute greatly in the selection of variables that determine dependent variables (Azar et al, 1994; Dale-Johnson & Brzeski, 2001; Páez et al, 2001). In early 19^{th} century Ricardo claims that land of higher productivity is not only directly used but also appears to have a greater value (Thrall, 2002). The particular claim posed the foundations in Ricardo's theory which however has an important weakness: in the calculations, parameter "location" is not taken into consideration. In the years that follow, Von Thünen introduces the concept of geographic location in the calculation of land values replacing the concept of *productivity* as is in Ricardo's theory with the concept of *"accessibility"*. In the sixties Alonso (_, 1960) expands Von Thünen's theory with applications in urban space and underlies that distance from the centre of town is an important factor in the configuration of land values.

The majority of modern theories do not confine themselves in the calculation of distance from the commercial centre but reinforce their models with parameters such as transports, socio-economic characteristics of population , level of social prosperity, environmental characteristics etc. (Mills and MacKinnon, 1996). From among these theories, the use of methods of regression analysis dominates in the research of land values analysis and configuration, as nowadays it constitutes an unquestionable statistical tool in the determination of the degree of cross-correlation and the definition of the type of relation between different variables.

2.1. Multiple linear regression analysis

The linear statistical relation of the correlation of variables can be substituted with a linear function in order that the information of the distribution is limited in a straight line depicted by the following equation: (Maloutas, 1994; Chalkos, 2000)

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \varepsilon_i \qquad (1)$$

With Y being the dependent variable, X_1, X_2, X_3, ..., X_κ the independent variables and β_1, β_2, ..., β_κ the coefficients of regression that determine the degree with which the independent variable influences the dependent. The quality control of the line, (produced by the regression's equation), is controlled by the following indicators (others as well): multiple correlation coefficient, partial correlation coefficient, coefficient of determination and the adjusted coefficient of determination.

The correlation analysis of variables and the formation of a regression model to use are extremely complicated procedures as they premise the outflank of three problems (Demetriades, 2002): collinearity or multicollinearity, the existence of outliers and autocorrelation.

2.2. Geographically weighted regression

In geographically weighted regression the observations taken into consideration for the formation of the model are weighted with regard to their location. The direct effect is that multiple linear regression involves analysis in the global level whereas geographically weighted regression involves analysis at local level. According to Fotheringham and Wong (Rogerson, 2001) regression analysis can possibly produce different results according to the spatial reference field of the application. In the case of global reference all local particularities are overlooked (Fotheringham et al, 2000). The ascertainment that intense diversifications characterize space enhances the transition of the analysis from global level to local level with the dynamic introduction of the spatial parameter (location).

The introduction of spatial dimension in the definition of the equation has been attempted in the past (Brunsdon, 1998) with more or less substantial results for local scale spatial analysis alternatively. The present article examines the dynamics of the analysis of geographically weighted regression. The equation 1 changes form:

$$y_i = \sum_j X_{ij} \beta_j (\rho_i) + \varepsilon_i \qquad (2)$$

where ρ_i is the geographic location of observation i.

In this particular method, each parameter $\beta_j(\rho_i)$ or β_{ij} is expressed in connection with its spatial location.

At this point, calibrating the regression model involves drawing around a certain ρ_i a circle of radius r and using only the observations that are found inside the circle. The β_j produced can be considered as an estimate of cross-correlation of observations inside and around ρ_i. These are estimations of β_{ij}. Calculating β_{ij} for each ρ_i produces a total of estimates of spatial differentiated parameters (Brunsdon, 1998). Particular caution is required in the determination of the radius r.

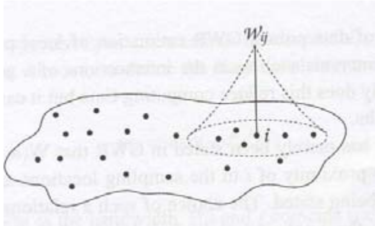

Figure 1: Schematic representation of the kernel' bandwidth and its shape (Fotheringham, 2000)

Each observation k (for a certain area ρ_i) is weighed with a weight w_{ik} (Brunsdon, 1998):

$$w_{ik} = \begin{cases} \{1-(d_{ik}/h)^2\}^2 & d_{ik} < r \\ 0 & d_{ik} > r \end{cases} \qquad (3)$$

Thus, whereas in multiple linear regression $\hat{\beta} = (X^{\mathrm{T}}X)^{-1}X^{\mathrm{T}}y$ (3), in geographically weighted regression goes as in equation 4:

$$\hat{\beta}_i = (X^{\mathrm{T}}W_iX)^{-1}X^{\mathrm{T}}W_iy \qquad (4)$$

where $W_i = \begin{pmatrix} w_{i1} & 0 & \dots & 0 \\ 0 & w_{i2} & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & w_{iN} \end{pmatrix}$ and N is the number of observations

The table of weights is a diagonal matrix whose diagonal elements correspond to the weights of the weighted regression around point p_i. Consequently the precedent function is more of a vector of equations rather than a simple equation for which each β_i corresponds in a column of the matrix whose elements are β_{ij}. This matrix is the most important outcome attributed to geographically weighted regression. The selection of each line separately aims the verification that each parameter corresponding to a different variable differs spatially. (Brunsdon, 1998)

The calculation of the kernel bandwidth is based on the least squares analysis, provided that no previous experience exists (depending on the case under review):

$$h = \sum_i \{y_i - \breve{y}_i(h)\}^2 \qquad (5)$$

Where the anticipated value of y_i is $\breve{y}_i(h)$, calculated in a geographically weighted regression model provided that for each estimate i the estimate of y_i is made without the inclusion of value I (to avoid significant statistical errors). (Brunsdon, 1998)

2.2.1 Specifying indicators

The *coefficient of determination* is a common statistical indicator of multiple linear regression and of geographically weighted regression (6).

$$r_i^2 = \frac{\sum_j w_{ij}(y_j - \bar{y})^2}{\sum_j w_{ij}(y_j - \hat{y})^2} \qquad (6)$$

Where the numerator is the total sum of squares and the denominator is the geographically weighted sum of squares of the residuals.

Two other statistical indicators measure the influence of each observation in the calculation of the model. The first indicator is well known as leverage values. If the value of the particular observation is substantial – for each observation – then the influence of the observation in the model is probably high. This indicator can also be mapped making it possible to reveal the existence of spatial patterns. The second indicator is Cook's Distance which also constitutes denotation of each observation's influence:

$$D_i = r_i^2 s_{ii} / p(1 - s_{ii}) \qquad (7)$$

where r_i is the standardized residual for point i and p is the number of parameters. Observations having a substantial influence in the model will have a value greater than one (>1).

2.2.2 Monte Carlo's significance test

The use of geographically weighted regression is justified by the following significance test (Fotheringham et al, 1998):

$$\begin{aligned} &\boldsymbol{H_0 : \beta_{ij} = \beta_j} \\ &\boldsymbol{H_1 : \beta_{ij}} \quad \textit{\textbf{spatially diversified}} \end{aligned} \qquad (8)$$

In other words, it's about the control of potential existence of local spatial differentiation of parameters. Generally this could be considered as measure of variation. For a given k let $\hat{a}_k(u_i, v_i)$ be the estimate of geographically weighted regression for $a_k(u_i, v_i)$. Be it also that we take into consideration n values of the parameter (one for each point inside the area), then the standard deviation (s_k) of the n estimates of the parameters is an estimator of the variation of the parameter.

Thereafter, follows the use of Monte Carlo test. The Monte Carlo technique is a method that creates random numbers observing how the fraction[i] reacts in every case. In effect, it is about the transformation of non random problems in random form so as to facilitate finding a solution to the problem via statistical sampling[ii].

Consequently, the observed prices of s_k can be compared with the prices that result from the random data rearrangement in space and the application of geographically weighted regression.

3. APPLICATION

Section of Volos' urban conglomeration is selected for the application of the afore-mentioned theoretical frame. The specific selection is based to a large extent to the selection of Volos as an Olympic city for the 2004 Olympic Games and to the potentials for growth of the city and the reevaluation of real estate values after the Games. The city rises dynamically from a period of slump after suffering a decade of socio-economic exhaustion. Moreover, the current geomorphology in the area reduces available space for extension inducing land values to change accordingly. The resulting model should be by all means realistic. This has been achieved with the following methodological frame.

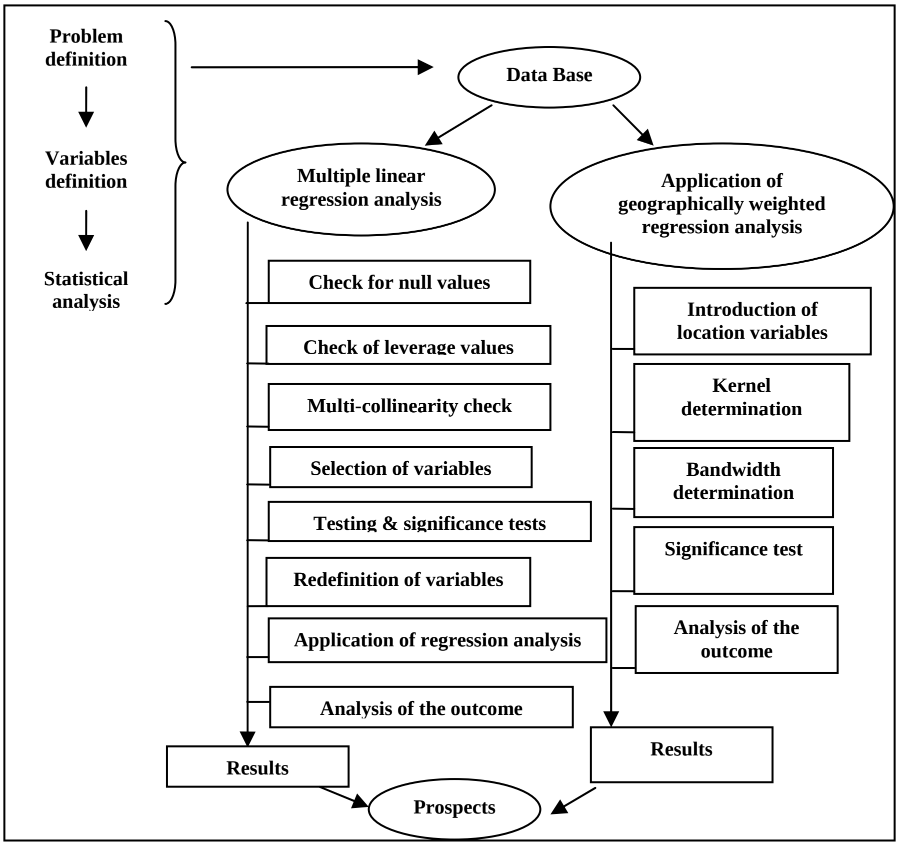

Diagram 1: Methodological diagram

The nature of the problem under review should be initially determined. A prime bibliographical analysis in the existing experience as well as in the existing theoretical background concerning the evolution of land values constitute the base for the determination of variables in use. With the completion of the data base, statistical analysis is constituted of two different processes that concern the application of two methods: multiple linear regression and geographically weighted regression. The resulting outcome defines the emerging prospects in the use of the two methods.

3.1 Data Base

The statistical analysis of variables presupposes an integrated data base. An important factor for the statistical calculations that follow is the determination of variables. Table 1 cites all variables used in the model.

Table 1: Methodological diagram

Variables	**Description**
TIMH_EU	The value for each building block
DIST_CBD	Distance of each building block from the commercial centre
SD_NEW	Indicator based on floor area ratio[iii]
DIST_UNI	Minimum distance from university buildings per building block
PROSWPO	Binary variable expressing face/no face in main artery per building block
POP	Population per building block for year 1991
PERC_1	Percentage of residence on first floor level
PERC_2	Percentage of commercial land uses on first floor level

3.2 Multiple linear regression model

Following all methodological steps the equation best approaching the cross-correlation of variables is generated with the following process: a) check for null observations and outliers, b) check for multicollinearity and autocorrelation and c) determination of the equation.

The data base contained 19 null registrations for variable “POP” that are removed from the sample. Moreover, the control for outliers brought out 121 such observations. The responsible variable for the majority of the outliers is “trade rate” and it is removed from the model.

Multicorrelation among the independent variables is intense (tables 2 and 3). On the contrary there is no autocorrelation. Moreover the removal of variable “dist_uni” will contribute in the avoidance of further problems in calculations.

Table 2: Model's indicators estimation – the Durbin Watson indicator

Model Summary[b]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	**Durbin-Watson**
1	.839[a]	.704	.702	33.4396	**1.986**

Table 3: Indicators of correlation and multi-collinearity

Coefficients[a]

Model		Correlations			Collinearity Statistics	
		Zero-order	Partial	Part	Tolerance	VIF
1	POP	.193	.016	.009	.829	1.206
	DIST_CBD	-.636	-.134	-.073	.197	5.077
	SD_NEW	.830	.678	.502	.322	3.103
	PROSWPO	.218	.068	.037	.834	1.200
	DIST_UNI	**-.638**	**.137**	**.075**	**.141**	**7.109**
	PERC_1	-.253	.100	.055	.547	1.827

From the total of seven initial independent variables only three of them participate in the resulting equation explaining 69.7% of the information of the dependent variable. A large percent of the information is explained alone by the variable "floor area ratio". The next variable that corresponds to the criteria of import in the equation is "percentage of residence in ground floor level" whereas the last variable imported in the equation is "facing a main artery".

Table 4: Regression coefficients and indicators per model

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.830[a]	.690	.689	34.1689
2	.834[b]	.696	.695	33.8416
3	.835[c]	.697	.697	33.7631

a. Predictors: (Constant), SD_NEW

b. Predictors: (Constant), SD_NEW, PERC_1

c. Predictors: (Constant), SD_NEW, PERC_1, PROSWPO

3.2.1 Conclusions

The equation resulting from the preceded analysis has as follows:

$$[TIMH\_EU] = 464{,}886 + 117{,}583 * [SD\_NEW] + 19{,}015 * [PERC\_1] + 6{,}043 * [PROSWPO] \qquad (9)$$

In practice, the equation 9 reveals a certain number of relations and correlations. Hence, the determination of land values per building block is connected to the floor area ratio, the percentage of ground floor residence and the fact that a building square is facing a main road. In all three cases, the correlation is positive creating a positive effect as they augment.

Figure 2 presents the difference between values produced by the model and the actual land values. Such a comparison can establish that the model is actually realistic.

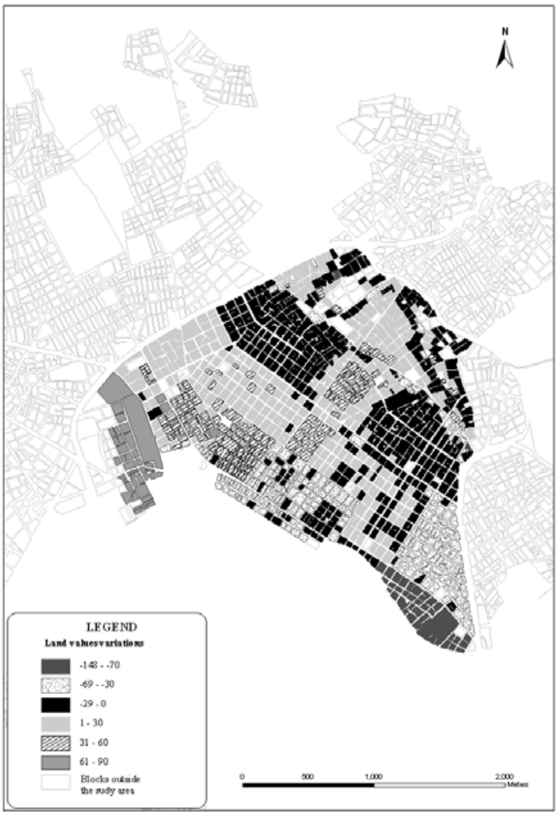

Figure 2: Comparison of actual land values and anticipated values

The question rising concerns the way prices resulting from the analysis differentiate from the land values of building blocks in the study area. Thus null prices or prices near zero indicate absolute success of model while in cases of over-estimates or undervaluations the model is proved weak to approach reality. In figure 2 one can easily observe two regions in the south-western and south-eastern department of city where land values have been overestimated and underestimated respectively. An insight look at the map makes it perceptible that the linear regression analysis achieved only partially to represent the real conditions for the configuration of land values in the city of Volos.

3.3 Geographically weighted regression analysis

Geographically weighted regression analysis responds to the vital question of spatial differentiation of parameters. Indeed the cartographic representation of results contributes to the comprehension of these differentiations as well as in their interpretation. The application revealed important differentiations of parameters of the equation 9 in the area under review.

3.3.1 Application results

The completion of the specific application required the use of computer programme GWR 2.x. For the creation of the model the kernel's form is selected to be adaptive, that is not constant at length of the study area. Regarding the bandwidth, it is selected to be minimizing the AIC criterion whereas a Monte Carlo method is used for a significance test.

The result of the analysis is a *.txt file containing all indicators of the geographically weighted regression parameters (table 5).

Table 5: Parameter estimation in GWR

Number of observations	1227
Number of independent variables	4
Kernel bandwidth	94
Number of locations that where included in the model	1227
Diagnostic information	
	226784.567
Number of parameters	97.5325762
Sigma	14.1700014
Akaike Criterion	10100.7758
Coefficient of determination	0.950766975

In this case, the number of observations coincides with the number of points taken into consideration in the calibration of the regression model and the kernel's bandwidth coincides with the number of neighbouring points included by the kernel. Also, the coefficient of determination is now 0.95, contrary to the initial 0.69. This indicator is created by the comparison of anticipated and observed values for each point of the regression analysis. The increase is in a degree anticipated regarding the difference in the degrees of freedom. Moreover the decrement of the AIC criterion indicates the superiority of the local scale model.

Thereafter, the following elements are calculated and mapped: the observed elements, the anticipated elements, the residuals, the standardized residuals, the local coefficient of determination, the influence and the Cook's distance. What follows is a test analysis of variation (ANOVA) that tests the null hypothesis that the geographically weighted regression analysis does not improve the model in global scale.

Table 6: Analysis of variation (ANOVA)

	SourceSS	DF	MS	F
OLS Residuals	1394252.3	4.00		
GWR Improvement	1167467.8	93.53	12481.9373	
GWR Residuals	226784.6	1129.47	200.7889	62.1645

The value in column "F" rejects the null hypothesis so the parameters in the local scale model are spatially differentiated and do not maintain the same value for all observations of the study area. Moreover, the significance test impoverishes the influence of variable representing face in main road artery" (table 7)

Table 7: Monte Carlo significance test results

Parameter	P-value
Intercept	0.00000
SD_NEW	0.00000
PROSWPO	0.15000
PERC_1	0.10000

3.3.2 Conclusions

In the specific application the parameters "floor area ratio" and "percentage of ground floor residence" are spatially unstable. Hence, we ought to analyse and map those parameters, whereas we consider that the parameter "facing a main artery" is constant in space as it has a p value equal to 0.15>0.1.

The parameter I has high values in the commercial centre, in the south-eastern segment as well as in the westwards of the city. Low values are observed in the southern western segment and the segment between the centre and the south-eastern segment of the city. The intense spatial differentiation of this parameter could possibly be decreased with the introduction of new variables in the model. Thus the model does not satisfactorily explain extreme values of certain observations so other elements should be considered in the calculations (such as the view, the adjacency in natural environment etc.)

In what concerns parameter (II) of the variable “floor area ratio” it presents higher values in a segment of the commercial centre as well as in the northern side of the city. In the global scale model the parameter is positive. On the other hand in the local scale model (figure 3) in three cases – areas the relation among floor area ratio and land values has a negative sign.

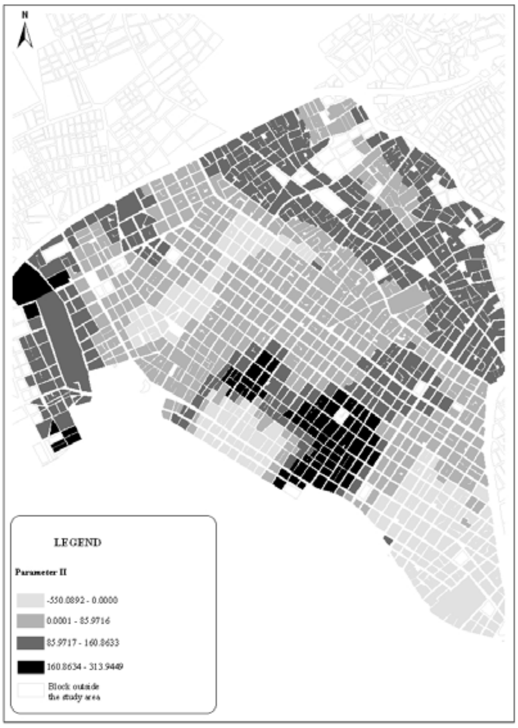

Figure 3: Cartographic representation of the results for parameter II

Regarding the spatial differentiation of parameter IV – concerning the variable “percentage of residence on ground floor level” – is intense (map 5.8). In the case of global scale analysis the parameter has a positive sign. In the local scale analysis, however, the situation is different with certain cases in which its cross-correlation with land values is negative. This is also connected to the fact that the commercial sub-centres are not included in the model. Furthermore, the variable “commercial land use on ground floor level” is removed from the model of multiple regression analysis as it presents great collinearity with other variables.

3.4 General conclusions

Conclusively, the analysis that preceded revealed the weaknesses of multiple linear regression analysis and brings to light the potential spatial differentiations in local scale. Moreover, it uncovers that the study area is in some way “divided” in sub-areas where the parameters are somewhat differentiated. The south-eastern segment of the study area is an explicit example of high values regarding the constant of the equation and low negative values for the rest of the parameters.

Contrary to the conclusions resulting from the use of multiple linear regression, the use of geographically weighted regression avoids the pitfall of the “general preview” and presents the special characteristics of the study area. In this way we are presented with the possibility of determining local weaknesses, possibilities but also new factors that were not included in the model and that however have a great influence to it.

The analysis therefore has resulted in the following conclusions:

- There exist three near the center areas where land values for each building block is negatively influenced by the high level of floor area ratio.
- There are two areas – one in the south-western boundary of the study area and one close to the commercial center – that present low initial land values and that are influenced positively by floor area ratio.
- The rest of the area follows the results of multiple linear regression with minor variations in the parameter values.

The application of geographically weighted regression has emphasized the need for redefining variables that influence the configuration of land values in the city of Volos as from the initial model only one variable is statistically significant and exhibits important spatial differentiation.

A substantial disadvantage of the application is the absence of socio-economic and environmental variables and should be taken into consideration in future practice.

4. PROSPECTS

The present article exposes the weaknesses of multiple linear regression analysis and the valuable contribution of modern statistical methods in urban space analysis. The existence of various methods as well as their continuous development strengthens their contribution in spatial analysis expanding at the same time their application field. Likewise, it has well been clarified that local scale methods outclass global scale methods.

It is evident that global models contain a large amount of spatial information (regarding spatial interaction) that emerges with the calculation of local parameters. In this way generalities are avoided and there is a greater focalization in local "exceptions". Moreover grows the possibility of representing potential variations and extracting conclusions with the aid of mapping.

The concrete case study is a typical example for analysing the operation of statistical methods in spatial analysis. The objective is the interpretation of the configuration in building blocks' land values with the use of variables that represent characteristics of location, demographic data, land uses etc..

In this paper, an abundance of problems were faced. Thus therefore, one of the issues in question is the lack of data in building block level which in fact is a common problem of research in the Greek reality. The use of personal data is prohibited whilst even available data in certain cases are problematic. Problems such as lack of data and discrepancies between data and digital backgrounds were overcame with personal processing.

On the other hand, the variables in use present significant problems of spatial autocorrelation rendering their use unavailable in the creation of a realistic problem. As a result, one of the variables presenting intense correlation with the dependent variable is removed from the initial model (variable: "Distance from commercial center"). It has been ascertained and underlined that the modulation of urban land values is connected to the distance by the commercial centre. Such questions are possible to be faced with the use of alternative methods or with the suitable modification of variables.

In what concerns the prospects for the determination of urban land values it should be marked that they are prominent and promising under certain conditions. The most important factor in the determination of a suitable and a realistic model is not so much the lack of theoretical background as mainly the lack of data that will be used in the enrichment of empirical analysis with more applications.

Consequently, it is becoming comprehensible that the urban local scale analysis permits the conception of spatial processes through various ways and offers an acceptable frame for the development of new statistical approaches in spatial analysis. Moreover, as it has already successfully been claimed (Jones and Hanham, 1995), local methods of spatial analysis can downsize the gap of opinions among realists and positivists.

---

[i]http://mathworld.wolfram.com/MonteCarloMethod.html

[ii]http://www.riskglossary.com/articles/monte_carlo_method.htm

[iii] We draw a circle with a radius of 100 meters from the centroid of each building block. For all polygons for which the circle is encompassed by polygons of floor area ratios the initial is maintained. For the remaining building squares located in the boundaries of regions with different floor area ratio the new indicator is calculated (Sum of percentage of the circles' surfaces * floor area ratio of each area respectively)

5. REFERENCES

1. **AlonsoW. (1960/1991)**, "A theory of the urban land market", in Cheshire's and Evans' *Urban and Regional Economics*, Edward Elgar Publications, UK

2. **Azar K., Ferreira J., Abed J., Kazzaz T. (1994)**, "Using GIS for estimating the spatial distribution of land value in metropolitan Beirut", *URISA* (Urban and Regional Information Association), pp.805-819

3. **Brunsdon C., Fotheringham S., Charlton M. (1998)**, "Geographically weighted regression – modeling spatial non-stationarity", *in The Statistician*, 47, Part 3, pp. 431-443

4. **Chalkos G. (2000),** *"Statistics: Theory, Applications and Use of statistic programmes in a pc",* Typothito editions, Athens

5. **Charlton M., Fotheringham S., Brunsdon C. (2000)**, "*Geographically Weighted Regression: Version 2.x. User's Manual and Installation Guide"*, University of Newcastle, UK

6. **Dale-Johnson D., Brzeski W. J. (2001)**, "Spatial Regression Analysis of Commercial Land Price Gradients", presented at the Asian Real Estate Society Sixth Annual Conference at Keio University, Tokyo, Japan

7. **Demetriades E (2002),** *"Descriptive Statistics",* Critiki Editions, Athens

8. **Fotheringham A., Brunsdon C., Charlton M. (2000)**, *"Quantitative Geography: Perspectives on Spatial Data Analysis"*, SAGE Publications, London

9. **Fotheringham A., Charlton M. (1998)**, "Geographically weighted regression: a natural evolution of the expansion method for spatial data analysis", *in Environment and Planning A*, 30, pp. 1905-1927

10. **Huang Y., Leung Y. (2002)**, "Analyzing regional industrialization in Jiangsu province using geographically weighted regression", in *Journal of Geographical Systems*, vol 4, pp. 233-249

11. **Li J., Rainwater J. (2003)**, "The real picture of land-use density and crime: a GIS application", in http://gis.esri.com/library/userconf/proc00/professional/papers

12. **Maloutas Th. (1994),** *"Quantitative Methods of Geographical Analysis: Lecture Notes",* University of Thessaly, Volos

13. **Mills E., MacKinnon J. (1973/1996)**, "Notes on the New Urban Economics", in Richardson's, Kenneth's, Button's, Nijkamp's and Park's *Analytical Urban Economics*, Edward Elgar Publishing Company, UK, pp. 3-11

14. **Páez A., Uchida T., Miyamoto K. (2001)**, «Spatial Association and Heterogeneity Issues in Land Price Models", in *Urban Studies*, Vol. 38, No. 9, pp. 1493-1508

15. **Rogerson P. (2001)**, *"Statistical Methods for Geography"*, SAGE Publications, London

16. **Thrall G. (2002)**, "Business Geography and New Real Estate Market Analysis", Oxford University Press, New York