Estimating Housing Prices: Kriging and Cokriging as an alternative.

José-María Montero, Beatriz Larraz, Gema Fernández-Avilés

University of Castilla-La Mancha (Spain)

Nowadays, housing prices is one of the burning issues in some of the current main economies in the world (such is the case of the United Estates of America, the United Kingdom, Spain, among others, as well as in Poland). This is the reason of the increasing literature about methodologies for estimating them, because the valuation of the real estate properties is crucial for fiscal issues, it influences the consumption and the rate of economic growth, etc. Model-driven orientation has been traditionally used for estimating housing prices (hedonic models, spatial autoregressive models, mixed regressive spatial autoregressive models, models with spatial dependence in the error term, spatial Durbin models, etc.). But there is another alternative approach, the data-driven approach, which includes, among others, shape functions, methods based on inverse distance weighting, kriging and cokriging, the last two ones having shown to provide excellent results due to the fact that they take into account the structure of the spatial dependence existing in the data. Having said that, the paper shows how kriging and cokriging work, and provides their equations in several situations (constant and unknown mean, existence of trend, homotopic, partial and complete heterotopic cases, etc.). It includes three different applications in the real estate market of Toledo (Spain).

JEL classification: C21, C31, R31, P25.
Keywords: kriging, cokriging, variogram, spatial dependence, real estate valuation.

1. Introduction

The importance of space as the fundamental concept underlying the essence of social science is unquestioned. Since 1950s, a large number of spatial theories and operational models have been developed, which have gradually disseminated into the practice of urban and regional policy and analysis. However, this theoretical contribution has not been matched by a similar advance in the methodology for the econometric analysis of data observed in space.

"Spatial" means each item of data has a geographical reference so we know where each case occurs on a map. This spatial indexing is important because it carries information that is relevant to the analysis of the data. The term "spatial analysis" has a background in geography that can be traced back at least the 1950s. Spatial analysis is a term widely used in the Geographical Information *Systems* (GIS) and Geographical Information *Science* (GISc) literature. A definition of spatial analysis is that represents a collection of techniques and models that explicitly use the spatial referencing associated with each data value or object that is specified within the system under study. The main idea when the spatial effects appear is how these effects can be measured. So, in this sense is very important to show that the spatial effects can be considered as special cases of general frameworks in standard econometrics, and to outline how they need a separate set of methods and techniques, encompassed within the field of spatial econometric.

The term "spatial econometrics" was coined by Jean Paelinck in the early 1970s to designate a growing body of the regional science literature that dealt primarily with estimation and testing problems encountered in the implementation of multiregional econometric models. In their book *Spatial Econometrics*, Paelink and Klaassen outline the following characteristics of the field in terms of the types of issues considered: the role of spatial interdependence in spatial models, the asymmetry in spatial relations, the importance of explanatory factors located in other spaces, and the space modelling.

On the one hand, the distinction between "spatial econometric" and "statistics econometric" is easy and essential. If activities such as the estimation of spatial interaction models, the statistical analysis of urban density functions, and the empirical implementation of regional econometric models are analysed applying standard econometric, the study of the models in question will tend to ignore specific spatial aspects. On the other hand, the distinction between "spatial econometrics" and "spatial statistics" is less straightforward, and methods tend to be categorized as belonging to one or the other field depending on the personal preference of the researcher. One possible categorization can be extracted from Haining (1986) and Anselin (1986). They refer to "the data driven orientation" in spatial statistics[1] and to "the model driven approach" in spatial econometrics[2]. Moreover, spatial econometric typically deals with models related to regional and urban economics, whereas a substantial body of the spatial statistical literature is primary focused on physical phenomena in biology and geology. This paper shows briefly "the model driven approach", studies widely "the data driven orientation" (kriging and cokriging methodologies), and provides several applications of these spatial techniques to the real estate case, due to the importance that this issue has nowadays.

Of course, it is clear the importance that the construction sector has into the economic activity, not only because of its contribution to the Gross Value Added, but also by its driving character of the rest of the economic sectors and its influence in the employment. Moreover, within the construction sector, the housing sub-sector has special characteristics due to the social impact it presents, being the house a liability of utmost importance, whose purchase requires a great economic effort. So, this is why the housing price indexes are one of the issues that worry the citizens and the economic authorities greatly. Due to the importance that the housing market has nowadays in the whole developed countries, it generates a lot of literature. Some interesting recent contributions from several points of view are, among others, the following ones: Dubin (1998) uses multiple listing data to predict house prices; Basu and Thibodeau (1998) examine spatial autocorrelation in transaction prices of single-family properties in Dallas, Texas; Gámez et al. (2000) deal with the estimation of punctual house prices in some particular locations; Din et al. (2001) compare various real estate valuation models and the manner in which they take into account environmental variables; Clapp et al. (2002) deal with spatial prediction models that take into account different sources of spatial variation; Fik el al. (2003) present an interactive variables approach and test its ability to explain price variations in an urban residential housing market; Case et al. (2004) report results from a competition on modelling spatial and temporal components of house prices; Han (2004) examines the spatial distribution pattern of housing prices in Beijing and Jakarta; Mathur et al. (2004) provide new evidence on the effects of impact fees on housing prices; Militino et al. (2004) compare different statistic techniques to analyze the behaviour of used dwelling market prices; Gelfand et al. (2004) incorporate spatial information in explaining the evolution of house prices over time; Ihlanfeldt and Shaughnessy (2004) present the results from estimating the effects of development impact fees on the prices of new and existing single-family homes; Montero and Larraz (2006a) suggest punctual kriging to estimate the price of a house in a particular location and Anderson and West (2006) use hedonic analysis to estimate the effects of proximity to open space on sales price.

[1] Representative examples are the most of the writings of Cliff and Ord (1973, 1981), Bartlett (1975), Bennett (1979), Griffith (1980, 1987), Ripley (1981), Watenberg (1985), Upton and Fingleton (1985) and Wackernagel (2003).

[2] Illustrative examples are the work of Hordijk (1974), Hordijk and Paelinck (1976), Blommestein (1983), Anselin (1980, 1988), Foster and Gorr (1986) and Fisher and Nijkamp (1984).

There is no doubt that prices of houses but also premises are spatially correlated (things that are closer are likely to be more similar than the remoter ones). This lack of independence has been widely commented, as can be seen, among others in Granelle (1970), Guigou (1982), Richardson (1986), Roca (1988), Chica (1994), Dubin (1998), Montero and Larraz (2006a) and Yiu and Tam (2007) and can be easily checked in any case throughout the existing autocorrelation test such as Moran's *I* or Geary's *c* statistics. The existence of this spatial correlation has lead us to suggest the above mentioned techniques (kriging and cokriging) as the more suitable ones when trying to estimate real estate prices.

The rest of the paper is organized as follows. Section two investigates briefly into the traditional spatial econometric approach. Section three sets up methodological issues related with the spatial statistics, suggesting kriging and cokriging as an alternative. In section four three different applications are shown. Finally, section five summarizes the paper and concludes.

2. The traditional econometric approach

As Anselin pointed out (1988), spatial effects are the essential reason for the existence of a separate field of spatial econometrics. Spatial effects can be classified in two general types: spatial dependence or spatial autocorrelation and spatial heterogeneity. While spatial dependence can be considered as the existence of a functional relationship between what happens at one point in space and what happens elsewhere, spatial heterogeneity implies that functional forms and parameters vary with location and are not homogeneous throughout the data set.

We are going to focus on spatial dependence but, due to the multidirectional nature of dependence in space, it is needed a different methodological framework than used in the one-directional situation in time. It is well known that the consequence of spatial dependence and heterogeneity is that the observations contain less information than if there had been independence. The statistical properties for estimators and hypothesis tests in the standard econometric approaches will not hold in the presence of those spatial errors. In order to obtain approximately the same degree of information as in an independent set of observations, the structural information of spatial pattern will be needed.

The main characteristic of spatial econometric is the way in which spatial effects (in our case, spatial autocorrelation) are taken into account. Typically, the use of a spatial weight matrix achieves that spatial models can be applied to many empirical contexts, provide that the spatial dependence is properly expressed in the weights and that the spatial heterogeneity is accounted for in the specification of the model. The most important taxonomies for spatial regression models have been written by Anselin (1988 and 2001) and Florax y Folmer (1992), and the rest of the section is devoted to briefly show the primary ones used when valuating properties.

The *simple regression model* (*SRM*) and *multiple linear regression model* (*MLRM*) have been widely used in the valuation of properties, but usually they did not take into account the location of properties. The introduction of the micro-localization factors and the spatial dependence within the regression models implies the definition of the concept of vicinity and the well known matrix **W** of vicinities. The ultimate objective of the use of **W** in the specification of spatial econometric models is to relate a variable at one point in space to the observations for that variable in other spatial units in the system.

There are two basic alternatives to introduce **W** in the model: including **WY** and/or **WX** as exogenous variables (substantive dependence) or including a spatial autoregressive disturbance, or both. These alternatives lead to the following models:

1) The *first-order spatial autoregressive model* (*SAR* (1)), was suggested by Besag in 1974 and consists of a linear relationship between a conditional expectation of the dependent variable and its values in the rest of the system. It can be expressed in the following form:

$$\begin{aligned} \mathbf{y} &= \rho\mathbf{W}\mathbf{y} + \mathbf{u} \\ \mathbf{u} &\approx N(\mathbf{0}, \sigma^2\mathbf{I}) \end{aligned} \tag{1}$$

in which **y** is a vector of observations on the dependent variable expressed in distances between data points and the mean, $\boldsymbol{\rho}$ the autoregressive parameter, **W** denotes a spatial weight matrix, **u** a vector of independently normally distributed errors with zero expectation and variances σ^2 and **I** the identity matrix. In *SAR* (1) the value of an observation in a point of the space is conditioned by the values of the observations in other points. In this sense, it is a pure autoregressive model. For example, in the case of properties valuation the price of a property is obtained from the prices of the properties nearby. It does not take into account exogenous variables **X**.

2) The s*patial lag model* (*SLM*) is a general spatial autoregressive model, in which explanatory variables include a spatial lag for the dependent variable as well as a set of exogenous variables. It can be expressed as:

$$\begin{aligned} \mathbf{y} &= \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \mathbf{u} \\ \mathbf{u} &\approx N(\mathbf{0}, \sigma^2\mathbf{I}) \end{aligned} \tag{2}$$

where **y** is the vector of responses, ρ is the autoregressive parameter, **Wy** is the spatial lag for **y**, **X** is a matrix of explanatory variables, β is an vector of parameters and **u** a vector of independent and normally distributed errors. *SLM* is suitable when the spatial dependence is significant and remains regardless of the explanatory variables added, then 'substantive spatial dependence' is present. This implies that the value of a variable observed at each location is truly jointly determined with that at other locations, so is unreliable the independent premise.

In the real estate case, this model could be useful to estimate, for example, housing values, when data about explanatory variables (quality of the construction, age of the building, width of the street in which is located, and height in the case of buildings, if it has or has not a lift, the number of bathrooms, etc.) are available but the price of the houses not only depend on such as exogenous variables but also on the price of properties in their neighbourhood.

3) The *spatial autoregressive model with spatial error dependence* consists of a linear relationship between a conditional expectation of the dependent variable and its values in the rest of the system with spatial dependent in the error terms. It can be written as:

$$\begin{aligned} \mathbf{y} &= \rho \mathbf{W}_1 \mathbf{y} + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W}_2 \mathbf{u} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &\approx N(\mathbf{0}, \sigma^2 \mathbf{I}); \end{aligned} \tag{3}$$

where $\mathbf{y}$ is the vector of observations on the dependent variable, ρ is the autoregressive coefficient, $\mathbf{W}$ denotes a spatial weight matrix, λ is the coefficient in a spatial autoregressive structure for the disturbance $\mathbf{u}$, ε is a vector of independently normally distributed errors whit zero expectation and variances σ^2 and $\mathbf{I}$ the identity matrix.

4) In the l*inear regression model with a spatial autoregressive disturbance* or *spatial error model (SEM)* the explanatory variables contain only exogenous variables but the error term follows a spatial autoregressive process. It is usually expressed as

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{u} \\ \mathbf{u} &= \lambda \mathbf{W} \mathbf{u} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} &\approx N(\mathbf{0}, \sigma^2 \mathbf{I}) \end{aligned} \tag{4}$$

In this specification, β is a vector of parameters associated with exogenous variables $\mathbf{X}$ and λ the coefficient in a spatial autoregressive structure for the spatially correlated errors. It can also be considered as a second order spatial autoregressive structure for the error term

$$\begin{cases} \mathbf{u} = \lambda_1 \mathbf{W}_3^1 \mathbf{u} + \lambda_2 \mathbf{W}_3^2 \mathbf{u} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} \approx N(\mathbf{0}, \sigma^2 \mathbf{I}) \end{cases} \tag{5}$$

where $\mathbf{W}^1, \mathbf{W}^2$ are spatial interaction matrices, or a moving average structure

$$\begin{cases} \mathbf{u} = \theta \mathbf{W}_3 \boldsymbol{\varepsilon} + \boldsymbol{\varepsilon} \\ \boldsymbol{\varepsilon} \approx N(\mathbf{0}, \sigma^2 \mathbf{I}) \end{cases} \tag{6}$$

where θ is a scalar parameter associated to $\mathbf{W}\boldsymbol{\varepsilon}$.

5) By combining the *SLM* and the *spatial autoregressive model with spatial error dependence*, and adding one or more spatially lagged exogenous variables it is

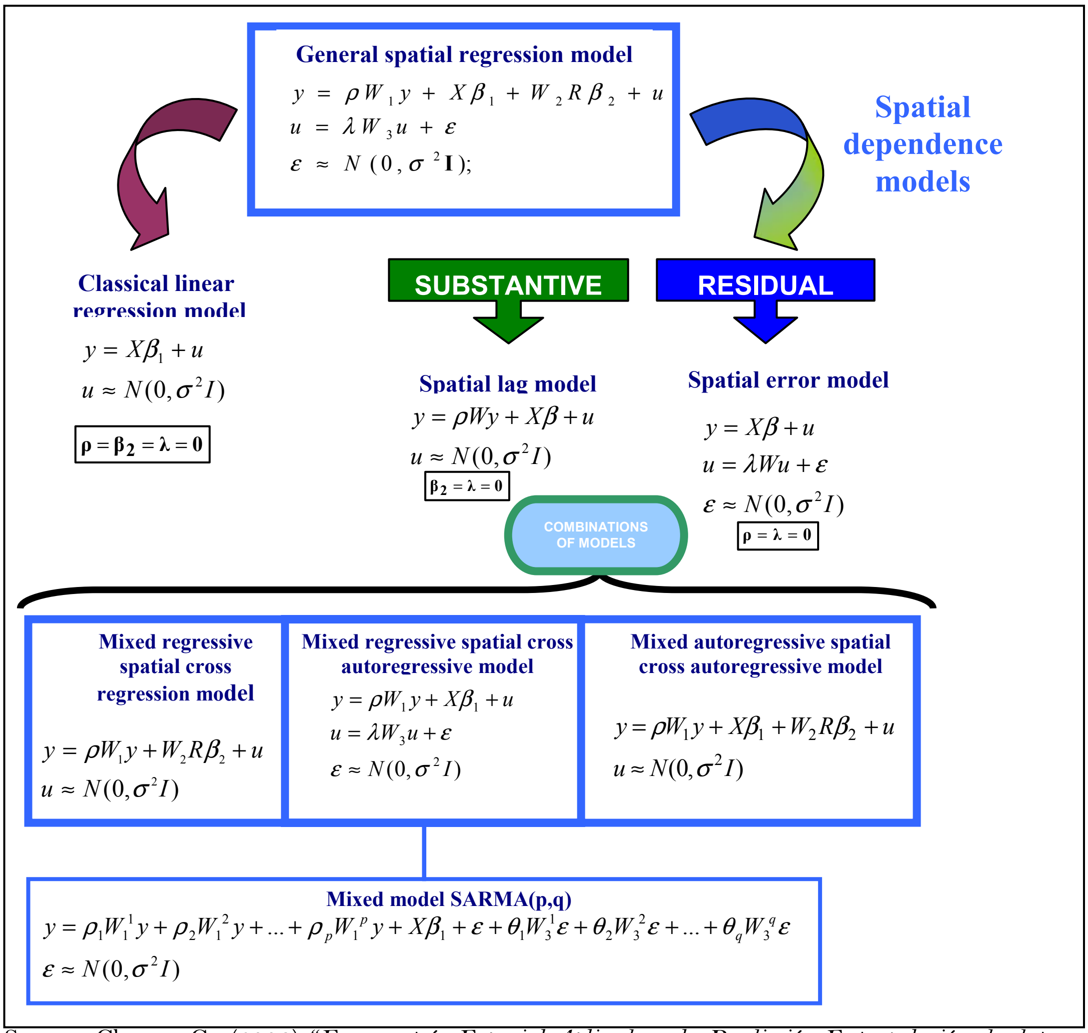
obtained the *mixed-regressive-spatial autoregressive model with a spatial autoregressive disturbance* or *general spatial regression model* (*GSRM*):

$$
\begin{aligned}
\mathbf{y} &= \rho\mathbf{W}_1\mathbf{y} + \mathbf{X}\beta_1 + \mathbf{W}_2\mathbf{R}\beta_2 + \mathbf{u} \\
\mathbf{u} &= \lambda\mathbf{W}_3\mathbf{u} + \boldsymbol{\varepsilon} \\
\boldsymbol{\varepsilon} &\approx N(\mathbf{0}, \sigma^2\mathbf{I});
\end{aligned}
\tag{7}
$$

where **R** is the matrix of the exogenous variables, which could entirely match up with the matrix **X** or not.. Obviously, the number of specifications can be duplicated by allowing heteroskedasticity of a specific form.

Taking as point of departure *GSRM* and setting to zero sub-vectors of the parameter vector $\boldsymbol{\Theta} = [\boldsymbol{\rho}, \boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \boldsymbol{\lambda}]$ the above mentioned models are easily obtained (among others), as can be seen in Figure 1.

Figure 1: Taxonomy of models with spatial dependence



Source: Chasco. C., (2003) "*Econometría Espacial Aplicada a la Predicción-Extrapolación de datos Microterritoriales*". Tesis Doctoral. Universidad Autónoma de Madrid.

3. Kriging and cokriging as as alternative.

These approaches facilitate realistic modelling that account for the complexity of natural and economic phenomena and help solve economic and development problems in mining, oil exploration, geology, hydrology, real estate, environmental engineering, and other real-world situations involving spatial uncertainty. In real world, it is impossible to get exhaustive values of data at every desired point because of practical or economic constraints. Thus, interpolation is important and fundamental to graphing, analysing and understanding of two-dimensional data. Among all the existing interpolation methods, kriging and cokriging takes advantage of the structure of the spatial dependence that the phenomena present. In this sense, kriging is a method of interpolation which estimates unknown values from data observed at known locations (univariate approach) while cokriging estimates unknown values using also additional information about other phenomena correlated with the main one (multivariate approach).

3.1. Kriging methods: Univariate approach

Kriging is a geostatistical technique that allows us to interpolate the value of a random function at an unobserved location from observations of its value at nearby locations. There are several natural and economic phenomena such as ore body —the extent and value of ore where it is located— , extraction of petroleum or coal, mineral concentrations, soil moisture, house prices, premises prices, among others, which can be considered as a random function or stochastic process, $X(\mathbf{s})$ (i.e. an infinite family of random variables constructed at all points of a given region) , being **s** the different locations where the data are or could been observed (e.g. where minerals are located or houses are sited). It means that each variable $X(\mathbf{s}_i)$ takes different values depending on its spatial location. This function used to be spatially autocorrelated, being the sample a particular realization $x(\mathbf{s})$ of $X(\mathbf{s})$. When this correlation between the variables exists, it is necessary to take it into account for estimating the mean value of the phenomena and for making punctual estimations at an unobserved location.

In this case, when the variables of the random function are spatially correlated, kriging is considered as a suitable estimation method because it incorporates the structure of that spatial correlation computing the best linear unbiased estimator. Depending on the purpose, it is necessary to take into account two versions of kriging: kriging the mean to estimate the mean value over a region (section 3.1.1.1.) and punctual kriging to estimate a value at a particular point of a region (section 3.1.2). And depending on the kind of stochastic process we deal with, three different types of punctual kriging can be distinguished: simple kriging —if the random function is second-order stationary with known mean—, ordinary kriging —if the random function is second-order or intrinsically stationary with unknown mean— and universal kriging —if the random function is non-stationary.

3.1.1. The importance of spatial correlation when estimating parameters.

The presence of spatial correlation between the variables of a random function or stochastic process affects to the traditional methodology used to estimate the population parameters. The estimator variance varies from the case of independence to the case of spatial dependence. Assuming that we deal with a random function, which presents positive autocorrelation, it will imply a higher estimator variance when estimating the population mean. Therefore, if the estimators that have been designed for the independence case are used when positive autocorrelation exists (the real estate case) then, the confidence intervals of the mean value in the interested area will be too narrow.

As an example, suppose that $X(s_1), X(s_2), \ldots, X(s_n)$ are random variables denoting the price of the houses located, respectively, in the geographic locations $s_1, s_2, \ldots, s_n$ of a domain. Suppose that every variable are independent and has unknown mean μ and known variance σ^2. The well-known unbiasedness and minimum variance estimator of the parameter μ is:

$$\overline{X} = \sum_{i=1}^{n} \frac{X(s_i)}{n} \tag{8}$$

with mean μ and variance $\dfrac{\sigma^2}{n}$.

Suppose now the more realistic assumption in the spatial case that the variables are positive autocorrelated and that such dependence decreases with the distance between them. This implies, for example, that[3]

$$C\big(X(s_i), X(s_j)\big) = \sigma^2 \rho^{|s_i - s_j|} \quad i, j = 1,2,\ldots,n \quad 0 < \rho < 1 \tag{9}$$

being the estimator variance

$$\begin{aligned} V(\overline{X}) &= \frac{1}{n^2} V\left(\sum_{i=1}^{n} X(s_i) \right) = \frac{1}{n^2}\left[n\sigma^2 + 2\sum_{i \neq j} \sigma^2 \rho^{|s_i - s_j|} \right] = \\ &= \frac{\sigma^2}{n} + 2\,\frac{\sigma^2}{n^2}\left[(n-1)\rho^1 + (n-2)\rho^2 + \ldots + \left[n-(n-1)\rho^{n-1} \right] \right] = \\ &= \frac{\sigma^2}{n} + 2\,\frac{\sigma^2}{n^2} \sum_{h=1}^{h=n-1} [n-h]\,\rho^h = \end{aligned}$$

---

[3] The covariance function has been obtained from the classic example writen by Cressie (1993), p. 14. This function is obtained from an autorregresive first-order processs $X(s_i) = \rho\, X(s_{i-1}) + \varepsilon(s_i) \quad i \geq 1$ where $\varepsilon(s_i)$ are random independent and equally distributed Gaussian variables with nule mean and variance $\sigma^2(1-\rho^2)$, independent from $X(s_{i-1})$. Nevertheless, in this paper $\mathbf{X} \equiv \big(X(s_1), \ldots, X(s_n)\big)$ are spatial data in ú[1] so that the prediction of $X(s_0)$ or $X(s_{3/2})$ is so appropriated as the one from $X(s_{n+1})$.

$$= \frac{\sigma^2}{n}\left[1+\frac{2}{n}\sum_{h=1}^{h=n-1}[n-h]\rho^h\right]=$$

$$= \frac{\sigma^2}{n}\left[1+2\frac{\rho}{1-\rho}\left(1-\frac{1}{n}\right)-2\left(\frac{\rho}{1-\rho}\right)^2\frac{1-\rho^{n-1}}{n}\right] \tag{10}$$

From the above expression it is deduced that the term in square brackets is greater than one if $\rho > 0$—as in the real estate case— while is lower if $\rho < 0$. Therefore, the presence of spatial dependence between the housing prices makes the estimator variance higher respect to the independence case if $\rho > 0$, and lower if $\rho < 0$ —though this is not the housing case. In other words, when spatial correlation exists, the variance of the arithmetic mean can be expressed as follows

$$V(\bar{X}) = \frac{\sigma^2}{n'} \tag{11}$$

being n' the number of independent observations that are equivalent to a certain number of observations, n, with the spatial dependence abovementioned.

$$n' = \frac{n}{1+\frac{2}{n}\sum_{h=1}^{n-1}(n-h)\rho^h} = \frac{n}{1+2\frac{\rho}{1-\rho}\left(1-\frac{1}{n}\right)-2\left(\frac{\rho}{1-\rho}\right)^2\frac{1-\rho^{n-1}}{n}} \tag{12}$$

So, the positive spatial autocorrelation affects substantially to the accuracy of the mean estimator. And it happens even when the sample size is high. Note that, in this case,

$$n' \approx \frac{n}{\frac{1+\rho}{1-\rho}} \tag{13}$$

Of course, the presented results depend on the correlation model that has been chosen. There are simpler covariance models —for example, the perfect spatial correlation— or more complex ones, but all they drive us to the same conclusion (see, for example, Grenander, 1954, Grenander and Rosenblatt, 1984 and Haining, 1988): The arithmetic sample mean lose accuracy estimating the population mean, increasing it with the spatial correlation. It means that in the case that a confidence interval was computed, if spatial correlation exists and it is not considered, we would obtain too narrow intervals. In other words, if we maintain the size of the confidence interval built under the independence assumption, the confidence level would be reduced.

Faced with the accuracy lost by the arithmetic mean estimating the population mean under the presence of positive spatial correlated variables, it can be deduced that this estimator is not the best linear one. In this sense, the aim will be to determine which the best among the linear ones is.

3.1.1.1. Kriging the mean as an alternative

Having been exposed that the arithmetic mean is not a BLUE estimator of the mean value over a region when sample variables are not independent, in the case of spatial correlation D.G. Krige proposes the following weighted mean:

$$\mu^* = \sum_{i=1}^{n} \lambda_i X(\mathbf{s}_i) \tag{14}$$

where the "best" weights are obtained by minimizing the variance of the estimation error respecting the unbiasedness condition. In the real estate case, $X(\mathbf{s}_i)$ is the price per square meter in the location $\mathbf{s}_i$, being n the sample size.

In order to determine the weights λ_i, assuming that the stochastic process is second-order stationary with constant mean μ, it would be desirable that the estimation error was nil in average, $E[\mu^* - \mu] = 0$.

This unbiasedness condition would be held if and only if the weights λ_i sum up to one. Indeed,

$$E[\mu^* - \mu] = E\left[\sum_{i=1}^{n} \lambda_i X(\mathbf{s}_i) - \mu\right] = \sum_{i=1}^{n} \lambda_i E[X(\mathbf{s}_i)] - \mu = \mu \sum_{i=1}^{n} \lambda_i - \mu = 0$$

$$\Leftrightarrow \sum_{i=1}^{n} \lambda_i = 1 \tag{15}$$

It would be also desirable the variance of the estimation error to be minimum under the unbiasedness condition. Calculating the variance as follows,

$$V(\mu^*) = E\left[(\mu^*)^2\right] - \mu^2 = E\left[\left(\sum_{i=1}^{n} \lambda_i X(\mathbf{s}_i)\right)^2\right] - \mu^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j E\left[X(\mathbf{s}_i) \cdot X(\mathbf{s}_j)\right] - \mu^2 =$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C\left[X(\mathbf{s}_i), X(\mathbf{s}_j)\right] + \mu^2 - \mu^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C\left[X(\mathbf{s}_i), X(\mathbf{s}_j)\right] \tag{16}$$

we consider the Lagrange function, φ, and we obtain the weights λ_i by the method of Lagrange Multipliers

$$\varphi(\lambda_i, \alpha) = V(\mu^* - \mu) - 2\alpha\left(\sum_{i=1}^{n} \lambda_i - 1\right) \tag{17}$$

Setting the partial derivative with respect to the Lagrange multiplier and the weights λ_i to zero it is obtained the following equation system

$$\begin{cases} \dfrac{\partial \varphi(\lambda_i, \alpha)}{\partial \lambda_i} = \displaystyle\sum_{j=1}^{n} \lambda_j C\left[X(\mathbf{s}_i), X(\mathbf{s}_j)\right] - \alpha = 0 \quad \forall i = 1, \ldots, n \\ \dfrac{\partial \varphi(\lambda_i, \alpha)}{\partial \alpha} = \displaystyle\sum_{i=1}^{n} \lambda_i - 1 = 0 \end{cases} \tag{18}$$

or with the usual notations[4],

$$\begin{cases} \sum_{j=1}^{n} \lambda_j C(\mathbf{s}_i - \mathbf{s}_j) - \alpha = 0, \quad \forall i = 1,\dots,n \\ \sum_{j=1}^{n} \lambda_j = 1 \end{cases} \tag{19}$$

In matrix form it can be written as

$$\begin{pmatrix} C(\mathbf{0}) & C(\mathbf{s}_1-\mathbf{s}_2) & C(\mathbf{s}_1-\mathbf{s}_3) & \cdots & C(\mathbf{s}_1-\mathbf{s}_n) & -1 \\ C(\mathbf{s}_2-\mathbf{s}_1) & C(\mathbf{0}) & C(\mathbf{s}_2-\mathbf{s}_3) & \cdots & C(\mathbf{s}_2-\mathbf{s}_n) & -1 \\ C(\mathbf{s}_3-\mathbf{s}_1) & C(\mathbf{s}_3-\mathbf{s}_2) & C(\mathbf{0}) & \cdots & C(\mathbf{s}_3-\mathbf{s}_n) & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C(\mathbf{s}_n-\mathbf{s}_1) & C(\mathbf{s}_n-\mathbf{s}_2) & C(\mathbf{s}_n-\mathbf{s}_3) & \cdots & C(\mathbf{0}) & -1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{20}$$

where the optimal $\lambda_1, \lambda_2, \dots, \lambda_n$ can be obtained from

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \\ \alpha \end{pmatrix} = \begin{pmatrix} C(\mathbf{0}) & C(\mathbf{s}_1-\mathbf{s}_2) & C(\mathbf{s}_1-\mathbf{s}_3) & \cdots & C(\mathbf{s}_1-\mathbf{s}_n) & -1 \\ C(\mathbf{s}_2-\mathbf{s}_1) & C(\mathbf{0}) & C(\mathbf{s}_2-\mathbf{s}_3) & \cdots & C(\mathbf{s}_2-\mathbf{s}_n) & -1 \\ C(\mathbf{s}_3-\mathbf{s}_1) & C(\mathbf{s}_3-\mathbf{s}_2) & C(\mathbf{0}) & \cdots & C(\mathbf{s}_3-\mathbf{s}_n) & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C(\mathbf{s}_n-\mathbf{s}_1) & C(\mathbf{s}_n-\mathbf{s}_2) & C(\mathbf{s}_n-\mathbf{s}_3) & \cdots & C(\mathbf{0}) & -1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \tag{21}$$

Finally, to obtain the estimator variance, the expression (19) is included in (16)

$$V(\mu^* - \mu) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C(\mathbf{s}_i - \mathbf{s}_j) = \sum_{i=1}^{n} \lambda_i \sum_{j=1}^{n} \lambda_j C(\mathbf{s}_i - \mathbf{s}_j) = \sum_{i=1}^{n} \lambda_i \alpha = \alpha \tag{22}$$

Note that in case of spatially independent data, the system for the kriging of the mean simplifies to:

$$\begin{cases} \lambda_i V[X(\mathbf{s})] = \alpha, \quad \forall i = 1,\dots,n \\ \sum_{i=1}^{n} \lambda_i = 1 \end{cases} \tag{23}$$

where all weights are equal to $\lambda_i = \alpha / V[X(\mathbf{s})]$ and sum up to one, so, all of them have to be $1/n$. Therefore, in the independent case, the kriging estimator of the mean becomes the arithmetic mean being the minimum variance

$$V(\mu^* - \mu) = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C(\mathbf{s}_i - \mathbf{s}_j) = \sum_{i=1}^{n} \lambda_i^2 \cdot V[X(\mathbf{s}_i)] = \frac{1}{n^2} n\, \sigma^2 = \frac{\sigma^2}{n} \tag{24}$$

---

[4] $C[X(\mathbf{s}_i), X(\mathbf{s}_j)] = C(\mathbf{s}_i - \mathbf{s}_j)$ because the covariance function only depends on the distance between the pair of locations under the assumption of second-order stationarity.

Before ending this section, it is important to stress two issues:
i) Solving the equations system (20) it can be appreciate that the weights λ_i does not depend on the available data but on the structure of the spatial dependence of the phenomena.
ii) It is necessary to have at our disposal a covariance function, $C(\mathbf{h})$, being **h** the distance (length and direction) that links the locations. This covariance function can be estimated from the classical estimator being adjusted, subsequently, to a theoretical model among the existing ones (e.g. see Emery, 2000).

3.1.2. Punctual Kriging

The more important aim of this paper is to estimate $x^*(\mathbf{s}_0)$, a value of the random variable in any location $\mathbf{s}_0$ out of the sample. To deal with this aim, it is proposed the kriging estimator

$$X^*(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i X(\mathbf{s}_i), \tag{25}$$

a weighted average of the prices in the sample locations $\{\mathbf{s}_i, i = 1,\ldots,n\}$. As we wish it to be the BLUE one, the error estimation has to be nil in average and minimum variance. Depending on the kind of random variable we deal with, different types of punctual kriging can be distinguised: simple kriging, ordinary kriging and universal kriging.

3.1.2.1. Simple Kriging

If the random function $X(\mathbf{s})$ is second-order stationary with known mean, μ, it can be written as $X(\mathbf{s}) = \mu + e(\mathbf{s})$, where $e(\mathbf{s})$ has null expected value. Since the mean μ is constant and known, it can be defined a new regionalized variable $Y(\mathbf{s}) = X(\mathbf{s}) - \mu$ with $E[Y(\mathbf{s})] = 0$ and a simple procedure to estimate $X(\mathbf{s}_0)$ consists of estimating the zero-mean variable $Y(\mathbf{s})$ in $\mathbf{s}_0$ by the linear estimator $Y^*(\mathbf{s}_0) = \sum_{i=1}^{n} \lambda_i Y(\mathbf{s}_i)$ adding the mean afterwards.

In this simple case, the estimation error is always zero on average for every set of weights, because $E[Y(\mathbf{s})] = 0$. To obtain the set of weights that minimize the variance of the estimation error, it must be expressed as follows,

$$\begin{aligned} & V[Y^*(\mathbf{s}_0) - Y(\mathbf{s}_0)] = E\left[(Y^*(\mathbf{s}_0) - Y(\mathbf{s}_0))^2\right] = \\ & = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j E[Y(\mathbf{s}_i)Y(\mathbf{s}_j)] - 2\sum_{i=1}^{n} \lambda_i E[Y(\mathbf{s}_i)Y(\mathbf{s}_0)] + E\left[(Y(\mathbf{s}_0))^2\right] = \\ & = \sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i \lambda_j C(\mathbf{s}_i - \mathbf{s}_j) - 2\sum_{i=1}^{n} \lambda_i C(\mathbf{s}_i - \mathbf{s}_0) + C(\mathbf{0}) \end{aligned} \tag{26}$$

where $C(\mathbf{s}_i-\mathbf{s}_j)$ refers to the covariance between the random variables sited in the locations $\mathbf{s}_i$ and $\mathbf{s}_j$, $C(\mathbf{s}_i-\mathbf{s}_0)$ refers to the covariance between the random variables in the locations $\mathbf{s}_i$ and the location of interest $\mathbf{s}_0$, and $C(\mathbf{0})$ is the variance of the random function.

The minimization procedure leads to the following system with *n* equations and *n* unknowns

$$\sum_{j=1}^{n}\lambda_j C(\mathbf{s}_i-\mathbf{s}_j)=C(\mathbf{s}_i-\mathbf{s}_0), \forall i=1,\ldots,n \tag{27}$$

This system can be expressed in matrix notations as $\mathbf{C}\cdot\boldsymbol{\lambda}=\mathbf{C}_0$, where

$$\mathbf{C}=\begin{pmatrix} C(\mathbf{0}) & C(\mathbf{s}_1-\mathbf{s}_2) & \cdots & C(\mathbf{s}_1-\mathbf{s}_n) \\ C(\mathbf{s}_2-\mathbf{s}_1) & C(\mathbf{0}) & \cdots & C(\mathbf{s}_2-\mathbf{s}_n) \\ \vdots & \vdots & \ddots & \\ C(\mathbf{s}_n-\mathbf{s}_1) & C(\mathbf{s}_n-\mathbf{s}_2) & \cdots & C(\mathbf{0}) \end{pmatrix}, \boldsymbol{\lambda}=\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\end{pmatrix}, \mathbf{C}_0=\begin{pmatrix} C(\mathbf{s}_1-\mathbf{s}_0)\\ C(\mathbf{s}_2-\mathbf{s}_0)\\ \vdots\\ C(\mathbf{s}_n-\mathbf{s}_0)\end{pmatrix} \tag{28}$$

and its solution is $\boldsymbol{\lambda}=\mathbf{C}^{-1}\cdot\mathbf{C}_0$.

Including this result in (26), the estimation variance gives

$$V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]=V[Y^*(\mathbf{s}_0)-Y(\mathbf{s}_0)]=C(\mathbf{0})-\sum_{i=1}^{n}\lambda_i C(\mathbf{s}_i-\mathbf{s}_0) \tag{29}$$

3.1.2.2. Ordinary Kriging

If the random function is second-order or intrinsically stationary, with an unknown mean, ordinary kriging is the suitable estimation method. When the mean is unknown, unbiasedness is not guaranteed and it implies to impose the condition $\sum_{i=1}^{n}\lambda_i=1$,

$$\begin{aligned} E[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)] &= E\left[\sum_{i=1}^{n}\lambda_i X(\mathbf{s}_i)-X(\mathbf{s}_0)\right]=\sum_{i=1}^{n}\lambda_i E[X(\mathbf{s}_i)]-E[X(\mathbf{s}_0)]= \\ &=\mu\sum_{i=1}^{n}\lambda_i-\mu=0\Leftrightarrow\sum_{i=1}^{n}\lambda_i=1 \end{aligned} \tag{30}$$

Under this condition, the variance of the estimation error can be expressed in variogram terms both if the random function is second-order stationary or if it is intrinsically stationary[5].

---

[5] See Isaaks and Srivastava (1989) for ordinary kriging system in covariance terms when the random function is second-order stationary.

$$V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]=E\left[(X^*(\mathbf{s}_0)-X(\mathbf{s}_0))^2\right]=E\left[\left(\sum_{i=1}^{n}\lambda_i X(\mathbf{s}_i)-X(\mathbf{s}_0)\right)^2\right]=$$
$$=E\left[\left(\sum_{i=1}^{n}\lambda_i\left(X(\mathbf{s}_i)-X(\mathbf{s}_0)\right)\right)^2\right]=$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j E\left[(X(\mathbf{s}_i)-X(\mathbf{s}_0))(X(\mathbf{s}_j)-X(\mathbf{s}_0))\right] \qquad (31)$$

But adding and subtracting $X(\mathbf{s}_0)$ in the variogram expression it follows that

$$\gamma(\mathbf{s}_i-\mathbf{s}_j)=\frac{1}{2}E\left[(X(\mathbf{s}_i)-X(\mathbf{s}_j))^2\right]=\frac{1}{2}E\left\{\left[(X(\mathbf{s}_i)-X(\mathbf{s}_0))-(X(\mathbf{s}_j)-X(\mathbf{s}_0))\right]^2\right\}=$$
$$=\frac{1}{2}E\left[(X(\mathbf{s}_i)-X(\mathbf{s}_0))^2\right]+\frac{1}{2}E\left[(X(\mathbf{s}_j)-X(\mathbf{s}_0))^2\right]-E\left[(X(\mathbf{s}_i)-X(\mathbf{s}_0))(X(\mathbf{s}_j)-X(\mathbf{s}_0))\right]=$$
$$=\gamma(\mathbf{s}_i-\mathbf{s}_0)+\gamma(\mathbf{s}_j-\mathbf{s}_0)-E\left[(X(\mathbf{s}_i)-X(\mathbf{s}_0))(X(\mathbf{s}_j)-X(\mathbf{s}_0))\right] \qquad (32)$$

where

$$E\left[(X(\mathbf{s}_i)-X(\mathbf{s}_0))(X(\mathbf{s}_j)-X(\mathbf{s}_0))\right]=\gamma(\mathbf{s}_i-\mathbf{s}_0)+\gamma(\mathbf{s}_j-\mathbf{s}_0)-\gamma(\mathbf{s}_i-\mathbf{s}_j) \qquad (33)$$

Finally the variance can be written as:

$$V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]=$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_i-\mathbf{s}_0)+\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_j-\mathbf{s}_0)-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_i-\mathbf{s}_j) \qquad (34)$$

Including the constraint on the weights, the equation (34) simplifies to

$$V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]=2\sum_{i=1}^{n}\lambda_i\gamma(\mathbf{s}_i-\mathbf{s}_0)-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma(\mathbf{s}_i-\mathbf{s}_j) \qquad (35)$$

The optimal weights are obtained from standard Lagrangian techniques by setting each of the n partial derivatives

$$\partial(\varphi(\lambda_i,\alpha))=\partial\left(\frac{1}{2}V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]-\alpha\left(\sum_{i=1}^{n}\lambda_i-1\right)\right) \qquad (36)$$

with respect to λ_i and α to zero. This procedure provides a system of $n+1$ linear equations in $n+1$ unknowns (the n weights and the Lagrange parameter) which is called ordinary kriging system.

$$\begin{cases}\sum_{j=1}^{n}\lambda_j\gamma(\mathbf{s}_i-\mathbf{s}_j)+\alpha=\gamma(\mathbf{s}_i-\mathbf{s}_0), \forall i=1,\dots,n\\ \sum_{i=1}^{n}\lambda_i=1\end{cases} \qquad (37)$$

This system can be expressed in matrix notations as $\mathbf{\Gamma}\cdot\boldsymbol{\lambda}=\mathbf{\Gamma}_0$ where

$$\mathbf{\Gamma}=\begin{pmatrix}\gamma(\mathbf{0}) & \gamma(\mathbf{s}_1-\mathbf{s}_2) & \cdots & \gamma(\mathbf{s}_1-\mathbf{s}_n) & 1\\ \gamma(\mathbf{s}_2-\mathbf{s}_1) & \gamma(\mathbf{0}) & \cdots & \gamma(\mathbf{s}_2-\mathbf{s}_n) & 1\\ \vdots & \vdots & \ddots & \vdots & \vdots\\ \gamma(\mathbf{s}_n-\mathbf{s}_1) & \gamma(\mathbf{s}_n-\mathbf{s}_2) & \cdots & \gamma(\mathbf{0}) & 1\\ 1 & 1 & \cdots & 1 & 0\end{pmatrix},\boldsymbol{\lambda}=\begin{pmatrix}\lambda_1\\ \lambda_2\\ \vdots\\ \lambda_n\\ \alpha\end{pmatrix},\mathbf{\Gamma}_0=\begin{pmatrix}\gamma(\mathbf{s}_1-\mathbf{s}_0)\\ \gamma(\mathbf{s}_2-\mathbf{s}_0)\\ \vdots\\ \gamma(\mathbf{s}_n-\mathbf{s}_0)\\ 1\end{pmatrix} \quad (38)$$

with solution $\boldsymbol{\lambda}=\mathbf{\Gamma}^{-1}\cdot\mathbf{\Gamma}_0$. Finally the minimum estimation variance can be written as

$$V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]=\sum_{i=1}^{n}\lambda_i\gamma(\mathbf{s}_i-\mathbf{s}_0)+\alpha \quad (39)$$

The variogram $\gamma(\mathbf{h})$ necessary to solve this ordinary kriging system is defined as the expected squared increment of the values between a pair of locations at distance **h:** $\gamma(\mathbf{h})=\frac{1}{2}V[X(\mathbf{s}+\mathbf{h})-X(\mathbf{s})]=\frac{1}{2}E\left[(X(\mathbf{s}+\mathbf{h})-X(\mathbf{s}))^2\right]$. It summarises the relationship between differences in pairs of measurements and the distance of the corresponding points from each other. It has to be estimated (from the observed sample values) fitting to the experimental variogram, usually calculated by the classic moments estimator (Matheron, 1965)[6] $\gamma^*(\mathbf{h})=\frac{1}{2N(\mathbf{h})}\sum_{\alpha=1}^{N(\mathbf{h})}(X(\mathbf{s}_\alpha+\mathbf{h})-X(\mathbf{s}_\alpha))^2$ being $N(\mathbf{h})$ the number of pairs at distance **h**, one or some theoric variogram models following the linear model of regionalization (Wackernagel, 2003)[7]. The usual variogram models can be seen in Emery (2000). The variogam depicts the structure of the spatial correlation that the process presents. It implies that a correct choice of the theoretical variograms is essential to obtain good estimates.

3.1.2.3. Universal Kriging

A significant problem with punctual kriging is that it will not work unless the regionalized variable is stationary. In the presence of a trend or a drift, a linear estimation is no longer unbiased. The computed estimates will be systematically shifted upwards and downwards from the true values upon the arrangement of the points and the direction of dip of the surface.

The universal kriging (also called unbiased kriging of order *k*) provides an unbiased linear estimator which takes the drift into account, provided that both the form of the drift and the variogram of the non-stationary random function are known. In this particular situation, the random function can be regarded as having two components: a deterministic one, the drift term $\mu(\mathbf{s})$, and an

[6] Cressie and Hawkins (1980), Amstrong, (1984), Omre, (1984), Dowd, (1984) provide alternative estimators.

[7] See Chilès and Delfiner 1999, pp. 104-124.

stochastic component, the residual term $e(\mathbf{s})$, which can be treated as an intrinsically stationary process with null expected value[8], $E(e(\mathbf{s}))=0$.

$$X(\mathbf{s})=\mu(\mathbf{s})+e(\mathbf{s}) \tag{40}$$

If, as usual, the drift is supposed to be a linear combination of any known functions $\mu(\mathbf{s})=\sum_{h=1}^{p} a_h f_h(\mathbf{s})$, where $\{f_h(\mathbf{s}), h=1,\ldots,p\}$ are p known functions, a_h constant coefficients and p the number of terms used in the approximation, then, the unbiasedness condition implies that $\sum_{i=1}^{n}\lambda_i\mu(\mathbf{s}_i)=\mu(\mathbf{s}_0)$. Taking the drift expression into account

$$\sum_{i=1}^{n}\lambda_i\sum_{h=1}^{p}a_h f_h(\mathbf{s}_i)=\sum_{h=1}^{p}a_h f_h(\mathbf{s}_0)\Leftrightarrow\sum_{h=1}^{p}a_h\sum_{i=1}^{n}\lambda_i f_h(\mathbf{s}_i)=\sum_{h=1}^{p}a_h f_h(\mathbf{s}_0) \tag{41}$$

that verifies if and only if

$$\sum_{i=1}^{n}\lambda_i f_h(\mathbf{s}_i)=f_h(\mathbf{s}_0)\,,\forall h=1,\ldots,p \tag{42}$$

being this last expressions the p unbiasedness equations equivalents to the unbiasedness condition $\sum_{i=1}^{n}\lambda_i=1$ relative to the ordinary kriging.

The variance of the estimation error, the unbiasedness conditions included, can be expressed as follows

$$V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]=E\left[\left(\sum_{i=1}^{n}\lambda_i X(\mathbf{s}_i)-X(\mathbf{s}_0)\right)^2\right]=$$

$$=E\left\{\left[\left(\underbrace{\sum_{i=1}^{n}\lambda_i\mu(\mathbf{s}_i)-\mu(\mathbf{s}_0)}_{0}\right)+\left(\sum_{i=1}^{n}\lambda_i e(\mathbf{s}_i)-e(\mathbf{s}_0)\right)\right]^2\right\}=$$

$$=E\left[\left(\sum_{i=1}^{n}\lambda_i e(\mathbf{s}_i)-e(\mathbf{s}_0)\right)^2\right] \tag{43}$$

Knowing that usually $f_1(\mathbf{s})=1$, it implies that $\sum_{i=1}^{n}\lambda_i=1$, and so

[8] It is considered intrinsic because is constant in mean (in fact is null) but not in variance.

$$E\left[\left(\sum_{i=1}^{n}\lambda_i e(\mathbf{s}_i)-e(\mathbf{s}_0)\right)^2\right]=E\left[\left(\sum_{i=1}^{n}\lambda_i (e(\mathbf{s}_i)-e(\mathbf{s}_0))\right)^2\right]=$$
$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j E\big[(e(\mathbf{s}_i)-e(\mathbf{s}_0))(e(\mathbf{s}_j)-e(\mathbf{s}_0))\big] \tag{44}$$

The expression (44) is similar to the one that appeared in the ordinary kriging procedure, but using $e(\mathbf{s})$ instead of the variables $X(\mathbf{s})$.Therefore, following in the same way, the estimation variance can be obtained in terms of the variograms of the residual random function:

$$V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]=2\sum_{i=1}^{n}\lambda_i\gamma_e(\mathbf{s}_i-\mathbf{s}_0)-\sum_{i=1}^{n}\sum_{j=1}^{n}\lambda_i\lambda_j\gamma_e(\mathbf{s}_i-\mathbf{s}_j) \tag{45}$$

To minimize this expression under the unbiasedness conditions it is necessary the construction of the Lagrange function

$$\varphi(\lambda_i,\alpha_k)=\frac{1}{2}V[X^*(\mathbf{s}_0)-X(\mathbf{s}_0)]-\sum_{k=1}^{p}\alpha_k\left[\sum_i\lambda_i f_k(\mathbf{s}_i)-f_k(\mathbf{s}_0)\right] \tag{46}$$

and, setting the partial derivatives with respect to the weights λ_i and the Lagrange multiplier α_k, it is obtained the universal kriging system[9]:

$$\begin{cases}\displaystyle\sum_{j=1}^{N(\mathbf{s}_0)}\lambda_j\gamma_e(\mathbf{s}_i-\mathbf{s}_j)+\sum_{k=1}^{p}\alpha_k f_k(\mathbf{s}_i)=\gamma_e(\mathbf{s}_i-\mathbf{s}_0),\forall i=1,\dots,N(\mathbf{s}_0)\\ \displaystyle\sum_{i=1}^{N(\mathbf{s}_0)}\lambda_i f_k(\mathbf{s}_i)=f_k(\mathbf{s}_0),\forall k=1,\dots,p\end{cases} \tag{47}$$

This system can be expressed in matrix notations as $\mathbf{\Gamma}\cdot\mathbf{\Omega}=\mathbf{\Gamma}_0$, where

$$\mathbf{\Gamma}=\begin{pmatrix}
\gamma_e(\mathbf{0}) & \gamma_e(\mathbf{s}_1-\mathbf{s}_2) & \cdots & \gamma_e(\mathbf{s}_1-\mathbf{s}_{N(\mathbf{s}_0)}) & 1 & f_2(\mathbf{s}_1) & \cdots & f_p(\mathbf{s}_1)\\
\gamma_e(\mathbf{s}_2-\mathbf{s}_1) & \gamma_e(\mathbf{0}) & \cdots & \gamma_e(\mathbf{s}_2-\mathbf{s}_{N(\mathbf{s}_0)}) & 1 & f_2(\mathbf{s}_2) & \cdots & f_p(\mathbf{s}_2)\\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots\\
\gamma_e(\mathbf{s}_{N(\mathbf{s}_0)}-\mathbf{s}_1) & \gamma_e(\mathbf{s}_{N(\mathbf{s}_0)}-\mathbf{s}_2) & \cdots & \gamma_e(\mathbf{0}) & 1 & f_2(\mathbf{s}_{N(\mathbf{s}_0)}) & \cdots & f_p(\mathbf{s}_{N(\mathbf{s}_0)})\\
1 & 1 & \cdots & 1 & 0 & 0 & \cdots & 0\\
f_2(\mathbf{s}_1) & f_2(\mathbf{s}_2) & \cdots & f_2(\mathbf{s}_{N(\mathbf{s}_0)}) & 0 & 0 & \cdots & 0\\
\vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & 0\\
f_p(\mathbf{s}_1) & f_p(\mathbf{s}_2) & \cdots & f_p(\mathbf{s}_{N(\mathbf{s}_0)}) & 0 & 0 & \cdots & 0
\end{pmatrix}$$

[9] Note that in this case $i=1,\dots,N(\mathbf{s}_0)$ instead of $i=1,\dots,n$ because the drift expression is only locally valid. In consequence, $N(\mathbf{s}_0)$ is the number of observations in the neighbourhood of $\mathbf{s}_0$ ($N(\mathbf{s}_0)\le n$).

$$\mathbf{\Omega} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_{N(\mathbf{s}_0)} \\ \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix} \quad \text{and} \quad \mathbf{\Gamma}_0 = \begin{pmatrix} \gamma_e(\mathbf{s}_1 - \mathbf{s}_0) \\ \gamma_e(\mathbf{s}_2 - \mathbf{s}_0) \\ \vdots \\ \gamma_e(\mathbf{s}_{N(\mathbf{s}_0)} - \mathbf{s}_0) \\ 1 \\ f_2(\mathbf{s}_0) \\ \vdots \\ f_p(\mathbf{s}_0) \end{pmatrix} \tag{48}$$

being its solution $\mathbf{\Omega} = \mathbf{\Gamma}^{-1} \cdot \mathbf{\Gamma}_0$.

The minimum estimation variance or "kriging variance" can then be written as

$$V[X^*(\mathbf{s}_0) - X(\mathbf{s}_0)] = \sum_{i=1}^{N(s_0)} \lambda_i \gamma_e(\mathbf{s}_i - \mathbf{s}_0) + \sum_{k=1}^{p} \alpha_k f_k(\mathbf{s}_0) = \mathbf{\Omega}^t \cdot \mathbf{\Gamma}_0^* \tag{49}$$

In general, this model needs additional specifications because the variogram of the residual random function can only be inferred in exceptional situations. In the geostatistical literature some solutions to this problem are suggested: Suppose that $\gamma_X \approx \gamma_e$ in a small neighbourhood (Samper and Carrera, 1996); estimate the drift through the assumption of an initial variogram model and use it to estimate the residuals by kriging (Davis, 1986); estimate the drift through the median or mean polish method and estimate the residuals by ordinary kriging (Cressie, 1993).

We do not want to finish this methodological section without stating that some authors as Christakos (2000) consider the kriging procedure as a particular case of the BME[10] analysis. When the general knowledge is limited to the mean and the (centered) covariance and the specifying knowledge includes only hard data (i.e., accurate measurements obtained from real-time observation devices, computational algorithms or simulation processes) the BMEmode estimates coincides with the simple kriging estimate; when the general knowledge is limited to the variogram and the specificatory knowledge includes only hard data, the BMEmode estimates coincides with the ordinary kriging estimate.

3.2. Cokriging methods as an alternative: Multivariate approach

Cokriging is a multivariate interpolation technique that allows one to better estimate map values if the distribution of a secondary process sampled more intensely than the primary process is known. If the primary process is difficult or expensive to measure, then cokriging can greatly improve interpolation

[10] *Bayesian Maximum Entropy.*

estimates without having to more intensely sample the primary process. According to Wackernagel (2003, pp.158), cokriging is a natural extension of kriging when a multivariate variogram (or covariance model) and multivariate data are available. Cokriging allows estimating a variable of interest —the so-called main process— at a specific location from data about itself and about auxiliary processes in the neighbourhood. This is an important advantage when the size of the sample taken from the variable of interest is small and other auxiliary variables, correlated with it and more sampled, exist. This multivariate method allows improving the univariate estimates using the spatial correlation among the processes involved in the study. As an example and related to the real estate market, estimating prices of premises at any place of a particular area is not an easy task because the information available regarding the price of premises (not as comprehensive as that of the price of houses) usually is not enough to provide good estimates. Furthermore, due to the fact that prices of the real estate properties are spatially correlated, methods which are able to incorporate the role of space into conventional estimates are needed. This two facts –little available information and spatial correlation- lead us to propose cokriging as a methodology for the estimation of premises prices (main process) when sample sizes are small, unfortunately the real situation, being helped by the house prices (auxiliary process) in the same area, as it will be shown in section 4.3.

Trying to classify the different cokriging methods developed in the literature, it is necessary to observe that the stochastic processes considered (the main process and the auxiliary ones) can be sampled in the same geographical locations or not and that the sampled sizes can be identical or not. In this sense, the data set can be isotopic —if data is available for each variable at all sampling points—, entirely heterotopic —if the variables have been measured on different sets of sample points and have no sample locations in common— and partially heterotopic — if the data set do not cover all variables at all sample locations.

These facts lead us to make the following punctual cokriging classification:
(i) Cokriging with partial heterotopy (Ordinary cokriging (OCK): when the processes are second order stationary or intrinsically stationary and Universal cokriging (UCK): when the processes are non stationary); (ii) Cokriging with isotopy (second order stationary or intrinsically stationary and non stationary); and (iii) Cokriging with total heterotopy (second order stationary or intrinsically stationary and non stationary). From now on it is going to be reviewed the partial heterotopy cases because the isotopic situations are particular cases of the partial heterotopy ones, when the observed values of the regionalized variables are sited in the same locations. In this situation, the notation is clearly simpler. Respect to the entirely heterotopy cases, when the sample sets are disjoint, the experimental cross-variograms cannot be computed. So, the cokriging system has to be expressed in direct and cross-covariance terms and it can only be made in the second order stationary situation (see Wackernagel, 2003, pp. 158).

3.2.1. Ordinary cokriging

This section reviews the theoretical considerations for Ordinary Cokriging (OCK), in the joint intrinsic hypothesis, in the partial heterotopy case (some variables share some sample locations). Consider $\mathbf{X} = (X_1, X_2, ..., X_m)^t$ a vector of intrinsic random functions, that is,

$$E[X_i(\mathbf{s}+\mathbf{h}) - X_i(\mathbf{s})] = 0, \forall i = 1,\dots,m \quad \text{and} \tag{50}$$

$$\operatorname{cov}[X_i(\mathbf{s}+\mathbf{h}) - X_i(\mathbf{s}), X_j(\mathbf{s}+\mathbf{h}) - X_j(\mathbf{s})] = 2\gamma_{ij}(\mathbf{h}), \ \forall i, j = 1,\dots,m \tag{51}$$

being $\mathbf{s}$ the locations where information is taken and $\mathbf{h}$ a vector linking any two points $\mathbf{s}$ and $\mathbf{s}+\mathbf{h}$ in the domain under study. Each variable is defined on a set of samples with sizes $n_1, n_2, ..., n_m$ respectively.

Under these conditions, the OCK estimator of a particular variable X_i at the point $\mathbf{s}_0$ is a weighted linear combination of the data values from the variables X_j ($j = 1,...,m$) located at sampled points in the neighbourhood of $\mathbf{s}_0$.

$$X_i^*(\mathbf{s}_0) = \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j X_j(\mathbf{s}_\alpha^j) \tag{52}$$

with $\{\mathbf{s}_\alpha^j, \alpha = 1,\dots,n_j\}$ being the set of locations where the variable X_j, $j = 1,\dots,m$, has been sampled. Expanding the first summation as follows,

$$X_i^*(\mathbf{s}_0) = \sum_{\alpha=1}^{n_1} \lambda_\alpha^1 X_1(\mathbf{s}_\alpha^1) + \sum_{\alpha=1}^{n_2} \lambda_\alpha^2 X_2(\mathbf{s}_\alpha^2) + \dots + \sum_{\alpha=1}^{n_i} \lambda_\alpha^i X_i(\mathbf{s}_\alpha^i) + \dots + \sum_{\alpha=1}^{n_m} \lambda_\alpha^m X_m(\mathbf{s}_\alpha^m) \tag{53}$$

the estimator expression can be better appreciated as a linear combination of the random variables corresponding to each of the m processes.

The weights λ_α^j, $\alpha = 1,\dots,n_j$, $j = 1,\dots,m$, are calculated to ensure that the estimator is optimal, in the sense that it is unbiased and minimum error-variance. The unbiasedness is warranted by choosing weights which sum up to one for the variable of interest and which have zero sums for auxiliary variables (see, for example, Wackernagel, 2003, pp. 160)

$$E\left(X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right) = E\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j X_j(\mathbf{s}_\alpha^j) - X_i(\mathbf{s}_0)\right) = \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j E\left(X_j(\mathbf{s}_\alpha^j)\right) - E\left(X_i(\mathbf{s}_0)\right) =$$

$$= \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j \mu_j - \mu_i = 0 \quad \Leftrightarrow \quad \sum_{\alpha=1}^{n_i} \lambda_\alpha^i = 1 \ \text{and} \sum_{\alpha=1}^{n_j} \lambda_\alpha^j = 0, \forall j \neq i \ j = 1,\dots,m \tag{54}$$

Due to the impossibility of expressing the OCK system in terms of covariances it is suitable to express the variance of the estimation error in terms of direct and cross variograms. Then

$$V[X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)] = E\left[\left(X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right)^2\right] - \underbrace{\left[E\left(X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right)\right]^2}_{0} =$$

$$= E\left[\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j X_j(\mathbf{s}_\alpha^j) - X_i(\mathbf{s}_0)\right)^2\right] \tag{55}$$

and taking into account the unbiasedness constraints (46), this last expression can be written as

$$V\left[X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right] = E\left[\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\left(X_j(\mathbf{s}_\alpha^j) - X_i(\mathbf{s}_0)\right)\right)^2\right] =$$

$$= E\left[\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\left(X_j(\mathbf{s}_\alpha^j) - X_i(\mathbf{s}_0)\right)\right)\left(\sum_{k=1}^{m}\sum_{\beta=1}^{n_k}\lambda_\beta^k\left(X_k(\mathbf{s}_\beta^k) - X_i(\mathbf{s}_0)\right)\right)\right] =$$

$$= \sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{\alpha=1}^{n_j}\sum_{\beta=1}^{n_k}\lambda_\alpha^j\lambda_\beta^k E\left[\left(X_j(\mathbf{s}_\alpha^j) - X_i(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\beta^k) - X_i(\mathbf{s}_0)\right)\right] \tag{56}$$

To express (56) in terms of direct and cross variograms we proceed as follows: In the analytical definition of the cross variogram for a fixed pair of variables X_j, X_k

$$\gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) = \frac{1}{2}E\left[\left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_\beta^k)\right)\left(X_k(\mathbf{s}_\alpha^j) - X_k(\mathbf{s}_\beta^k)\right)\right] \tag{57}$$

we add and subtract $X_j(\mathbf{s}_0)$ in the first bracket and $X_k(\mathbf{s}_0)$ in the second one

$$\gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) =$$
$$= \frac{1}{2}E\left[\left(\left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_0)\right) - \left(X_j(\mathbf{s}_\beta^k) - X_j(\mathbf{s}_0)\right)\right)\left(\left(X_k(\mathbf{s}_\alpha^j) - X_k(\mathbf{s}_0)\right) - \left(X_k(\mathbf{s}_\beta^k) - X_k(\mathbf{s}_0)\right)\right)\right] =$$
$$= \frac{1}{2}E\left[\begin{matrix}\left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\alpha^j) - X_k(\mathbf{s}_0)\right) - \left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\beta^k) - X_k(\mathbf{s}_0)\right) - \\ - \left(X_j(\mathbf{s}_\beta^k) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\alpha^j) - X_k(\mathbf{s}_0)\right) + \left(X_j(\mathbf{s}_\beta^k) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\beta^k) - X_k(\mathbf{s}_0)\right)\end{matrix}\right] \tag{58}$$

obtaining

$$\begin{aligned}\gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) &= \\ = \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_0) &- \frac{1}{2}E\left[\left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\beta^k) - X_k(\mathbf{s}_0)\right)\right] - \\ - \frac{1}{2}E\left[\left(X_j(\mathbf{s}_\beta^k) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\alpha^j) - X_k(\mathbf{s}_0)\right)\right] &+ \gamma_{jk}(\mathbf{s}_\beta^k - \mathbf{s}_0)\end{aligned} \tag{59}$$

At this point it is introduced the additional assumption related to the symmetry of the covariances, that is $C_{ij}(\mathbf{h}) = C_{ji}(\mathbf{h})$. It implies that

$$\begin{aligned}&E\left[\left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_0)\right)\cdot\left(X_k(\mathbf{s}_\beta^k) - X_k(\mathbf{s}_0)\right)\right] = \\ &= E\left[\left(X_j(\mathbf{s}_\beta^k) - X_j(\mathbf{s}_0)\right)\cdot\left(X_k(\mathbf{s}_\alpha^j) - X_k(\mathbf{s}_0)\right)\right]\end{aligned} \tag{60}$$

The inclusion of this assumption leads to the expression

$$\gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) =$$
$$= \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_0) + \gamma_{jk}(\mathbf{s}_\beta^k - \mathbf{s}_0) - E\left[\left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\beta^k) - X_k(\mathbf{s}_0)\right)\right] \quad (61)$$

having that

$$E\left[\left(X_j(\mathbf{s}_\alpha^j) - X_j(\mathbf{s}_0)\right)\left(X_k(\mathbf{s}_\beta^k) - X_k(\mathbf{s}_0)\right)\right] =$$
$$= \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_0) + \gamma_{jk}(\mathbf{s}_\beta^k - \mathbf{s}_0) - \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) \quad (62)$$

Finally, substituting (62) in (56), the variance of the estimation error can be rewritten as

$$V\left[X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right] = \sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{\alpha=1}^{n_j}\sum_{\beta=1}^{n_k} \lambda_\alpha^j \lambda_\beta^k \left(\gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_0) + \gamma_{jk}(\mathbf{s}_\beta^k - \mathbf{s}_0) - \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k)\right) =$$
$$= 2\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j \gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{\alpha=1}^{n_j}\sum_{\beta=1}^{n_k} \lambda_\alpha^j \lambda_\beta^k \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) \quad (63)$$

Remembering that the criterion to define the best weights λ_α^j will be to reduce as much as possible the variance of the estimation error, expression (63) has to be minimized with the constraints (54) to obtain such weights. Following the method of Lagrange we define the objective function

$$\varphi(\lambda_\alpha^j, \omega_j) = V\left[X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right] - 2\omega_i\left(\sum_{\alpha=1}^{n_i} \lambda_i - 1\right) - 2\sum_{\substack{j=1 \\ j\neq i}}^{m} \omega_j \sum_{\alpha=1}^{n_j} \lambda_\alpha^j,$$
$$j = 1,\ldots,m;\ \alpha = 1,\ldots,n_j \quad (64)$$

that contains *m* Lagrange multipliers ω_j, $j = 1,\ldots,m$, the number of constraints that guarantees the unbiasedness of the OCK estimator.

Setting the first order partial derivatives to zero the OCK system (65) is obtained.

$$\begin{cases} \sum_{k=1}^{m}\sum_{\beta=1}^{n_k} \lambda_\beta^k \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) + \omega_j = \gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) & \forall j = 1,\ldots,m;\quad \forall \alpha = 1,\ldots,n_j \\ \sum_{\alpha=1}^{n_j} \lambda_\alpha^j = \delta_{ij} = \begin{cases} 1 & si\ \ i = j \\ 0 & si\ \ i \neq j \end{cases} \end{cases} \quad (65)$$

The solution of this system provides the optimal weights λ_α^j for the estimation of the particular variable X_i at the point $\mathbf{s}_0$. Finally, rewriting the first equation of (65) as

$$\sum_{k=1}^{m}\sum_{\beta=1}^{n_k} \lambda_\beta^k \gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) = \gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \omega_j \qquad \forall j = 1,\ldots,m;\quad \forall \alpha = 1,\ldots,n_j \quad (66)$$

and including (66) into the variance expression (63), the minimal cokriged estimation variance is obtained as

$$V\left[X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right] =$$

$$= 2\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \sum_{j=1}^{m}\sum_{k=1}^{m}\sum_{\alpha=1}^{n_j}\sum_{\beta=1}^{n_k}\lambda_\alpha^j\lambda_\beta^k\gamma_{jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) =$$

$$= 2\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\omega_j\right) =$$

$$= 2\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\gamma_{ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \omega_i\right) =$$

$$= \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_\alpha^j\gamma_{ji}(\mathbf{s}_\alpha^i - \mathbf{s}_0) + \omega_i \tag{67}$$

Cross and direct variograms, γ_{ij} and γ_{ii} respectively, are obtained in two steps: First, point estimates of the variograms are obtained using the classical variogram estimator based on the method-of-moments $\gamma_{ij}^*(\mathbf{h}) = \frac{1}{2N(\mathbf{h})}\sum_{N(\mathbf{h})}\left(X_i(\mathbf{s}_\alpha + \mathbf{h}) - X_i(\mathbf{s}_\alpha)\right)\cdot\left(X_j(\mathbf{s}_\alpha + \mathbf{h}) - X_j(\mathbf{s}_\alpha)\right)$ (it is supposed constant-mean, see Lark and Papritz, 2003). The second step is to fit a theoretical variogram function to the sequence of average dissimilarities, according to the linear model of corregionalization (see, for example, Journel and Huijbregts, 1978, pp. 171-175; Goovaerts, 1997, pp. 108-115 and Wackernagel, 2003, pp. 175-176) because it is the usual strategy to ensure a positive definite model.

3.2.2. Universal cokriging

Consider that we deal with a non-stationary vector of stochastic processes $\mathbf{X} = (X_1, X_2, ..., X_m)^t$. It means that at least one process presents a drift term: $E[X_i(\mathbf{s})] = \mu_i(\mathbf{s})$ for any location $\mathbf{s} \in S \subset \mathbf{R}^2$ and for some $i = 1,...,m$. Under this assumption, as in the univariate case, each process X_i can be broken up in a deterministic component and a stochastic one as follows:

$$X_i(\mathbf{s}) = \mu_i(\mathbf{s}) + e_i(\mathbf{s}), \quad \forall i = 1,...,m \tag{68}$$

Now, suppose that $\boldsymbol{\mu}(\mathbf{s}) = (\mu_1(\mathbf{s}),...,\mu_m(\mathbf{s}))^t$ is the vector of the drifts corresponding to the m random fuctions[11]. As in the univariate case, suppose that each of them can be expressed as a linear combination of p any known functions $f_1^j(\mathbf{s}),...,f_p^j(\mathbf{s}), \quad j = 1,\dots,m$.

[11] For those processes with no trend, $\mu_i(\mathbf{s}) = \mu_i \ \forall \mathbf{s}$, remembering that a vector is considered as non stationary if at least one process has trend.

$$\mu_j(\mathbf{s}) = \sum_{h=1}^{p} a_h^j f_h^j(\mathbf{s}) \tag{69}$$

being a_h^j constant coefficients. Imposing to the estimation error the null expected value,

$$E\left(X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right) = E\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j X_j(\mathbf{s}_\alpha^j) - X_i(\mathbf{s}_0)\right) = \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j E\left(X_j(\mathbf{s}_\alpha^j)\right) - E\left(X_i(\mathbf{s}_0)\right) =$$

$$= \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j \mu_j(\mathbf{s}_\alpha^j) - \mu_i(\mathbf{s}_0) = 0 \tag{70}$$

we obtain

$$\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_\alpha^j \left[\sum_{h=1}^{p} a_h^j f_h^j(\mathbf{s}_\alpha^j)\right] = \sum_{h=1}^{p} a_h^i f_h^i(\mathbf{s}_0) \tag{71}$$

that verifies if and only if

$$\sum_{\alpha=1}^{n_j} \lambda_{ij}^\alpha f_h^j(\mathbf{s}_\alpha^j) = f_h^i(\mathbf{s}_0)\delta_{ij} = \begin{cases} f_h^i(\mathbf{s}_0) & \text{si} \quad i = j \\ 0 & \text{si} \quad i \neq j \end{cases} \qquad \forall h = 1,\ldots,p \tag{72}$$

Following the classic procedure to find the best linear estimator, it is necessary to minimize the variance of the estimation error, which can be written as follows,

$$V\left[X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right] = E\left[\left(X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right)^2\right] - \underbrace{\left[E\left(X_i^*(\mathbf{s}_0) - X_i(\mathbf{s}_0)\right)\right]^2}_{0} =$$

$$= E\left[\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_{ij}^\alpha X_j(\mathbf{s}_\alpha^j) - X_i(\mathbf{s}_0)\right)^2\right] = E\left[\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_{ij}^\alpha \left(\mu_j(\mathbf{s}_\alpha^j) + e_j(\mathbf{s}_\alpha^j)\right) - \left(\mu_i(\mathbf{s}_0) + e_i(\mathbf{s}_0)\right)\right)^2\right] =$$

$$= E\left[\left(\underbrace{\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_{ij}^\alpha \mu_j(\mathbf{s}_\alpha^j) - \mu_i(\mathbf{s}_0)}_{0} + \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} e_j(\mathbf{s}_\alpha^j) - e_i(\mathbf{s}_0)\right)^2\right] \tag{73}$$

If, as usually, we suppose that $f_1^i = 1$, then the first unbiasedness condition can be written as

$$\sum_{\alpha=1}^{n_i} \lambda_{ii}^\alpha = 1 \text{ and } \sum_{\alpha=1}^{n_j} \lambda_{ij}^\alpha = 0, \quad \forall j \neq i \tag{74}$$

Under this condition, the expression (73) can be rewritten as

$$E\left[\left(\underbrace{\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} \lambda_{ij}^\alpha \mu_j(\mathbf{s}_\alpha^j) - \mu_i(\mathbf{s}_0)}_{0} + \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j} e_j(\mathbf{s}_\alpha^j) - e_i(\mathbf{s}_0)\right)^2\right] =$$

$$= E\left[\left(\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\left(e_j(\mathbf{s}_\alpha^j) - e_i(\mathbf{s}_0)\right)\right)^2\right] =$$

$$= \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\sum_{k=1}^{m}\sum_{\beta=1}^{n_k}\lambda_{ij}^{\alpha}\lambda_{ik}^{\beta}E\left(e_j(\mathbf{s}_\alpha^j) - e_i(\mathbf{s}_0)\right)\left(e_k(\mathbf{s}_\beta^k) - e_i(\mathbf{s}_0)\right) \tag{75}$$

This last expression (75) can be expressed in cross-variograms error terms as in (61), only changing the process X by the errors e.

$$\gamma_{e\,jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) =$$
$$\gamma_{e\,jk}(\mathbf{s}_\alpha^j - \mathbf{s}_o) + \gamma_{e\,jk}(\mathbf{s}_\beta^k - \mathbf{s}_0) - E\left[\left(e_j(\mathbf{s}_\alpha^j) - e_j(\mathbf{s}_0)\right)\left(e_k(\mathbf{s}_\alpha^k) - e_k(\mathbf{s}_0)\right)\right] \tag{76}$$

Working out the value of $E\left[\left(e_j(\mathbf{s}_\alpha^j) - e_j(\mathbf{s}_0)\right)\left(e_k(\mathbf{s}_\beta^k) - e_k(\mathbf{s}_0)\right)\right]$ and replacing it in (75) we finally obtain the variance of the estimation error,

$$V\left(X_i^*(\mathbf{s}_0) - X_i\,(\mathbf{s}_0)\right) = 2\sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\lambda_{ij}^{\alpha}\gamma_{e\,ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0) - \sum_{j=1}^{m}\sum_{\alpha=1}^{n_j}\sum_{k=1}^{m}\sum_{\beta=1}^{n_k}\lambda_{ij}^{\alpha}\lambda_{ik}^{\beta}\gamma_{e\,jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) \tag{77}$$

To minimize this variance it is necessary to construct the following Lagrange function:

$$\begin{aligned}\varphi(\lambda_{ij}^{\alpha}, \omega_{hj}) &= V\left(X_i^*(\mathbf{s}_0) - X_i\,(\mathbf{s}_0)\right) - 2\sum_{h=1}^{p}\omega_{hi}\left[\sum_{\alpha=1}^{n_i}\lambda_{ii}^{\alpha}f_h^i(\mathbf{s}_\alpha^i) - f_h^i(\mathbf{s}_0)\right] - \\ &- 2\sum_{h=1}^{p}\sum_{j=1}^{m}\omega_{hj}\sum_{\alpha=1}^{n_j}\lambda_{ij}^{\alpha}f_h^j(\mathbf{s}_\alpha^j), \quad \forall j = 1,\ldots,m, \quad \forall\alpha = 1,\ldots,n_j, \quad \forall h = 1,\ldots,p\end{aligned} \tag{78}$$

being ω_{hj} the mxp Lagrange multipliers. Setting the first order partial derivatives to zero the universal cokriging (UCK) system (79) is obtained:

$$\sum_{k=1}^{m}\sum_{\beta=1}^{n_k}\lambda_{ki}^{\beta}\gamma_{e\,jk}(\mathbf{s}_\alpha^j - \mathbf{s}_\beta^k) + \sum_{h=1}^{p}\omega_{hj}f_h^j(\mathbf{s}_\alpha^j) = \gamma_{e\,ji}(\mathbf{s}_\alpha^j - \mathbf{s}_0), \quad \forall j = 1,\ldots,m, \quad \forall\alpha = 1,\ldots,n_j$$
$$\sum_{\alpha=1}^{n_j}\lambda_{ij}^{\alpha}f_h^j(\mathbf{s}_\alpha^j) = f_h^i(\mathbf{s}_0)\cdot\delta_{ij}$$
$$\tag{79}$$

As it can be seen in system (79), the variograms that appear correspond to the stochastic component of the initial break-up (68). Nevertheless, as in the univariate case, with unknown trends it is not possible to compute these variograms. This is why, in practice, one can use any of the methods mentioned in the univariate case —e.g. *OCK over the residuals*— in order to solve this problem.

A thorough presentation of the theory of multivariate geostatistics is available in Chilès and Delfiner (1999), Wackernagel (2003) and Webster and Oliver (2004), among others.

4. Application in the real estate market

This section includes three applications about economic phenomena involving spatial autocorrelation related with the real estate market. The first one gives an alternative to the official housing statistic in Spain proposing the kriging the mean to estimate mean housing prices. The second one estimates the house prices in any location of the city of Toledo (Spain) using ordinary kriging to map the price of houses in this historical area. Finally, the third application consists on estimating premises prices in Toledo, using ordinary cokriging because of the small sample size of the premises compared with the valuated houses.

4.1. Spatial correlation in housing price indexes

From an official point of view it is necessary to deal with statistical methods that provide a credible mean price of houses in geographical areas of interest (municipality, regions and the whole country), being the final aim to provide credible housing price indexes. As it has been pointed out in section 3.1.1., we want to stress at this point the importance that to consider the space or not has when estimating a parameter. In spite of this fact, in Spain, official authorities deal with the available data as if prices were independent and they continue using the arithmetic average of sampled prices as an estimator of the mean price of houses in an area. Obviously, in the case of spatial correlation the arithmetic mean is not the BLUE estimator. This is why in this section we suggest to compare several estimations of the mean house price in order to show the differences that could appear between the kriging the mean estimator and the official methodology. Trying to show how using one or other methodology the final results can vary, we have compared the results obtained in the estimation of housing prices in the last two years —2005 and 2006, quarterly data— in particular, estimating the mean price in Madrid (Spain), the Spanish city where the housing price per square meter used to be the highest over the country. We have computed our results for the non state-subsidized market because the official statistic does not provide the number of the valuated state-subsidized houses in each quarter. This is why we have not been able to aggregate, finally, the mean price to any type of house. We have tried to reply the data provided in the official web site because, obviously, we do not have the data at our disposal due to its secret character. Therefore, we have generated normal random sets of data with the mean matching up with the official one, as it is reported in Table 1.

Table 1: Official housing prices corresponding to the non state-subsidized market in Madrid (Spain) (€/m²).

Time	Price per square meter			Number of valuations		
	Less than 2 years	More than 2 years	Total	Less than 2 years	More than 2 years	Total
2005 Q1	3632.4	3289.2	3342.8	2185	11785	13973
2005 Q2	3741.0	3437.1	3504.7	3136	10973	14109
2005 Q3	3799.5	3483.1	3564.6	1755	5061	6816
2005 Q4	4014.4	3489.0	3603.1	2319	8359	10678
2006 Q1	3916.9	3576.1	3657.0	3675	11799	15474
2006 Q2	3990.2	3633.7	3699.7	3080	13558	16638
2006 Q3	3976.0	3643.1	3717.2	2172	7585	9757
2006 Q4	3914.6	3660.3	3728.4	3452	9430	12882

For each data set, one for each quarter, we also do not know the exact spatial location of the corresponding valuated house but this does not seem to be an impediment to continue the procedure because the official methodology does not use this spatial characteristic, so, its results would not change in any case, whatever the spatial location is. Therefore, the available information has been spatially disposed in order to obtain a geographical distribution according to the existence of a positive autocorrelated process.

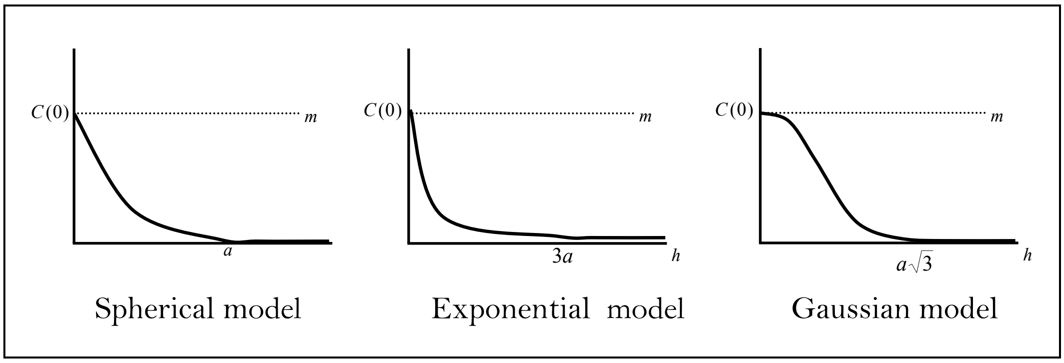
Three different covariance functions —which describes the correlation between the variables at any pair of locations **s** and **s+h**— have been selected to asses the estimating power of the investigated technique, in order to compare the different results that could appear, depending on the type of spatial structure that the stochastic process would follow. In this point, we assume that the simulated data are spatially disposed so that the structure of the spatial dependence between the random variables of the stochastic process is according to one of the theoretical models reported in Table 2 and represented in Figure 2.

Tabla 2: Theoretical covariance functions

Model	Omnidirectional Covariogram	Range (meters)
Spherical	$C(h) = m\left[1 - \frac{3}{2}\frac{h}{a} + \frac{1}{2}\left(\frac{h}{a}\right)^3\right]$	1500
Exponential	$C(h) = m\, e^{-\frac{h}{a}}$	1500
Gaussian	$C(h) = m\, e^{-\frac{h^2}{a^2}}$	1500

Note: m : sill; a : range; h : distance.

Fig. 2: Graphic representation of theoretical covariance functions



Assuming that the range remain constant during the lag period considered (2005-2006) and using the above mentioned covariance functions, 48 kriging the mean equation systems have been solved in order to obtain the estimation of the housing price per square meter, for the eight quarters considered, for new and second-hand houses in Madrid. The estimation of the optimal weights λ_i that have to be included in the kriging the mean estimator to obtain the housing price estimation in presence of spatial dependence have been computed using R-language and Isatis software developed by Geovariances and L'Ecole des Mines de Paris. The kriging the mean estimations are reported in Table 3.

Table 3: Kriging the mean housing prices corresponding to the non state-subsidized market in Madrid (Spain) (€/m²).

	Price per square meter								
Model	Spherical			Exponential			Gaussian		
Time	Less than 2 years	More than 2 years	Total	Less than 2 years	More than 2 years	Total	Less than 2 years	More than 2 years	Total
2005 Q1	3543.3	3045.3	3122.5	3506.1	3003.4	3081.4	3645.4	3349.1	3394.7
2005 Q2	3665.9	3159.7	3272.2	3670.9	3170.4	3281.6	3751.2	3489.6	3547.7
2005 Q3	3708.4	3190.2	3323.6	3688.7	3145.0	3285.0	3765.2	3395.0	3490.3
2005 Q4	3987.5	3204.4	3374.5	4012.1	3166.9	3350.5	4008.8	3421.1	3548.7
2006 Q1	3887.4	3322.8	3456.9	3954.4	3307.7	3461.3	3985.7	3688.9	3759.4
2006 Q2	3923.8	3456.1	3542.7	3987.2	3374.0	3487.5	3889.0	3556.3	3617.9
2006 Q3	3957.0	3477.0	3583.9	4123.3	3687.2	3784.3	4123.9	3782.6	3858.6
2006 Q4	3891.3	3489.6	3597.2	3990.9	3556.4	3672.8	3925.6	3677.5	3744.0

As it can be appreciated in the study that has been carried out, the official methodology shows different housing prices that those computed through the kriging the mean estimator, which considers the spatial dependence that the housing prices presents. Moreover, the results obtained through the three different covariance functions considered in the analysis are also quite different, which talks about the importance of considering the structure of the spatial correlation that the house prices presents in each geographic area and about the need to being able to estimate it well. The great differences can be clearly seen in Figure 3a to 3c where the housing price series are depicted. The figures compare the housing price series for the new houses (3a), the second-hand houses (3b) and the total (3c).

Figure 3a: New and non state-subsidized housing price series in Madrid (Spain) (€/m²).

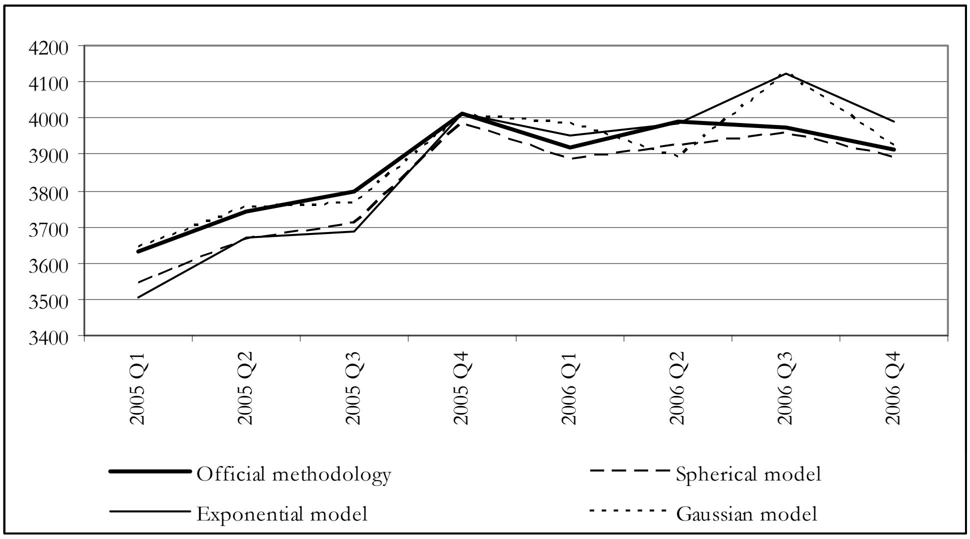


Figure 3b: Second-hand and non state-subsidized housing price series in Madrid (Spain) (€/m²).

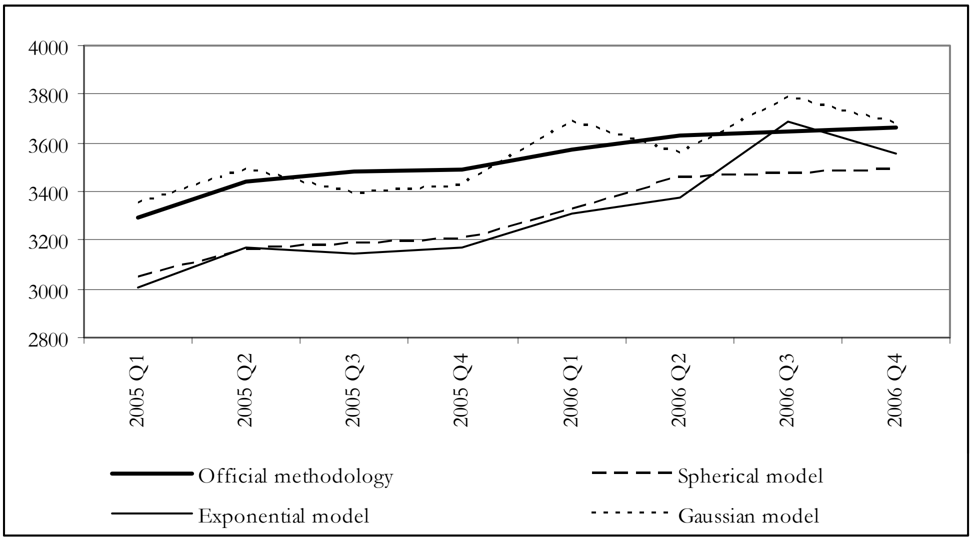
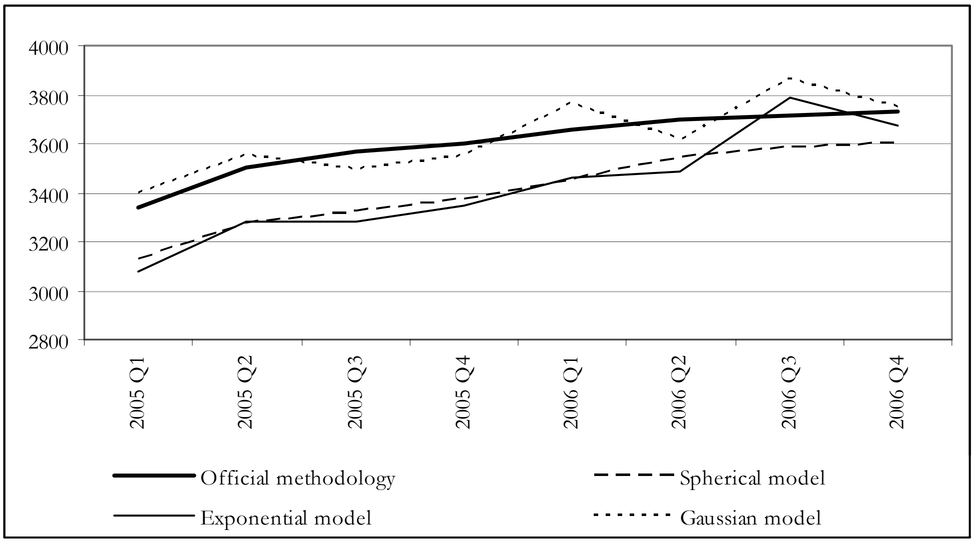


Figure 3c: Non state-subsidized housing price series in Madrid (Spain) (€/m²).



However, this is only an example, a simulation carried out through a particular layout of the data in the surface of the geographical area considered. The real important issue is that the presence of spatial correlation moves us further away from the results provided by the Housing Department and, as a consequence, it generates quite different housing price series.

Once we have obtained the kriging the mean estimations of the housing price for each quarter, we have proceeded to compute the housing price indexes when spatial correlation exists. To develop this task we have first computed the simple indexes for each quarter and for each type of house (new or second hand), being the base period 1995 Q1. Then, we have calculated the complex spatial indexes as a weighted average of the simple ones. The results are reported in Table 4.

Table 4: Spatial housing price indexes corresponding to the non state-subsidized market in Madrid (Spain) (€/m²).

	Spatial Price Indexes								
Model	Spherical			Exponential			Gaussian		
Time	Less than 2 years	More than 2 years	Total	Less than 2 years	More than 2 years	Total	Less than 2 years	More than 2 years	Total
2005 Q1	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0
2005 Q2	103.5	103.8	103.7	104.7	105.6	105.5	102.9	104.2	104.1
2005 Q3	104.7	104.8	104.7	105.2	104.7	104.8	103.3	101.4	101.6
2005 Q4	112.5	105.2	105.8	114.4	105.4	106.1	110.0	102.1	102.8
2006 Q1	109.7	109.1	109.2	112.8	110.1	110.4	109.3	110.1	110.1
2006 Q2	110.7	113.5	113.3	113.7	112.3	112.4	106.7	106.2	106.2
2006 Q3	111.7	114.2	114.0	117.6	122.8	122.4	113.1	112.9	113.0
2006 Q4	109.8	114.6	114.1	113.8	118.4	118.0	107.7	109.8	109.6

Together with these housing price indexes, which consider the spatial dependence that the stochastic process presents, we have also compared these results with the indexes deduced from the official methodology. Table 5 shows the cited comparison.

Table 5: Spatial and Official housing price indexes corresponding to the non state-subsidized market in Madrid (Spain) (€/m²).

	Spatial Price Indexes			**Official Indexes**
Model	Spherical	Exponential	Gaussian	
Time				
2005 Q1	100.0	100.0	100.0	100.0
2005 Q2	103.7	105.5	104.1	104.4
2005 Q3	104.7	104.8	101.6	105.8
2005 Q4	105.8	106.1	102.8	106.4
2006 Q1	109.2	110.4	110.1	108.6
2006 Q2	113.3	112.4	106.2	110.4
2006 Q3	114.0	122.4	113.0	110.7
2006 Q4	114.1	118.0	109.6	110.9

From Table 5 it can be deduced that while the official authorities provides a 10.9 growing rate at the end of 2006, the spherical model, for example, provides that the non-state subsidized houses have increased its prices in a 14.1 percent. Similar differences can be seen between the covariance functions considered and the official methodology as well as in the rest of the quarters.

Another very important issue we would like to stress is what would happen with the confidence intervals in the case that they were provided by the official authorities. Only trying to serve as an example, we have computed the confidence intervals of the mean house price, at a 90% confidence level, in three different cases, depending on the estimator used and on the hypothesis about the independence or spatial dependence of the variables assumed. The results are reported in Table 6.

Table 6: Confidence intervals for the housing prices of the non state-subsidized market in Madrid (Spain)

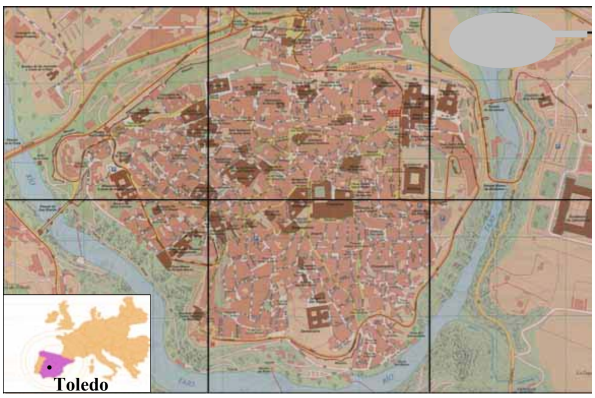
2006 Q4	**Confidence Intervals**			
	Model			
	Spherical	Exponential	Gaussian	Official price
House Price (€/m²)	3597.2	3672.8	3744.0	3728.4
Official mean/Independence	-	-	-	?
Official mean/Spatial dependence	-	-	-	[3311.5 ; 4145.3]
Kriging estimator/Spatial dependence	[3231.2 ; 3963.3]	[3293.3 ; 4052.3]	[3361.0 ; 4126.9]	-

In the first case, assuming what the official methodology says is true (the independence of the sample), we cannot construct the confidence interval because they do not provide the data and so its standard deviation. In the case that it exists, it would be narrower than the ones when considering the spatial correlation, but this is not the real situation. In the second case, considering that the spatial dependence exists but still estimating with the arithmetic mean, and assuming normal distribution, we conclude that the real mean price lies in the interval [3311.5 ; 4145.3] with a 90% confidence level. Finally, if we try to do the things well-done, considering that the spatial dependence exists and so, using the kriging the mean estimator, and assuming normal distribution, we conclude that the real mean price lies in the interval [3231.2 ; 3963.3], [3293.3 ; 4052.3] or [3361.0 ; 4126.9], dependeing on the covariance model considered, with a 90% confidence level. This last intervals are quite different that the second one. They are narrower —because the standard deviation of the kriging the mean estimator is the minimum— and settled in a different point —the kriging the mean estimation.

4.2. Mapping house prices using kriging in the historical area of Toledo city (Spain)

In the real estate case, the aim of estimating a value of the random variable in any location $\mathbf{s}_0$ out of the sample, $x^*(\mathbf{s}_0)$, seems to be more important than the estimation of the housing mean price due that a particular citizen would be interested on knowing the price of a particular house instead of the mean price over a region. To apply punctual kriging method to house prices it has been selected the historical area of Toledo city, in Spain. The reason for having chosen this emblematic area is that it represents an ideal site because it has neither geographic accident nor artificial barriers inside the walls that could benumb the spatial dependence structure. However, the study can be performed without *a priori* any difficulty, in any area, village or city. The study area and its position in Spain are depicted in Figure 4.

Fig. 4: Historical part of Toledo city



The database was collected during September 2003 and so the results will be referred to that date[12]. The database contains information about 121 houses and it has been obtained from the following Real Estate Agencies: Imagil Gestión Inmobiliaria, Zocopiso, Amian Inmobiliaria, Imperial Inmabel S.L., Agencia Inmobiliaria Gudiel, Inmobiliaria Castaño, Agrufinca, Acrópolis, Albatros, Teleinmobiliaria, Inmobiliaria Época, Inmobiliara Ábaco, Simar, Agencia Inmobibliaria Orgaz and Fondo Piso Toledo. The information we asked for was referred to the location of the house, condition (complete renovation needed, little renovation needed, in a enough good condition, new or completely renovated), surface (less than 60 m^2, from 60 to 120 m^2, from 120 to 200 m^2 and more than 200 m^2), garage (yes or not), elevator (yes or not), a particular use yard (yes or not), floor, views (yes or not) and, of course, sale price. However, whatever the house's location, only its condition, surface and if it had garage or not had significant effects on the price. This is the reason why only these three factors, besides the geographical location, were taken into account in the analysis. Percentages of houses with each level of the above mentioned factors are shown in Figures 5 to 8.

Fig. 5: House Condition

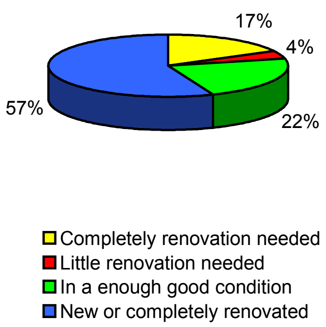
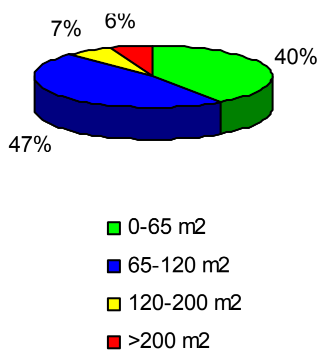


Fig. 6: House Surface



[12] Obviously, it woluld be desirable to carry out this analysis mensualy or quarterly. It would imply the incorpporation of a temporal dimension to the kriging methods (see Li and Revesz (2004), De Iaco, Myers and Posa (2002), among others). However, in Spain mensual or quarterly collection of data is not possible for researchers. Nevertheless, the objective we pursuit is to show that for a particular temporal reference not only is possible the estimation of the mean price of the houses in a particular area, but also, and more important, the estimation of prices in any location of the area.

Fig. 7: Garage

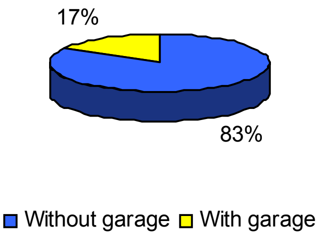
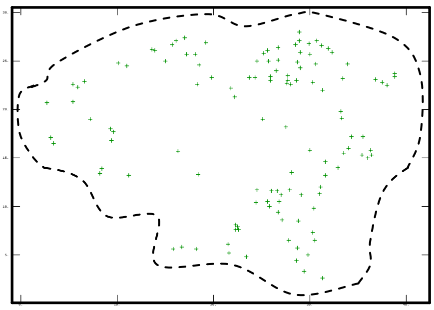


Fig. 8: Sample location base map



Making the houses "equivalent": In order to isolate the spatial component of the house prices (the analysis has been developed in price per square meter terms) the effects on prices of the levels of factors have been estimated by analysis of variance. Once these effects have been eliminated from the house prices, the new prices are supposed to correspond to equivalent houses (new or renovated, from 60 to 120 m^2, and without garage). The differences in the prices of the "equivalent" houses are supposed to be due to its geographical location. The structure of the spatial correlation among the prices of the "equivalent" houses has been estimated through a variogram that has been incorporated in the kriging equations to map the price of houses in the historical area of Toledo city or to kriging the mean. Evidently, in the first case, once the punctual estimates have been obtained, the effects of the factors "house condition", "surface" and "garage" are incorporated.

For the estimation of the effects on price of houses of the above mentioned factors we have proceeded as follows[13]:

1. Firstly, we have planned a 2-way multiplicative model (fixed effects) with unbalanced data, being the α and β factors "surface" and "garage" respectively (see Appendix 1b.)

$$y_{ijk} = \mu \cdot \alpha_i \cdot \beta_j \cdot e_{ijk} \tag{80}$$

where y_{ijk} are the houses prices.

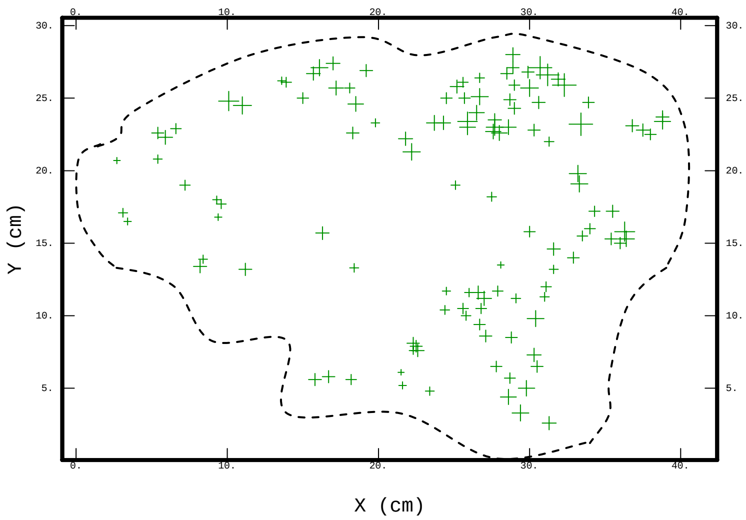
2. Once the differences between α_i and β_j effects are estimated it is created a data base of prices corresponding to “equivalent” houses with respect to these two factors. To estimate the "house condition" effect we have planned a 1-way additive model (fixed effects) with unbalanced data (see Appendix 1e.)

$$y_{ij}^{*} = \mu + \alpha_i + e_{ij} \tag{81}$$

where y_{ij}^{*} are prices of "equivalent" houses with respect to the "surface" and "garage" factors. The following Base Map provides a first view of these prices.

[13] Estimations were made with SPSS 11.5.

Fig. 9: Equivalent house prices Base Map



Once the prices of "equivalent" houses have been calculated we have tested randomness versus positive correlation by using the Moran's I statistic, with a contiguity matrix whose elements are the inverses of the distances among locations. The sample value of the statistic I was 0.125 being $E(I) = -0.0083$ and $V(I) = 0.00032$ in the random case[14]. Therefore, the standardized value of the Moran's statistic was 7.443 and this value leads to the rejection, at the 5% level of significance, of the null hypothesis of randomness in favour of positive autocorrelation.

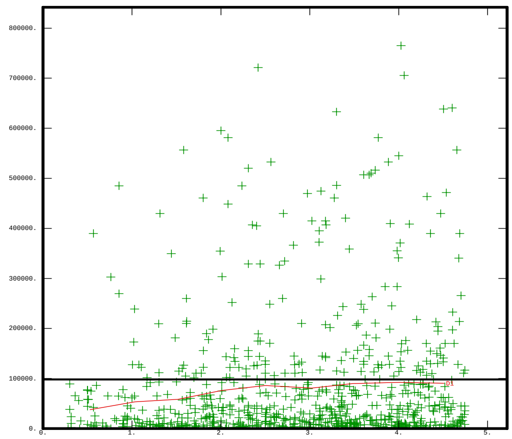
The structure of that positive autocorrelation was estimated by the following experimental variogram of prices ("equivalent" houses)

$$\gamma^*(\mathbf{h}) = \frac{1}{2N(\mathbf{h})}\sum_{\alpha=1}^{N(\mathbf{h})}\left(X(\mathbf{s}_\alpha + \mathbf{h}) - X(\mathbf{s}_\alpha)\right)^2 . \qquad (82)$$

From Figure 10 it can be appreciated that the stochastic process is supposed to be second-order stationary.

[14] R language for statistical computing was used to compute the sample value of the I statistic (see http://www.cran.r-project.org)

Fig. 10. Variogram cloud and experimental variogram of house prices.

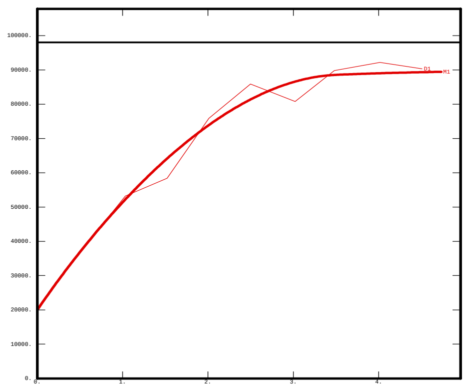


According to the linear model of regionalization, the following combination of theoretical models has been selected to fit the experimental variogram.

Table 7. Linear Model of Regionalization

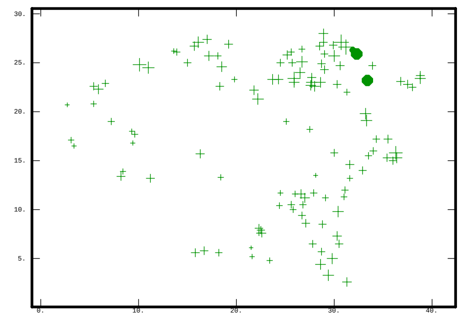
Model	Sill	Range
Nugget effect	20000	
Spheric	50000	120 meters
Exponential	20002	136 meters

Fig. 11. Fitted variogram $\gamma_{\text{viviendas}}$

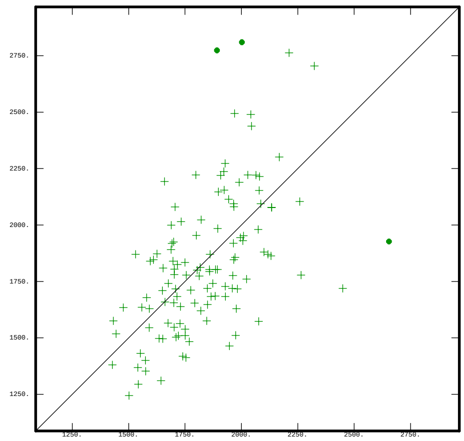


The fitted variogram has been checked using a cross-validation procedure, and 118 robust prices have been obtained out of 121 that form the original set, because in 118 cases the standardized kriging estimates belong to the interval $[-2'5; 2'5]$ (see Figure 12)

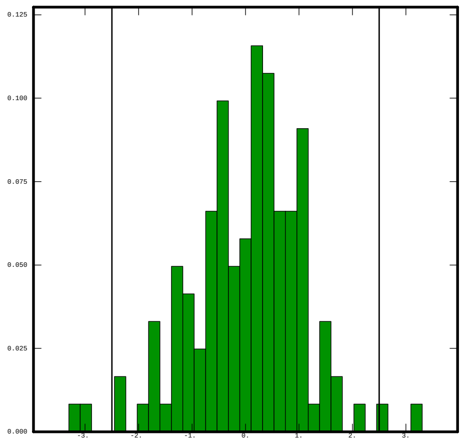
Fig. 12. (a) Base Map showing the non robust data, (b) scatter diagram of the true data versus the estimated value, (c) histogram of the standardized errors and (d) scatter diagram of the standardized estimations errors versus the estimated values.



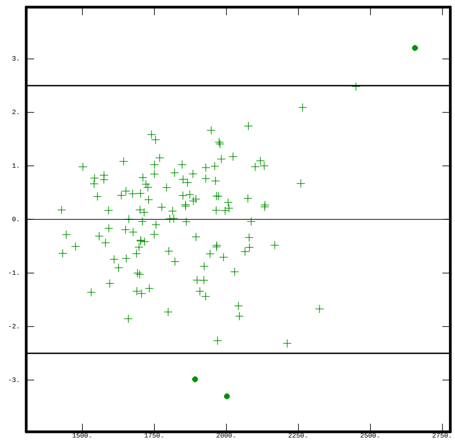
(*a*)



(*b*)



(*c*)



(*d*)

Once we have available a variogram that indicates the structure of the spatial dependence among the prices of houses in the historical area of Toledo city, the aim can be carried out: The estimation of particular house prices in any point of such an emblematic zone —taking into account the mentioned spatial correlation— obviously more useful for individual citizens than the mean price.

Once the combination of theorical variograms that better fits the structure of the spatial dependence in the house prices has been determined, it is now time to estimate the house price in all and each one of the locations sited inside the walls of the city of Toledo. As that variogram stabilizes around the estimated variance of the process, the random function relative to the price of "equivalent" houses can be considered second-order or intrinsically stationary. Therefore, the procedure to map the estimations will be the ordinary kriging. In order to carry out the kriging estimations we have drawn a regular grid of 36 meter side over the map of the historic site of Toledo city, having performed the estimations in the nodes of the grid. As the neighbourhood was a moving one, with a radius of 136 meters, 95550 estimations were carried out.

The basic statistics of mentioned estimates can be seen in Table 8 and all of them can be seen, graphically, in the estimation map (Figure 13) and the estimation surface (Figure 14) where the house price is seen as an elevator variable.

Table 8. Basic statistics of ordinary kriging estimations of prices of "equivalent" houses

	Min.	Q_{25}	Q_{50}	Q_{75}	Max.	Mean	Standard Deviation	Variation coefficient
Estimated price of "equivalent" houses (m^2)	1244.0	1618.82	1805.41	1937.73	2810.0	1797.57	230.91	0.128457
Standard Deviation	0	268.44	311.76	355.36	421.91	313.19	60.39	0.19283

Fig. 13. Kriging house prices. Estimation map.

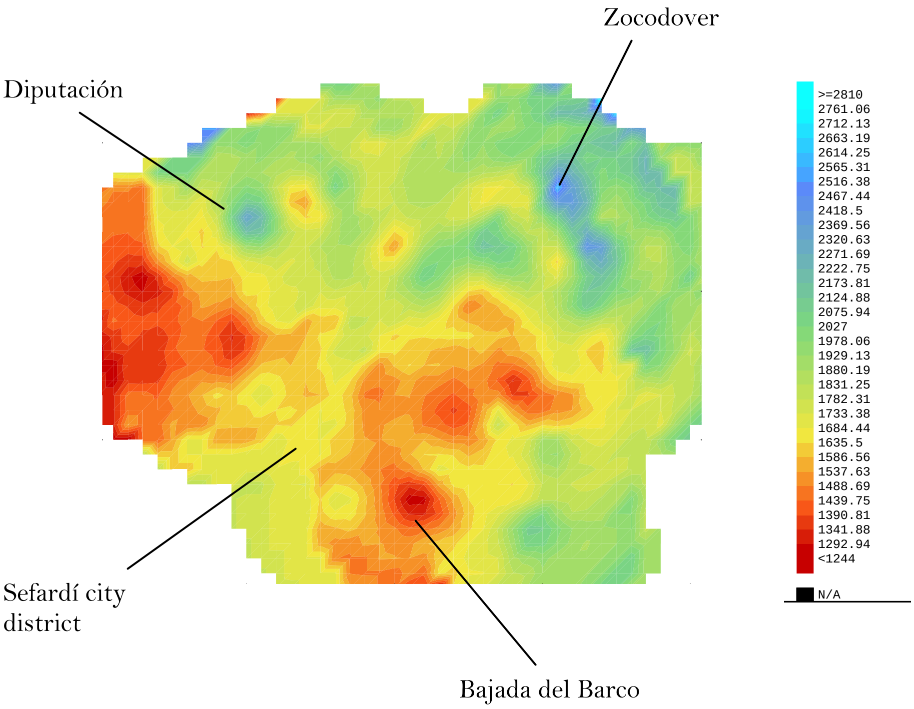
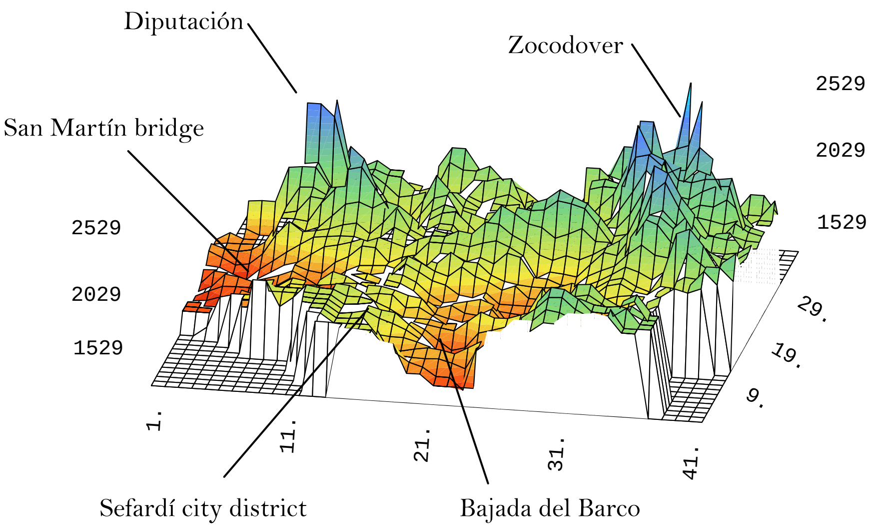


Fig.14. Kriging house prices. Estimation surface.



The Base Map (Figure 9), the experimental variogram (Figure 10), the fitted variogram (Figure 11) the estimation map (Figure 13) and the estimation surface (Figure 14) have been obtained using ISATIS, a spatial statistical program jointly developed by Geovariances[15] and L'Ecole des Mines from Paris.

As it can be appreciated, in the estimation map as well as in the estimation surface, the characteristics with respect to the factors "house condition", "surface" and "garage" being equal, it is easy to recognize the areas where house prices are cheaper (red zones) and the locations where it has been estimated a higher price (blue zones).

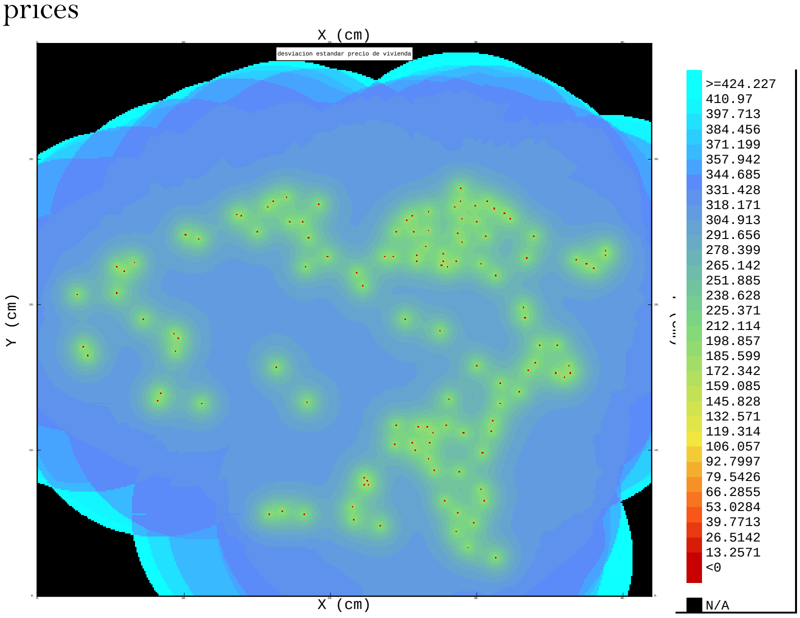
It can be remarked two well differenced zones where the price per square meter overcomes 2200 €: northwest inside the walls, where it is sited the *Diputación*, including the *Real* Street and *La Merced* Street; in the northeast, nearer *La Sillería* Street, *San Agustín* square and surrounding, so as the emblematic *Zocodover* square and the nearest streets. In the central zone of the historical area, near to *San Juan Bautista* church and the *Sefardí* city district, prices oscillate between 1600 y 2200 euros per square meter. However, prices in the *Sefardí* district are very close to the mean price, possibly due to the limited information available in that particular zone. Finally, we can note prices lower than 1600 €/m^2 in *Bajada del Barco* street, *la Cornisa* zone and *San Martín* bridge.

Figure 15 shows the map of the standard deviation of the estimation error. When the colour becomes green, the standard deviation decreases. It can be seen 121 points with the maximum colour intensity (in red colour) and, consequently with null standard deviation. They correspond to the sampled locations because

[15] See http://www.geovariances.fr

kriging estimator is an exact interpolator. Note that where sampled locations are more concentrated the variability of the estimation error, in standard deviation terms, oscillate between 100 and 200 €/m² while in zones with few sampled locations the standard deviation increases until 250-350 €/m². That is why the more accurate estimations have been placed near the *Diputación* zone, surrounding *Sillería* street, near the *Zocodover* square and the *Alcázar*, so as the *Pozo Amargo* area, decreasing the precision, logically in the centre of the historical area, due to the amount of historic monuments there placed.

Fig. 15. Standard deviation of estimation error obtained by kriging the house prices



Therefore, the ordinary kriging procedure carried out in the historical area of Toledo provides estimates in all and each location of this emblematic zone of the city. These prices would correspond to an equivalent set of houses, and real estimates would be easily computed by incorporating the factor effects relative to each house.

4.3. Cokriging estimation of premises prices in the Historic City of Toledo.

Having been exposed the theoretical concepts of cokriging methodology in the case of the joint intrinsic hypothesis and partial heterotopy in section 3.2.1., we next proceed to show the results obtained from the application of this multivariate spatial estimation procedure to the premises prices in the old part of Toledo city (Spain), taking the price of houses in that area as an auxiliary process.

There are several reasons for having chosen this emblematic area: (i) It is a World Heritage City, (ii) it is an excellent area for exploring the commercial real state market due to its tourist character and (iii) it has neither geographic accidents nor artificial barriers inside the walls that could break down the spatial dependence structure. The database contains information about 223 houses and 123 premises sited in the historical part of Toledo city. The locations of the houses and premises within the study area are depicted in Figure 16.

Fig. 16. Premises (+) and houses (∘) locations, being the symbols proportional to the prices.

The information has been provided by the real estate agencies that operate in this historical area and it refers to the market price, age, location, condition and surface. Additionally, it is known whether the premises have basement or not, and, in the case of houses, whether they have parking or not. Evidently, the age of a property usually has an important influence in its price, but in a historical part of a city as Toledo the influence of this factor vanishes. This is the reason why it has not been considered in the analysis. Moreover, we have detected some deficiencies in the measurement of the surface, having decided to consider it as a categorical variable. Obviously, there also exist more explanatory variables but unfortunately they are not provided by the real estate agencies for research purposes. The data correspond to commercial premises and houses for sale in the third quarter of 2004.

Obviously, in the original database the prices are unadjusted for housing and premises mix. So, we do not know at this point if the higher prices in some areas reflect higher property values per square meter or if the houses or premises in those areas possess some features that make them more expensive.

In order to isolate the spatial component of premises and house prices we have proceeded to adjust for housing and premises mix as follows: It has been tested if all the levels of every characteristic of premises and houses we have information about (see Tables 9 and 10), have the same effect on the price. In the case that this hypothesis is rejected, the significant differences have been estimated and removed from prices. Once removed these differences, houses and premises are equivalent with regard to the features considered (in this sense we have an "equivalent class" of houses and another one of premises) and the variability of the "new" prices is attributable to the spatial location of the properties. In concrete, factors and levels considered have been the following ones:

Table 9: Premises. Factors and levels

	Condition	Surface	Basement
	Ready for business	Less than 50 m^2	Yes
Premises	Some renovation needed	From 50 to 100 m^2	Not
	Complete renovation needed	From 100 to 200 m^2	
	Unfinished	More than 200 m^2	

Table 10: Houses. Factors and levels

	Condition	Surface	Parking
	New or completely renovated	Less than 65 m^2	Yes
Houses	In a good condition	From 65 to 120 m^2	Not
	Little renovation needed	More than 120 m^2	
	Complete renovation needed		

From now to the end, the premises and house prices we work with are the equivalent ones.

As pointed out in the introduction, from our point of view, the problem of estimating premises prices can only be adequately analyzed by taking into account the relative locations of the observations because spatial correlation is a typical characteristic of price of properties. So, after having constructed the databases of equivalent prices, we have next proceeded to represent the spatial dependence in both houses and premises markets by the appropriate theoretical variogram model, as well as to account for cross-dependence between both processes selecting the suitable cross variogram.

The experimental cross and direct variograms appear in Figure 17 with their respective fitted models. The values of the parameters are reported in Table 11.

Fig. 17: Experimental and fitted (a) *premises prices* direct variogram, (b) *house prices* direct variogram, (c) *premises prices-house prices* cross-variogram.

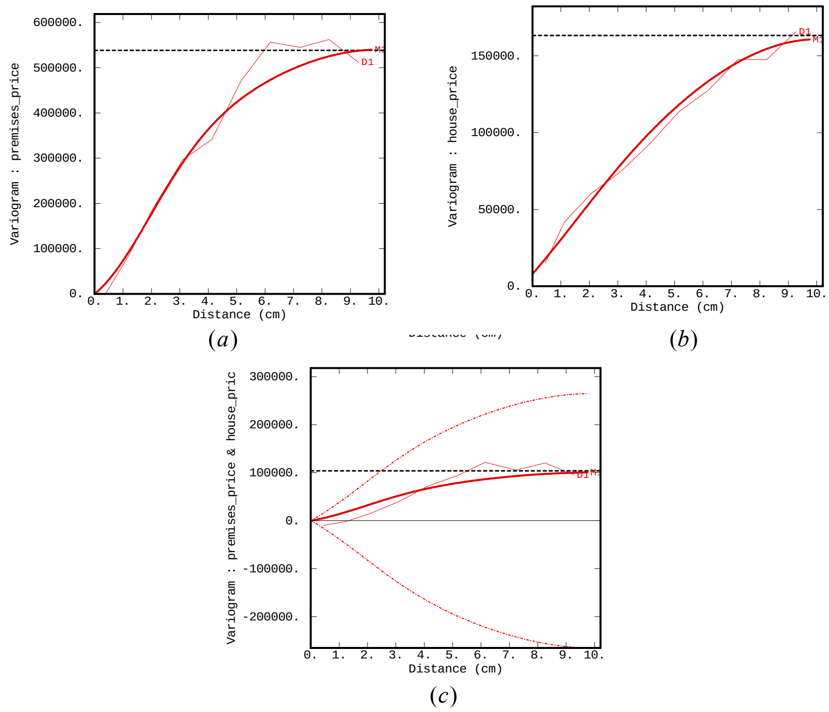


Table 11. Linear Model of Coregionalization

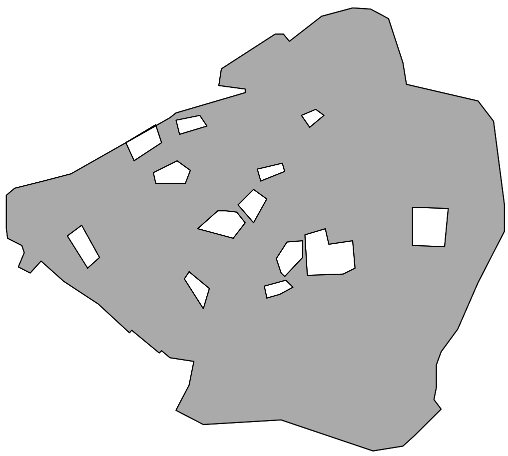
Model	Sill		
	Premises prices direct variogram	House prices direct variogram	Premises prices-house prices cross variogram
Spherical – 330m. range	340978.332	142783.006	70505.189
Nugget effect	1	8000	-85
Gaussian – 165m. range	200000	10000	30000

Cokriging results are dependent on the autocorrelation model of the principal and auxiliary processes, as well as on the cross-correlation model; hence, the variogram modeling process and the need for high-quality models are of paramount importance. The validation procedure may be done rigorously by having a separate set of sample data against which to compare cokriged estimates, but in research studies this means a waste of information, and validation usually is done by cross-validation. This is the reason for what in our case the proposed variograms have been checked for validity using the cross-validation or "leave-one-out" procedure (see, for example, Sinclair and Blackwell, 2002, pp. 221). In concrete, models from Table 11 provide 119 robust estimates when estimating premises prices (96,7% from a total of 123) and 214 in the case of the house prices (96% from a total of 223), a estimate being robust when its standardized value belongs to the interval $[-2.5; 2.5]$ (see Emery, 2000, pp. 117-118 and Chilès and Delfiner, 1999, pp. 112-113). These percentages of robust estimates (greater than 95%) lead us to consider models from Table 11 and Figure 14 valid for cokriging estimation.

Once decided the combination of theoretical variograms that better fits the structure of the spatial dependence in the area under study, we can proceed to estimate the premises prices by using the cokriging methodology. In concrete, as the fitted variograms stabilize around the variance of the data, the random functions relative to the price of premises and houses can be considered second-order stationary and OCK is used to map the estimates. We have also estimated the price of premises by OK. The aim is to compare both procedures and check, as expected from the theoretical literature, that OCK is more accurate than OK.

To perform the OCK estimation we have initially designed a polygon representing the outline of the study area (Figure 18), the holes corresponding to the places occupied by cultural buildings, as the Cathedral, the Alcazar, Christian churches, Islamic monuments, Synagogues, etc.

Fig. 18: Historic part of Toledo city's polygon.



Subsequently, we have drawn a regular grid of 3.30 meter mesh over the above mentioned polygon, having performed the estimation in the nodes of the grid. As the neighborhood was a moving one with a radius of 132 meters, 68911 estimations were carried out.

Finally, these 68911 estimates are depicted in the OCK estimation map (Figure 19, where the price per square meter is considered as an XY projection) and the OCK estimation surface (Figure 20, where the premises price is seen as an elevator variable). The basic descriptive statistics of these OCK estimates are reported in Table 12.

Fig. 19. Cokriging estimation map for the “equivalent” premises prices (€/m²)

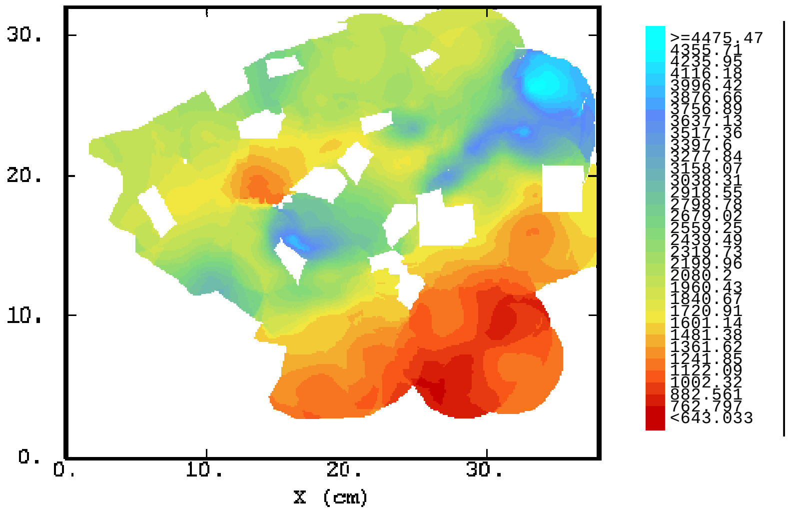


Fig 20. Cokriging estimation surface for the “equivalent” premises prices (€/m²)

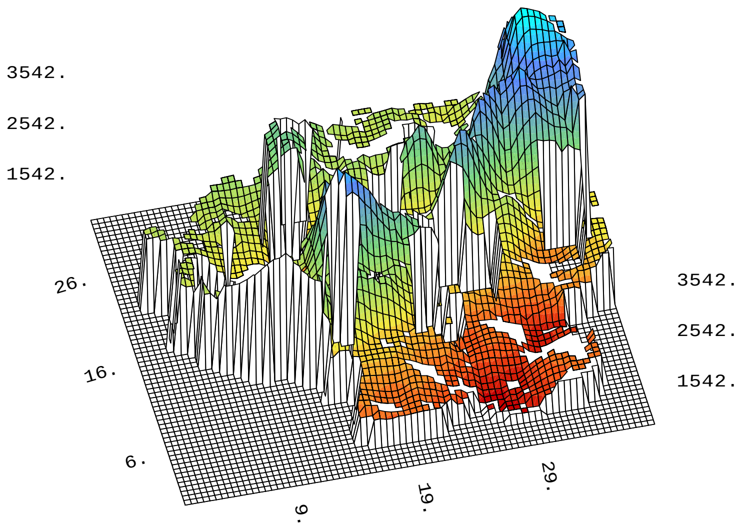


Table 12. Basic statistics for ordinary cokriged estimates of the price of “equivalent” premises

	Min.	**Q_{25}**	**Q_{50}**	**Q_{75}**	**Max.**	**Mean**	**Standard deviation**	**Variation coefficient**
Cokriged price (€ /m²)	643.03	1399.28	1968.76	2316.51	4475.47	1975.77	738.18	0.374
Standard deviation	0	229.74	339.98	544.91	852.95	395.83	197.74	0.499

As it can be appreciated from the estimation map, the areas where premises prices are cheap (red zones) are easily distinguished from the areas where they are more expensive (blue zones). The estimation surface shows, even most clearly than the estimation map, a general view of the cokriged prices of premises across the area under study and the differences among them.

In particular, the OCK estimation map (Figure 16) reveals that the highest prices per square meter appear, as expected, in the tourist zone: (i) the north-east part of the study area, corresponding to the emblematic *Zocodover* Square, the Cathedral surroundings and the streets that connect both zones, and (ii) the Sefardí district, in the south-west of the polygon. In both areas prices exceed $3000\,{}^{€}\!/_{m^2}$. There are other two areas with prices between 2500 and $3000\,{}^{€}\!/_{m^2}$, corresponding to the nearby of the place where tourist buses leave visitors: the escalator to the old city (in the north-west) and the old city's main entry point (*Bisagra* Gate, in the south-west). In the north, prices range from 1500 to $2500\,{}^{€}\!/_{m^2}$, while, finally, in the south-east area (the darkest one) prices are lower than $1000\,{}^{€}\!/_{m^2}$.

Fig. 21. Standard deviation map obtained by cokriging

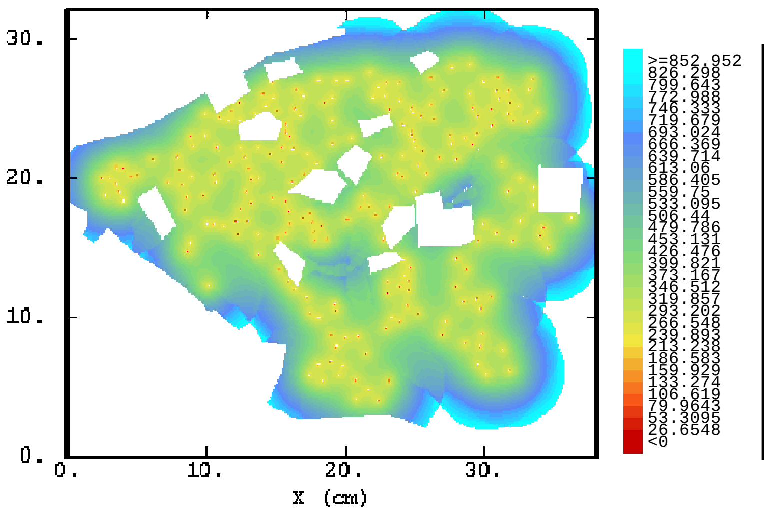


Figure 21 shows the standard deviation map: The darker the colour, the higher the standard deviation. It can be clearly appreciated 123 points in red, that is, with null standard deviation; obviously they correspond to the sampled locations, as OCK is an exact multivariate interpolator. Note that in the areas more sampled, the variability of the estimation error, in standard deviation terms, ranges between 100 and 200 €/m^2, while in zones with few sampled locations the standard deviation increases to 300-350 €/m^2 . It can also be appreciated that the more is the distance between the estimated points to the sampled locations, the more increases the standard deviation; this fact implies that the accuracy of the estimates decreases dramatically in locations separated from the sample site.

The ordinary cokriging procedure carried out provides estimates in all and each location of the area under study. These prices would correspond to an equivalent set of premises, and real estimates would be easily computed by incorporating the factor effects relative to each premises.

Figures 16 to 21 have been obtained by using ISATIS, a spatial statistical program jointly developed by Geovariances[7] and L'Ecole des Mines de Paris.

Once obtained the OCK estimation and surface maps, next we proceed to compare the results obtained by OCK multivariate methodology and the OK univariate procedure. Comparison is based on the same actual real estate data. OK estimates have been computed incorporating in the weighting mechanism the premises prices direct variogram reported in Table 11, in order to compare them with the OCK ones. The comparison criterion is the interpolation accuracy when carrying out a cross validation procedure. In particular, cokriging versus. kriging estimation variances are compared. Aditionally, OCK and OK standardized errors $\left({(x^*(s_0)-x^*(s_0))}/{\sigma^*(s_0)} \right)$ have been calculated. Another possibility could have been to create a test set and, eventually, compute the MAE and RMSE (see Li and Revesz, 2004) of the differences between the estimated value of price per square meter and the true value. However, taking into account the sample size, it has been preferred the cross validation method. The comparison results are reported in Table 13.

Table 13. Cross-validation results

Interpolation Method	**Error**		**Standardized error**	
	Mean	Variance	Mean	Variance
Kriging	-1.672	91186.233	-0.015	1.097
Cokriging	1.501	80621.447	0.0045	0.922

From Table 13 it can be concluded that when comparing OK versus OCK results, as expected (the correlation coefficient between premises and house prices, computed with the 65 pairs of prices corresponding to the homotopic case, is $\rho = 0.696$), OCK procedure has several advantages: (i) The mean estimation error decreases by 10.2% (from -1.672 to 1.5011) the OK result, (ii) the mean error, in standardized terms, decreases by 70% (from -0.015 to 0.0045), (iii) the variance of the estimation errors decreases by 11.6% (from 91186.233 to 80621.447) the variance of the OK ones and (iv) the reduction in variance increases to 15.95% (from 1.097 to 0.922) when standardized errors are considered.

5. Summary and conclusions

This paper has analyzed how the space can be taken into account when estimating a value from two different points of view: a spatial econometric and a spatial statistic approaches. Related to spatial econometric, the models with spatial dependence have been briefly shown and related to spatial statistic, kriging and cokriging have been widely exposed. The aim of kriging is to estimate the value of a random function at an unobserved point, given the values of the function at some other points, when the variables are spatially correlated. It provides the best linear estimator, in the sense that it is unbiased and minimum error variance in the presence of spatial dependence. This method uses variogram to express the spatial variation being its main advantage over other univariate statistic methods, i.e., it takes into account the structure of the spatial dependence existing between the variables of the stochastic process. However, if the process of interest is difficult or expensive to measure and if some additional information about other phenomena correlated with the principal one exists, then, cokriging improve kriging estimations in the sense that it obtains estimations with less variance of the estimation error using the multivariate variogram and the multivariate data available.

In first and second applications, we have shown how the importance of house prices knowledge seems to be evident in an actual context where the activity of the economic agents is very intensive. To estimate the mean of house prices in a particular area, and taking into account that prices of houses are spatially correlated, it has been proposed the kriging of the mean estimator because it is the best linear one in this case. Evidently, the kriging estimator of the mean provides a different estimate than the arithmetical mean does and its variance decreases notably with respect to the variance of the Spanish official estimator. But what is really useful for citizens is to know the price of houses anywhere they want to buy or sell them. To obtain an estimation map and an estimation surface of house prices over a geographical area, taking into account that prices of houses are spatially correlated, it has been proposed some punctual kriging methods applied to the called “equivalent" houses, that is, houses with the same characteristics. These procedures have been applied to the historical area of Toledo city, having estimated, firstly, the structure of the spatial dependence among the prices of the equivalent houses by a linear regionalized model and, finally, having obtained an estimation map and an estimation surface. Of course, these methodologies are easy to expand to another village or city. The important thing is that the estimation maps and the estimation surfaces, with no doubt, would be more useful to sellers, buyers and economic authorities than the now known mean prices.

In the third application also related to the real estate market, we have shown the importance of considering the structure of the spatial dependence among the prices of properties when estimating them. Furthermore, the existing correlation between the prices of different types of properties (in our case, houses and premises) has been used to obtain more accurate estimates, as available information about premises prices is usually smaller than available information about house prices. Before obtaining any estimates, we have proposed a procedure to study the spatial structure. It comprises three steps: (i) Adjusting for housing and premises mix in order to isolate the spatial component of premises and house prices; (ii) Testing for spatial correlation and identifying a global pattern of that spatial correlation; (iii) Modelling the direct and cross-variograms according to the linear model of corregionalization to ensure a positive definite model. Next, we have considered the OCK methodology to improve OK estimates by adding an auxiliary stochastic process corresponding to the house prices in the study area. As expected, and in agreement with specialized literature on geostatistics, our results have shown that OCK has a clear advantage over OK in comparison with estimation using only the data of the target variable. Therefore, the results have indicated that the use of an auxiliary variable improves OK estimates, what is crucial when the extent of the information on the main variable is not the desirable. This is precisely the case when estimating premises prices as information on them is usually scarce.

References

Amstrong, M. (1984), Improving the Estimation and Modelling of the Variogram, *Geostatistics for Natural Resources Characterization, Verly et al. eds.* 1, 1-20.

Anderson, S.T. and West, S.E., (2006) Open Space, Residential Property Values, and Spatial Context, *Regional Science and Urban Economics* 36(6), 773-789.

Anselin, L., (1980) *Estimation Methods for Spatial Autoregressive Structures*, Ithaca, NY: Cornell University, Regional Science Dissertation and Monograph Series.

Anselin, L., (1986) Some Further Notes on Spatial Models and Regional Science, *Journal of Regional Science*, 26, 799-802.

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

Anselin, L., (1995) Local Indicator of Spatial Association-LISA, *Geographical Analysis*, 27 (2).

Anselin, L., (2001) *Spatial Externalities, Spatial Multipliers and Spatial Econometrics*, Discussion paper of Regional Economics Applications Laboratory, REAL, 01-T-11

Anselin, L., (2005) *Exploring Spatial Data With Geoda: A Workbook*, Department Of Geography University of Illinois, Urbana-Champing.

Barlett, M. (1975) *The Statistical Analysis of Spatial Pattern*, London. Chapman & Hall.

Basu, S. and Thibodeau, T.G., (1998) Analysis of Spatial Correlation in House Prices, *Journal of Real Estate Finance and Economics* 17 (1), 61-85.

Bennett, R. (1979) *Spatial Time Series*, London. Pion.

Blommestein, H. (1983) Specification and Estimation of Spatial Econometrics Models: A Discussion of Alternative Strategies for Spatial Economic Modelling, *Regional Science and Urban Economics*, 13, 251-270.

Buchanan, G.R. (1995) *Finite Element Analysis.* New York: Mcgraw-Hill.

Case, B., Clapp, J.M., Dubin, R.A. and Rodríguez, M. (2004) Modeling Spatial and Temporal House Price Patterns: A Comparison of Four Models, *Journal of Real Estate Finance and Economics* 29 (2), 167-191.

Chasco. C., (2003) *Econometría Espacial Aplicada a la Predicción-Extrapolación de Datos Microterritoriales.* Tesis Doctoral. Universidad Autónoma De Madrid.

Cheshire, P. and Sheppard, S.C. (1995) On the Price of Land and the Value of Amenities, *Economica*, 62 (246), 247-267.

Chica Olmo, J., (1994) *Teoría De Las Variables Regionalizadas. Aplicación en Economía Espacial y Valoración Inmobiliaria.* Universidad de Granada. Granada.

Chilès, J.P. and Delfiner, P. (1999) *Geostatistics: Modelling Spatial Uncertainty.* New York: Wiley & Sons.

Chistakos, G. (2000). *Modern Spatiotemporal Geostatistics*, Oxford University Press, New York.

Clapp, J.M., Kim, H.J. and Gelfand, A.E. (2002) Predicting Spatial Patterns of House Prices Using LPR and Bayesian Smoothing, *Real Estate Economics* 30(4), 505-532.

Cliff, A. and Ord, J. (1973) *Spatial Autocorrelation*, London. Pion.

Cliff, A. and Ord, J. (1981) *Spatial Processes: Models And Applications*, London: Pion Ltd.

Cressie, N and Hawkins, D.M. (1980) Robust Estimation of the Variogram, *Journal of The International Association of Mathematical Geology.* 12 (2), 115-125.

Cressie, N (1993) *Statistics for Spatial Data.* Wiley Series in Probability and Mathematical Statistics. John Wiley & Sons, Inc. New York.

Davis, J.C., (1986). *Statistics and Data Analysis in Geology.* John Wiley & Sons.

Deutsch, C.V. and Journel, A.G. (1998) *GSLIG: Geostatistical Software Library and User's Guide.* 2ndedition. New York: Oxford University Press.

Din, A., Hoesli, M. and Bender, A. (2001) Environmental Variables and Real Estate Prices, *Urban Studies* 38(11), 1989-2000.

Dowd, P.A. (1984), The Variogram and Kriging. Robust and Resistant Estimators, *Geostatistics for natural resources characterization* 221-236. Eds. M. Guarascio, M. David, y C. Huijbregts, D. Reil, Dordrecht.

Dubin, R.A. (1998) Predicting House Prices Using Multiple Listing Data, *Journal of Real Estate Finance and Economics*, 17 (1), 144-164.

Emery, X. (2000) *Geoestadística Lineal.* Departamento de Ingeniería de Minas. Facultad de Ciencias Físicas y Matemáticas, Universidad de Chile. Chile.

Fik, T.J., Ling, D.C. and Mulligan, G.F. (2003) Modelling Spatial Variation in Housing Prices: A Variable Interaction Approach, *Real Estate Economics* 31(4), 623-646.

Fisher, M. and Nijkamp (1984), *Theory And Testing In Empirical Sciences.* Aldershot. Gower.

Florax R. and Folmer H., (1992), Specification and Estimation of Spatial Linear Regression Models: Monte Carlo Evaluation of Pre-Test Estimators. *Regional Science and Urban Economics*, 22, 405-432.
Foster, S., and Gorr, W. (1986) An Adaptive Filter for Estimating Spatially-Varying Parameters: Application to Modelling Police Hours Spent in Response to Calls for Service. *Management Science*, 32, 878-889.
Fotheringham, A.S., Brunsdon, C. and Charlton, M. (2002) *Geographically Weighted Regression. The Analysis of Spatially Varying Relationships.* Chichester: John Wiley & Sons, LTD.
Gámez, M., Montero, J.M. and García, N. (2000) Kriging Methodology for Regional Economic Analysis: Estimating the Housing Price in Albacete, *International Advances In Economic Research*, 6 (3), 438-451.
Gelfand, A.E., Ecker, M.J., Knight, J.R. and Sirmans, C.F. (2004), The Dynamics of Location in Home Price, *Journal of Real Estate Finance and Economics* 29 (2), 149-166.
Goodman, A.C. (1978) Hedonic Prices, Price Indices and Housing Markets, *Journal of Housing Research*, 3, 25-42.
Goodman, J.E. and O'Rourke, J. (1997) *Handbook of Discrete and Computational Geometry.* Boca Raton, New York: CRC Press.
Goovaerts, P. (1997) *Geostatistics for Natural Resources Evaluation.* New York: Oxford University Press.
Granelle, J.J., (1970) Espacio Urbano y Precio Del Suelo (Sirey, Paris).
Grenander, U., (1954) On The Estimation of Regression Coefficients in the case of an Autocorrelated Disturbance, *Annals of Mathematical Statistics* 25, 252-272.
Grenander, U. and M. Rosemblatt, (1984) Statistical Analysis of Stationary Time Series (2ª Ed.) (Chelsea Publishing Co., New York).
Griffith, D. (1980) Towards a Theory of Spatial Statistics, *Geographical Analysis*, 12, 325-339
Griffith, D. (1987) Towards a Theory of Spatial Statistics: Another Step Forward, *Geographical Analysis*, 19, 69-82.
Guigou, J.L., (1982) La Rente Foncière: Les Theories Et Leur Evolution Depuis 1650 (Economica, Paris).
Haining, R. (1986) Spatial Models and Regional Science: A Comment on Anselin's Paper and Research Directions, *Journal of Regional Science*, 26, 793-798.
Haining, R.P., (1988) Estimating Spatial Means with an Application to Remotely Sensed Data, *Communications in Statistics. Theory and Methods* 17, 573-597.
Han, S.S., (2004) Spatial Structure of Residential Property-Value Distribution in Beijing and Jakarta, *Environment And Planning A*, 36 (7), 1259-1283.
Harbaugh, J.W. and Preston, F.W. (1968) Fourier Analysis in Geology, in: Englewood Cliffs: B J L Berry, D F Marble (Eds.) *Spatial Analysis: A Reader.* Prentice-Hall
Hepple, L.W. (1998) Exact Testing For Spatial Autocorrelation Among Regression Residuals, *Environment And Planning A*, 30 (1), 85-108.
Hordijk (1974), Spatial Correlation in the Disturbance of a Linear Interregional Model, *Regional Science and Urban Economics*, 4, 117-140.

Hordijk, L. and Paelink, J., (1976), Some Principles and Results in Spatial Econometrics, *Recherches Economiques de Louvain*, 42, 175-197.
Ihlanfeldt, K.R. and Shaughnessy, T.M. (2004) An Empirical Investigation of the Effects of Impact Fees on Housing and Land Markets, *Regional Science and Urban Economics* 34(6), 639-661.
Isaaks, E. H., and Srivatstava, R. M., (1989). A*n Introduction to Applied Geostatistics*, Oxford University Press.
Johnston, K., Ver Hoe, J.M., Krivoruchko, K. and Lucas, N. (2001) *Using Arcgis Geostatistical Analyst.* Redlands: ESRI Press.
Journel, A.G. and Huijbregts, C.J. (1978) *Mining Geostatistics.* London: Academic Press.
Krige, D.G., (1951) *A Statistical Approach to Some Mine Valuation and Allied Problems on the Witwatersrand*, M.Sc. Eng. Thesis, University of Witwatersrand, Johannesburg.
Krige, D.G., 1951, A Statistical Approach to Some Basic Mine Valuation Problems on the Witwatersrand, *Journal of Chemical, Metallurgical and Mining Society of South Africa* 52, 119 A 139.
Lark, R.M. and Papritz, A. (2003) Fitting a Linear Model of Corregionalization for Soil Properties Using Simulate Annealing, *Geoderma*, 115, 245-260.
Larraz, B., (2004) *Técnicas de Cokrigeado para el Análisis Económico. Estimación de Precios de Bienes Inmuebles en el Casco Histórico de la Ciudad de Toledo*" Tesis Doctoral. Universidad de Castilla-La Mancha.
Lesage, J., (1999) *Spatial Econometrics.* Department of Economics, University of Toledo.
Lesage, J., and Pace, K., (2004) *Spatial and Spatiotemporal Econometrics*, Elsevier.
Lesage, J., and Pace, K., (2004) Models for Spatially Dependent Missing Data, *Journal of Real Estate Finance and Economics*, 29 (2), 233-254.
Leung, Y., Mei, C.L. and Zhang, W.X. (2003) Statistical Test for Local Patterns of Spatial Association, *Environment And Planning A*, 35 (4), 725-744.
Li, L. and Revesz, P. (2002) A Comparison of Two Spatio-Temporal Interpolation Methods, *Lecture Notes In Computer Science*, 2478, 145-160.
Li, L. and Revesz, P. (2004) Interpolation Methods for Spatio-Temporal Geographic Data, *Computers, Environment and Urban Systems*, 28 (3), 201-227.
Matheron, G. (1965). *Les Variables Régionalisées et leur Estimation. Une Application de la Théorie des Functions Aléatories aux Sciences de la Nature*, Masson & Cie, Eds.
Matheron, G. (1989) *Estimating and Choosing* (Springer-Verlag, Berlín).
Mathur, S., Waddell, P. and Blanco, H. (2004) The Effect of Impact Fees on the Price of New Single-Family Housing, *Urban Studies* 41 (7), 1303-1312.
Militino, A.F., Ugarte, M. D. and García-Rinaldos, L. (2004), Alternative Models for Describing Spatial Dependence Among Dewelling Prices, *Journal of Real Estate Finance and Economics* 29 (2), 193-209.
Ministerio De Vivienda (2005) Estadística De Precios De Vivienda.
Ministerio De Vivienda (2006) Estadística De Precios De Vivienda.
Ministerio De Vivienda (2007) Notas Metodológicas. Http://Www.Mviv.Es

Montero Lorenzo, J.M. (2004) El Precio Medio del Metro Cuadrado de la Vivienda: Una Aproximación desde la Perspectiva de la Geoestadística,.*Estudios de Economía Aplicada*, 22-3, 675-693.

Montero Lorenzo, J.M. and Larraz Iribas, B., (2006a) Estimación Espacial del Precio de la Vivienda Mediante Métodos de Krigeado, *Revista Estadística Española* 48 (162), 201-240.

Montero-Lorenzo, J.M. and B. Larraz-Iribas, (2006b), Estimating Housing Prices: Kriging the Mean - Research Note, *International Advances in Economic Research* 12 (2), 419.

Montero-Lorenzo, J. M. and B. Larraz-Iribas, (2006c), Cokriging Versus Univariate Interpolation Methods: An Application to the Premises Market, *International Workshop on Spatio-Temporal Modelling (METMA3)*, 223-226.

Moran, P., (1948) The Interpretation of Statistical Maps. *Journal of the Royal Statistical Society B.*, 10, 243-251.

Moreno R. and Vayá E. (2002) *Econometría Espacial: Nuevas Técnicas para el Análisis Regional. Una Aplicación a las Regiones Europeas.* Investigaciones Regionales, España.

Omre, H. (1984), The Variogram and its Estimation. *Geostatistics for Natural Resources Characterization.* 1, 107-125.Verly et al. (Eds.). NATO ASI Series, Reidel.

Paelink, J.H.P. and Klaasen, L.H. (1979) *Spatial Econometrics.* Farnborough, Saxon House.

Pavía-Miralles, J.M. (2005). Forecasts from Non-random Samples: The Election Night Case. *Journal of the American Statistical Association*, 100, 1113-1122.

Richardson, H.W., (1986), *Economía Regional y Urbana* (Alianza Editorial, Madrid).

Ripley, B. (1981) *Spatial Statistics*, New York. Academic Press.

Roca, J., (1988), La Estructura De Valores Urbanos: Un Análisis Teórico Empírico (Instituto De Estudios De Administración Local, Madrid).

Samper, F.J. and Carrera, J. (1996), *Geoestadística. Aplicaciones a la hidrología subterránea.* Ed. CIMNE

Sinclair, A.J. and Blackwell, G.H. (2002) *Applied Mineral Inventory Estimation.* Cambridge: Cambridge University Press.

Tiefelsdorf, M. and Boots, B. (1995) The Exact Distribution of Moran´s I, *Environment And Planning A*, 27 (6), 985-999.

Upton, G. and Fingleton, B., (1985) *Spatial Data Analysis by Example.* Wiley Series in Probability and Mathematical Statistics. Jon Wiley & Sons.

Wackernagel, H., (1988) Geostatistical Techniques for Interpreting Multivariante Sptial Information. *Quantitative Analysis of Mineral and Energy Resources.* 393-409.

Wackernagel, H., (2003) *Multivariate Geostatistics. An Introduction with Applications.*3ª Ed. Springer-Verlag, Berlín.

Wartenbeg, D. (1985) Multivariate Spatial Correlation: A Method for Exploratory Geographical Analysis. *Geographical Analysis*, 17, 263-283.

Watson, D.F. (1981), Computing the N-Dimensional Delaunay Tessellation with Applications to Voronoi Polytopes, *The Computer Journal*, 24 (2), 167-172.
Webester, R. and Oliver, M.A. (2004) *Geostatistics For Environmental Scientists.* Chichester: John Wiley & Sons.
Yiu, C.Y. and Tam, C.S. (2007) Housing Price Gradient with Two Workplaces — An Empirical Study in Hong Kong, *Regional Science And Urban Economics* 37(3), 413-429
Zienkiewicz, O.C. and Taylor, R.L. (2000) Finite Element Method, in: Butterworth Heinemann (Ed.) *The Basis* 1, London.
Zurflueh, E.G. (1967) Applications of Two Dimensional Linear Wavelength Filtering, *Geophysics*, 32, Pp. 1015-1035.