

El precio medio del metro cuadrado de la vivienda libre: Una aproximación metodológica desde la prespectiva de la Geoestadística

MONTERO LORENZO, JOSÉ MARÍA

Departamento de Economía y Empresa. Universidad de Castilla-La Mancha.

C/ Cobertizo de San Pedro Martir, s/n. 45071 TOLEDO. E-mail: jose.mlorenzo@uclm.es

RESUMEN

El precio medio del metro cuadrado de la vivienda libre y su evolución en los últimos años es una de las cuestiones que más preocupan, tanto a la ciudadanía como a las autoridades políticas y económicas. La única referencia oficial sobre la cuestión es la publicada por el Ministerio de Fomento, obtenida a partir de las tasaciones hipotecarias. La capacidad de los promedios obtenidos en dicha estadística para una determinada área geográfica se ve notablemente mermada por el hecho de no considerar la existencia de correlación espacial de los precios de los bienes urbanos. Esta lacra nos ha incitado a la elaboración de un nuevo estimador lineal de dicho precio medio, un estimador krigeado, especialmente diseñado para el caso de existencia de correlación en un dominio espacial. Este estimador, (el mejor entre los lineales insesgados), conduce a unas conclusiones más realistas en términos del precio medio del metro cuadrado de vivienda libre.

Palabras clave: Estimador krigeado, covariograma, correlación espacial, efecto pedanía, precio de la vivienda, economía espacial.

The mean housing price per square meter: A methodological approach from a geostatistic perspective

ABSTRACT

The price of houses and its evolution in the last years is one of the issues that Spanish citizens and, of course, political and economic authorities are worried about it. The unique official reference is published by the Spanish "Ministerio de Fomento", on the basis of mortgages evaluations. The means obtained by the aforementioned statistics have some limitations because they do not take into account the spatial correlation of the real property prices. This handicap of the ministerial methodology has lead us to elaborate a new linear estimator, the kriging estimator of the mean price of houses, which has been specially designed for the case of existing correlation in a spatial domain. This new estimator, the best unbiased linear estimator, will lead us to more realistic estimations of the mean price of houses in the geographical areas in which we are interested.

Keywords: Kriging estimator, covariogram, spatial correlation, small district effect, house prices, regional economics.

Clasificación JEL: C13, C21, R10, R31.

Artículo recibido en febrero de 2004 y aceptado para su publicación en octubre de 2004.

Artículo disponible en versión electrónica en la página www.revista-eea.net, ref.: E-22311.

1. INTRODUCCIÓN

Resulta evidente, incluso para los no iniciados, la importancia del sector de la construcción en la actividad económica, y no sólo por su aportación al Valor Añadido Bruto, sino también por su carácter de sector impulsor de los demás sectores y por su especial incidencia en el empleo. Pero además, dentro del sector de la construcción, el subsector vivienda reviste unas características especiales por las imbricaciones sociales que conlleva, por ser la vivienda un bien de primera necesidad que requiere para su adquisición un gran esfuerzo económico. Por tanto, no es extraño que el precio de la vivienda sea, en la actualidad, uno de los problemas que más preocupan a la ciudadanía.

En nuestro país, la única referencia oficial disponible, publicada con regularidad, es la proporcionada por la Subdirección General de Información y Estadística del Ministerio de Fomento en la publicación “Índice de precios de las viviendas. Estadística de precio medio del m² (datos obtenidos de las tasaciones hipotecarias)”, en la cual se puede encontrar información, desde 1987, relativa al precio medio del metro cuadrado de la vivienda libre tasada en el ámbito de las Comunidades Autónomas y de las capitales de provincia (distinguiéndose la vivienda nueva de la usada), fundamentalmente.

La mencionada Subdirección General estima el precio medio del metro cuadrado de la vivienda mediante un laborioso pero sencillo proceso. Dicho proceso parte del precio del metro cuadrado de las viviendas individuales que han sido objeto de tasación por parte de una serie de entidades financieras de tasación; a partir de los precios medios individuales de las viviendas tasadas se genera un precio medio del metro cuadrado de la vivienda a nivel de código postal; posteriormente, y con base en el precio medio del metro cuadrado de la vivienda en los distintos códigos postales, se determina el precio a nivel de municipio; a partir de éstos se obtiene el precio representativo de las Comunidades Autónomas y, por último, tomando como base los precios de las Comunidades Autónomas se estima el precio por metro cuadrado a escala nacional.

En concreto, el proceso comienza con el cómputo, para cada vivienda tasada, del precio del metro cuadrado de superficie construida mediante la expresión donde $PMM_{(ijkh)}^2$ es el precio por metro cuadrado de la h -ésima tasación llevada a cabo en el código postal k del municipio perteneciente a la Comunidad Autónoma i

$$PMM_{(ijkh)}^2 = \frac{V_{(ijkh)}}{M_{(ijkh)}^2} \quad [1]$$

$V_{(ijkh)}$ y $M_{(ijkh)}^2$ son, respectivamente, el valor y el número de metros cuadrados de la tasación llevada a cabo.

El precio medio del metro cuadrado a nivel de código postal, es decir, para el código postal k del j -ésimo municipio de la Comunidad Autónoma i , se obtiene como sigue:

$$PMM_{(ijk)}^2 = \frac{\sum_{h=1}^{NT_{(ijk)}} PMM_{(ijkh)}^2}{NT_{(ijk)}} \quad [2]$$

donde $NT_{(ijk)}$ es el número de tasaciones realizadas en el código postal k del municipio j de la Comunidad Autónoma i .

El precio medio del metro cuadrado por municipio se calcula, a partir de la expresión anterior, como media aritmética de los precios por metro cuadrado en los distintos códigos postales del municipio:

$$PMM_{(ij)}^2 = \frac{\sum_{k=1}^{NCP_{(ij)}} PMM_{(ijk)}^2}{NCP_{(ij)}} \quad [3]$$

donde $NCP_{(ij)}$ es el número de códigos postales del municipio j de la Autonomía i .

El precio medio por metro cuadrado en cada Comunidad Autónoma se obtiene ponderando la media agregativa expuesta con el número de habitantes de hecho de los municipios que componen la Comunidad Autónoma i ($H_{(ij)}$).

$$PMM_{(i)}^2 = \frac{\sum_{j=1}^{NM_{(i)}} PMM_{(ij)}^2 H_{(ij)}}{\sum_{j=1}^{NM_{(i)}} H_{(ij)}} \quad [4]$$

donde $NM_{(i)}$ es el número de municipios que componen la Comunidad Autónoma i .

Finalmente, el precio medio del metro cuadrado a escala nacional se obtiene como

$$PMM_{(nac)}^2 = \frac{\sum_{i=1}^{NCA=17} PMM_{(i)}^2 H_{(i)}}{\sum_{i=1}^{NCA=17} H_{(i)}} \quad [5]$$

siendo NCA el número de Comunidades Autónomas españolas y $H_{(i)}$ el número de habitantes de la i -ésima Autonomía.

Sin embargo, consideramos que la metodología que utiliza la mencionada Subdirección General está viciada de raíz por obviar un hecho absolutamente relevante: los precios de los bienes urbanos están correlacionados espacialmente¹. Y ello afecta directamente al primer nivel de estimación (nivel de código postal) y, en consecuencia, a los demás². El precio medio del metro cuadrado en un determinado código postal se estima mediante la media aritmética de los precios medios de las viviendas tasadas. Sin embargo, en presencia de correlación espacial la media aritmética no es precisamente el mejor estimador de la media poblacional ya que, si bien es insesgado, no es el estimador de varianza mínima.

Pues bien, esta cuestión es la que se pone de manifiesto en el epígrafe 2. En el tercero se propone un nuevo estimador, el estimador krigeado, especialmente diseñado para el caso de correlación espacial entre los precios de los bienes urbanos. Finalmente, se ejemplificarán, mediante simulación, las diferencias entre los precios del metro cuadrado de vivienda libre obtenidos con el estimador actualmente utilizado por la Subdirección General de Información y Estadística del Ministerio de Fomento y a través del estimador krigeado propuesto, pudiéndose comprobar que llevan a series de precios bien distintas.

2. IMPORTANCIA DE LA CORRELACIÓN ESPACIAL EN LA ESTIMACIÓN Y PREDICCIÓN

La existencia de correlación espacial afecta al proceso tradicional de estimación de los parámetros poblacionales y a la predicción en los puntos de interés, en el sentido de que modifica la variabilidad de dichas estimaciones y predicciones respecto de la situación de independencia. En nuestro caso concreto, en el que la correlación espacial es, en la inmensa mayoría de los casos, positiva, constituirá un elemento gravoso a la hora de estimar el precio medio del metro cuadrado poblacional en un área geográfica concreta, ya que elevará la variabilidad de las estimaciones, mientras que, por el contrario, favorecerá la predicción del precio por metro cuadrado en cualquier localización particular de dicha área por cuanto la variabilidad de dichas predicciones se verá reducida a medida que aumenta la intensidad de la correlación.

¹ A estos efectos, véase Chica (1994), especialmente el capítulo III relativo a las consideraciones teóricas sobre la autocorrelación espacial del precio de los bienes urbanos. También resultan interesantes los trabajos de Granelle (1970), Guigou, (1982) y Roca (1988). También son recomendables los trabajos relativos al gradiente del precio de los bienes urbanos de Richardson (1975, 1986).

² Además, en el ámbito municipal tiene lugar un efecto secundario, el "efecto pedanía", derivado de la no ponderación de los distintos códigos postales del municipio por el número de tasaciones que en ellos se han llevado a cabo. Las consecuencias de este efecto, infraestimación del precio medio, pueden verse en Montero et al. (1998).

Por consiguiente, si en presencia de correlación espacial (positiva en el caso inmobiliario) se utilizan los estimadores propios del caso de independencia, los intervalos de confianza del precio medio del área de interés serán demasiado estrechos mientras que los de las predicciones en puntos concretos de tal área serán demasiado anchos.

Como se ha avanzado en el epígrafe introductorio, en el nivel primario de código postal la metodología ministerial utiliza como estimador la media aritmética de los precios por metro cuadrado de las viviendas tasadas en el mismo. Éste sería el mejor estimador lineal del precio medio del metro cuadrado en dicho código postal si los precios de los inmuebles no presentasen correlación espacial alguna; sin embargo, la incorrelación espacial de los precios de los bienes urbanos es un supuesto que no se puede aceptar. En el caso más realista de existencia de correlación espacial entre los mismos, la varianza del estimador media muestral se modifica sustantivamente al alza (si la correlación es positiva) o a la baja (si es negativa) y tanto más cuanto mayor sea dicha correlación.

A modo de ilustración, supóngase que $Z(s_1), Z(s_2), \dots, Z(s_n)$ son variables aleatorias que representan el precio de viviendas ubicadas, respectivamente, en las localizaciones geográficas $s_1, s_2, \dots, s_n$ de un determinado dominio. Supóngase igualmente que todas ellas son independientes y tienen media μ desconocida y varianza σ^2 conocida.

El estimador insesgado y de varianza mínima del parámetro μ es:

$$\bar{Z} = \frac{\sum_{i=1}^n Z(s_i)}{n} \quad [6]$$

con esperanza μ y varianza σ^2/n .

Considérese ahora la suposición más realista en el caso espacial de que las variables están correlacionadas positivamente y que dicha correlación disminuye a medida que aumenta la distancia entre ellas. Esta suposición podría venir representada por³

$$C(Z(s_i); Z(s_j)) = \sigma^2 \rho^{|s_i - s_j|} \quad i, j = 1, 2, \dots, n \quad 0 < \rho < 1 \quad [7]$$

³ La función de correlación está tomada del ya clásico ejemplo de Cressie (1993), p. 14. Tal y como afirma el mencionado autor, "Es bien conocido que tal función de correlación resulta de un proceso autorregresivo de primer orden $Z(s_i) = \rho Z(s_{i-1}) + \varepsilon(s_i) \quad i \geq 1$ donde $\varepsilon(s_i)$ i son parte de una secuencia de variables aleatorias gaussianas i.i.d. con media nula y varianza $\sigma^2(1 - \rho^2)$ independientes de $Z(s_{i-1})$ Sin embargo, éste no es el mecanismo generador que yo tengo en mente; aquí $\mathbf{Z} \equiv (Z(s_1), \dots, Z(s_n))$ son datos espaciales en R^1 tal que la predicción de $Z(s_0)$ o de $Z(s_{3/2})$ es tan apropiada como la de $Z(s_{n+1})$ ". La notación se ha adaptado a la del texto.

Ahora la varianza de $\bar{Z}$ ya no es σ^2/n sino

$$\begin{aligned}
 V(\bar{Z}) &= \frac{1}{n^2} V\left(\sum_{i=1}^n Z(s_i)\right) = \frac{1}{n^2} \left[n\sigma^2 + 2 \sum_{i \neq j} \sigma^2 \rho^{|s_i - s_j|} \right] = \\
 &= \frac{\sigma^2}{n} + 2 \frac{\sigma^2}{n^2} \left[(n-1)\rho^1 + (n-2)\rho^2 + \dots + [n - (n-1)]\rho^{n-1} \right] = \\
 &= \frac{\sigma^2}{n} + 2 \frac{\sigma^2}{n^2} \sum_{h=1}^{h=n-1} [n-h] \rho^h = \\
 &= \frac{\sigma^2}{n} \left[1 + \frac{2}{n} \sum_{h=1}^{h=n-1} [n-h] \rho^h \right] = \\
 &= \frac{\sigma^2}{n} \left[1 + 2 \frac{\rho}{1-\rho} \left(1 - \frac{1}{n} \right) - 2 \left(\frac{\rho}{1-\rho} \right)^2 \frac{1-\rho^{n-1}}{n} \right] \quad [8]
 \end{aligned}$$

De la expresión anterior se deduce fácilmente que el término entre corchetes es superior a la unidad en caso de que $\rho > 0$, situación habitual en el mercado inmobiliario, mientras que es inferior cuando $\rho < 0$. En consecuencia, en presencia de correlación espacial entre los precios de los bienes urbanos, la varianza de la media muestral de dichos precios aumentará, respecto de la situación de independencia, si $\rho > 0$, y disminuirá si $\rho < 0$ (si bien este no es el caso inmobiliario).

En definitiva, cuando existe dependencia espacial entre las variables, la varianza de la media muestral puede expresarse como donde, en el caso expuesto,

$$V(\bar{Z}) = \frac{\sigma^2}{n'} \quad [9]$$

$$n' = \frac{n}{\left[1 + \frac{2}{n} \sum_{h=1}^{h=n-1} [n-h] \rho^h \right]} = \frac{n}{\left[1 + 2 \frac{\rho}{1-\rho} \left(1 - \frac{1}{n} \right) - 2 \left(\frac{\rho}{1-\rho} \right)^2 \frac{1-\rho^{n-1}}{n} \right]} \quad [10]$$

interpretándose n' como el número de observaciones independientes que equivalen a un determinado número de observaciones, n , con dependencia espacial del tipo de la expuesta.

Por tanto, la correlación espacial de carácter positivo no sale “gratis” en cuanto a la precisión del estimador media muestral se refiere. Y ello ocurre incluso cuando el tamaño muestral es muy elevado. Obsérvese que, en este caso,

$$n' \approx \frac{n}{\frac{1+\rho}{1-\rho}} \quad [11]$$

Evidentemente, se pueden tomar modelos de correlación espacial más sencillos (por ejemplo, asumir correlación espacial perfecta) o más complicados que el presentado a modo de ilustración, pero la conclusión sería la misma⁴: La media muestral pierde precisión como estimador de la media poblacional y, tanta más, cuanto mayor sea la correlación espacial⁵.

Ante la pérdida de precisión de la media muestral como estimador lineal del precio medio del metro cuadrado en una determinada área geográfica en presencia de correlación espacial de carácter positivo, la cuestión que surge de inmediato es si dicho estimador, estimador ministerial para el precio medio del metro cuadrado de vivienda libre (vivienda tasada), sigue siendo el mejor estimador lineal y, en caso contrario, el objetivo será determinar cuál es el mejor estimador dentro de la clase de los lineales. De ello se ocupa el siguiente epígrafe.

3. PROPUESTA METODOLÓGICA DE ESTIMACIÓN DEL PRECIO MEDIO DEL METRO CUADRADO DE VIVIENDA LIBRE

La cuestión que se trata de resolver en este epígrafe es la siguiente: ¿Cómo se debería estimar el precio medio del metro cuadrado en un código postal, estrato básico de la metodología actual del Ministerio de Fomento en la estimación de precios medios, o en cualquier otra zona geográfica de interés?.

El estimador que se propone es un estimador del tipo

$$\bar{Z}_k = \sum_{i=1}^n \lambda_i Z(s_i) \quad [12]$$

⁴ A estos efectos, véanse los trabajos de Grenander (1954), Grenander y Rosenblatt (1984) y Haining (1988).

⁵ En caso de que se construyesen intervalos de confianza, si existe dependencia espacial positiva y no se considera, tal es el caso de la estimación ministerial, se obtendrían intervalos demasiado estrechos, o lo que es lo mismo, si se mantiene la amplitud del intervalo construido bajo la suposición de independencia se reduce el nivel de confianza del mismo

donde $Z(s_i)$ son los precios por metro cuadrado de las viviendas en las localizaciones s_i del código postal de interés en las que se dispone de información⁶.

La cuestión que se plantea a continuación es la relativa a la elección de los ponderadores λ_i . Asumamos que el proceso que genera la realización disponible en el código postal de interés es estacionario de segundo orden y que su media es μ . Sería deseable que el valor esperado del error de estimación fuese nulo, es decir

$$E[\bar{Z}_k - \mu] = E\left[\sum_{i=1}^n \lambda_i Z(s_i) - \mu\right] = 0 \quad [13]$$

o, en otros términos, que el estimador propuesto fuese insesgado.

Para ello, basta con requerir que la suma de los ponderadores sea unitaria. Efectivamente:

$$\begin{aligned} E[\bar{Z}_k] &= E\left[\sum_{i=1}^n \lambda_i Z(s_i)\right] = E\left[\sum_{i=1}^n \lambda_i Z(s)\right] = \sum_{i=1}^n \lambda_i E[Z(s)] \\ &= \sum_{i=1}^n \lambda_i \mu = \mu \sum_{i=1}^n \lambda_i = \mu \end{aligned} \quad [14]$$

Además de ser insesgado también sería deseable que el error cuadrático medio del estimador fuese mínimo o, bajo la condición de insesgadería, que la varianza del estimador fuese la menor posible. La varianza del estimador propuesto se puede escribir como

$$\begin{aligned} V[\bar{Z}_k] &= V\left[\sum_{i=1}^n \lambda_i Z(s_i)\right] = E\left[\sum_{i=1}^n \lambda_i Z(s_i)\right]^2 - \mu^2 = E\left[\sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j Z(s_i) Z(s_j)\right] - \mu^2 = \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j E[Z(s_i) Z(s_j)] - \mu^2 = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C[Z(s_i); Z(s_j)] + \mu^2 - \mu^2 = \\ &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(Z(s_i); Z(s_j)) \end{aligned} \quad [15]$$

⁶ Evidentemente no existe razón alguna por la cual limitarse a un área geográfica delimitada por la actividad postal. Es más, sería más interesante considerar áreas más homogéneas en lo que a precios de viviendas se refiere. Sin embargo, consideramos los códigos postales debido a que, para mantener el secreto estadístico, la información relativa a la ubicación de la vivienda hace referencia únicamente al código postal.

donde puede observarse que los productos cruzados $\lambda_i \lambda_j$ van ponderados por la covarianza entre cada par de localizaciones s_i y s_j .

Definida la varianza del estimador que, como puede apreciarse, no coincide con la tradicional varianza poblacional (en este caso del proceso) dividida por el tamaño de la muestra ya que el covariograma no se supone del tipo efecto pepita puro (caso de incorrelación espacial), el problema es elegir los ponderadores de tal forma que dicha varianza sea mínima.

Para ello no tenemos más que derivar la expresión anterior respecto de los ponderadores, eso sí, incluyendo la condición de inesgadez. En otros términos, a partir del lagrangiano

$$L(\lambda_i; \theta) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(Z(s_i); Z(s_j)) - \theta \left(\sum_{i=1}^n \lambda_i - 1 \right) \quad [16]$$

donde θ es el multiplicador de Lagrange. Derivando e igualando a cero se obtiene el siguiente sistema de ecuaciones:

$$\frac{\partial L(\lambda_i; \theta)}{\partial \lambda_i} = \sum_{j=1}^n \lambda_j C(Z(s_i); Z(s_j)) - \theta = 0 \quad i = 1, 2, 3, \dots, n$$

$$\frac{\partial L(\lambda_i; \theta)}{\partial \theta} = \sum_{i=1}^n \lambda_i - 1 = 0 \quad [17]$$

que matricialmente puede ser expresado como⁷

$$\begin{pmatrix} C(0) & C(s_1 - s_2) & C(s_1 - s_3) & \dots & C(s_1 - s_n) & -1 \\ C(s_2 - s_1) & C(0) & C(s_2 - s_3) & \dots & C(s_2 - s_n) & -1 \\ C(s_3 - s_1) & C(s_3 - s_2) & C(0) & \dots & C(s_3 - s_n) & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C(s_n - s_1) & C(s_n - s_2) & C(s_n - s_3) & \dots & C(0) & -1 \\ 1 & 1 & 1 & \dots & 1 & 0 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \\ \theta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad [18]$$

⁷ $C(s_i - s_j)$ representa la covarianza entre las variables $Z(s_i)$ y $Z(s_j)$. Ésta es la notación usual en la literatura geoestadística al uso. Por otra parte $C(0) = C(s_i - s_i) = V(Z(s_i))$ es la varianza del proceso.

de donde

$$\begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \vdots \\ \lambda_n \\ \theta \end{pmatrix} = \begin{pmatrix} C(0) & C(s_1 - s_2) & C(s_1 - s_3) & \cdots & C(s_1 - s_n) & -1 \\ C(s_2 - s_1) & C(0) & C(s_2 - s_3) & \cdots & C(s_2 - s_n) & -1 \\ C(s_3 - s_1) & C(s_3 - s_2) & C(0) & \cdots & C(s_3 - s_n) & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ C(s_n - s_1) & C(s_n - s_2) & C(s_n - s_3) & \cdots & C(0) & -1 \\ 1 & 1 & 1 & \cdots & 1 & 0 \end{pmatrix}^{-1} \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad [19]$$

Para obtener la varianza de este estimador no tenemos más que incluir las ecuaciones anteriores en la expresión (15), obteniéndose:

$$\begin{aligned} V \left[\sum_{i=1}^n \lambda_i Z(s_i) \right] &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(Z(s_i); Z(s_j)) = \sum_{i=1}^n \lambda_i \sum_{j=1}^n \lambda_j C(Z(s_i); Z(s_j)) \\ &= \sum_{i=1}^n \lambda_i \theta = \theta \sum_{i=1}^n \lambda_i = \theta \end{aligned} \quad [20]$$

Por tanto, los ponderadores que hacen mínima la varianza del estimador la igualan al multiplicador de Lagrange.

Nótese que, en caso de incorrelación espacial, las ecuaciones (20) se transforman en las siguientes:

$$\frac{\partial L(\lambda_i; \theta)}{\partial \lambda_i} = \lambda_i V(Z(s)) - \theta = 0 \quad i = 1, 2, 3, \dots, n$$

$$\frac{\partial L(\lambda_i; \theta)}{\partial \theta} = \sum_{i=1}^n \lambda_i - 1 = 0 \quad [21]$$

de donde se desprende (primera ecuación) que todos los ponderadores son iguales a

y, como todos ellos deben sumar la unidad, valen .

En consecuencia, en una tesitura de independencia, teniendo en cuenta que la covarianza entre dos variables distintas es nula, la varianza mínima del estimador $\bar{Z}_k$ será

$$V \left[\sum_{i=1}^n \lambda_i Z(s_i) \right] = \sum_{i=1}^n \sum_{j=1}^n \lambda_i \lambda_j C(Z(s_i); Z(s_j)) = \sum_{i=1}^n \lambda_i^2 \sigma^2 = \frac{1}{n^2} n \sigma^2 = \frac{\sigma^2}{n} \quad [22]$$

que coincide con la varianza de la media muestral cuando suponemos independientes las variables muestrales y que, a su vez, iguala la cota de Cramér-Rao para estimadores insesgados.

De lo anterior se deduce que el estimador propuesto tiene mayor precisión que la media muestral a la hora de estimar el precio medio del metro cuadrado en una determinada área geográfica en presencia de correlación espacial por lo que, sin duda alguna, será preferible a esta última. Dicho estimador es conocido en la literatura geoestadística como estimador krigado de la media (en honor a Krige⁸) y ha sido ampliamente utilizado en las denominadas Ciencias de la Tierra.

Antes de finalizar el epígrafe se ha de dejar constancia de un par de cuestiones:

a) En la resolución del sistema de ecuaciones (19) puede apreciarse que los ponderadores λ_i no dependen de los precios de las viviendas en las localizaciones en las que se dispone de información.

b) Resulta necesario disponer de una función de autocovarianzas, $C(h)$, siendo h la distancia (longitud y orientación) que separa las localizaciones. Esta última cuestión es fácilmente solventable por cuanto se puede acudir a cualquiera de los modelos teóricos de covariograma expuestos en la literatura geoestadística al uso⁹ (con la meseta, alcance y distancia integral adecuadas al área geográfica objeto de estimación) en caso de que se disponga de información acerca de la forma funcional que adopta la estructura de dependencia espacial. En caso contrario dicha función de autocovarianzas puede estimarse a partir de la expresión

$$C^*(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} (Z(s_i + h) - \bar{Z})(Z(s_i) - \bar{Z}) \quad [23]$$

siendo $N(h)$ el número de parejas de localizaciones a distancia h , ajustándose posteriormente el covariograma empírico a uno teórico.

4. ESTIMACIÓN KRIGEADA DEL PRECIO MEDIO DEL METRO CUADRADO DE LA VIVIENDA LIBRE: EL CASO DE TOLEDO CAPITAL

Con el objeto de poder comparar las estimaciones proporcionadas por la Subdirección General de Información y Estadística del Ministerio de Fomento en la publicación

⁸ Daniel Gerardus Krige planteó el uso de este tipo de técnicas para estimar la distribución del contenido de oro de determinados yacimientos sudafricanos. El desarrollo posterior de las mismas se debe a Georges Matheron y su grupo de investigación en la Escuela de Minas de París. Actualmente el Centro de Geoestadística de Fontainebleau es la referencia en cuanto a desarrollos y generalizaciones de las técnicas ideadas por Krige.

⁹ Una referencia muy completa a estos efectos es Chilès y Delfiner (1999).

“Índice de precios de las viviendas. Estadística de precio medio del m² (datos obtenidos de las tasaciones hipotecarias)” y por el estimador krigeado de la media propuesto en el epígrafe precedente, se ha utilizado la información relativa al precio de las viviendas tasadas, tanto nuevas como usadas, en Toledo capital, para una referencia temporal de 5 años (20 trimestres)¹⁰. El número total de viviendas utilizadas en la simulación fue de 2.986. De ellas, 1.180 fueron tasadas en 1995, 437 en 1996, 326 en 1997, 446 en 1998 y 597 en 1999. Por otra parte, 231 se ubican en el código postal 40001, 268 en el 45002, 286 en el 45003, 502 en el 45004, 579 en el 45005, 396 en el 45006, 436 en el 45007 y 88 en los códigos 45000, 45008 y 450009.

El precio y la ubicación de las viviendas fueron puestos a disposición de la unidad promotora por las siguientes entidades financieras de tasación:

- Tasaciones Inmobiliarias (TINSA).
- Banco Hipotecario.
- Tasaciones, Valoraciones y Servicios (TASER).
- Tasaciones Hipotecarias S.A. (excepto en la parte correspondiente al Banco Hipotecario).
- TASAMADRID.
- SERVATAS S.A.
- Sociedad de Tasación.
- Compañía Hispana de Tasaciones S.A. (COHISTASA).

El objetivo no es otro que comparar los resultados obtenidos por la mencionada Subdirección General con los del estimador krigeado propuesto. A estos efectos, hemos de poner de manifiesto que, debido al secreto estadístico, no se dispone de la ubicación exacta de las viviendas en los códigos postales en los que se ubican; no obstante, ello no es impedimento para llevar a cabo el ejercicio de simulación puesto que el resultado obtenido mediante la metodología ministerial sería exactamente el mismo fuese cual fuese la ubicación de los inmuebles. En consecuencia, la información disponible se dispuso de manera que cubriese la superficie de los códigos postales correspondientes (de tal forma que los resultados sean comparables con los de la estadística “Índice de precios de las viviendas. Estadística del precio medio del metro

¹⁰ En la actualidad la Dirección General de Programación Económica y Presupuestaria de la Subdirección General de Estadística y Estudios del Ministerio de Fomento no facilita la información pormenorizada relativa al precio medio del metro cuadrado de las viviendas y la adquisición de dicha información a las entidades tasadoras resulta económicamente inviable. Sin embargo, el autor dispone de dicha información desde 1987 (inicio de la serie) hasta 1999 para los municipios de Castilla-La Mancha como consecuencia de anteriores investigaciones. Esta es la razón por la cual la simulación se lleva a cabo hasta 1999. El hecho de que comience en 1995 radica en el escaso número de tasaciones registradas hasta 1994, lo cual llevaría a resultados escasamente representativos para el caso de Toledo capital.

cuadrado”), de acuerdo con el covariograma omnidireccional estimado en cada uno de los códigos postales¹¹. Dichos covariogramas se exponen en la Tabla 1.

Tabla 1

Código Postal	Covariograma omnidireccional		Alcance (metros)
45001	Esférico	$C(h) = m \left[1 - \frac{3}{2} \frac{h}{a} + \frac{1}{2} \left(\frac{h}{a} \right)^3 \right]$	1.150
45002	Esférico	$C(h) = m \left[1 - \frac{3}{2} \frac{h}{a} + \frac{1}{2} \left(\frac{h}{a} \right)^3 \right]$	1.300
45003	Gaussiano	$C(h) = m e^{-\frac{h^2}{a^2}}$	1.175
45004	Exponencial	$C(h) = m e^{-\frac{h}{a}}$	1.750
45005	Gaussiano	$C(h) = m e^{-\frac{h^2}{a^2}}$	1.800
45006	Esférico	$C(h) = m \left[1 - \frac{3}{2} \frac{h}{a} + \frac{1}{2} \left(\frac{h}{a} \right)^3 \right]$	1.250
45007	Gaussiano	$C(h) = m e^{-\frac{h^2}{a^2}}$	2.100

m : meseta; a : alcance; h : distancia.

¹¹ Aunque en Toledo capital existen 10 códigos postales, se han obviado el 45000, 45008 y 45009 debido al escasísimo número de tasaciones. De esta forma se evita el «efecto pedanía» puesto de manifiesto en Montero et al. (1998). Así mismo, y tras múltiples pruebas, se ha adoptado el supuesto de la no variabilidad del alcance en el tiempo, circunstancia razonable, por otra parte, dado que el periodo temporal considerado es muy corto.

A partir de los anteriores covariogramas se llevó a cabo la resolución de 140 sistemas de ecuaciones de krigado de la media (uno para cada código postal y trimestre de los cinco años de la simulación). La estimación de los ponderadores del estimador krigado se realizó mediante la programación de las rutinas pertinentes en lenguaje R, llevándose a cabo estimaciones puntuales en ubicaciones no pertenecientes al dominio mediante el software ISATIS elaborado por Geovariances y L'Ecole de Mines de Paris, con el objeto de analizar la consistencia de las soluciones.

Una vez obtenidas de la manera anteriormente expuesta las estimaciones de la media poblacional del precio del metro cuadrado de la vivienda libre para cada uno de los códigos postales en los veinte trimestres incluidos en la referencia temporal, se calcularon las estimaciones de las medias anuales en dos etapas:

- 1) Cálculo de la media espacio-ponderada de las estimaciones en los códigos postales para cada trimestre.
- 2) Cálculo de la media tempo-ponderada de las estimaciones trimestrales de cada año.

El anterior procedimiento, al utilizar como ponderadores el número de tasaciones, evita tanto el “efecto pedanía” (que tiende a infravalorar el precio medio del metro cuadrado) como las consecuencias derivadas de la estacionalidad en la actividad del mercado inmobiliario.

Los resultados trimestrales obtenidos y su comparación con los deducidos de la metodología empleada en la publicación “Índice de precios de las viviendas. Estadística del precio medio del m²” se exponen en la Tabla 2.

Tabla 2: Precio medio del metro cuadrado de la vivienda libre. Toledo capital.

Año	Trimestre	Estimación Krigeada	Estimación Ministerial	Diferencia
1995	1	651,62	595,70	55,92
	2	672,38	581,01	91,37
	3	582,66	584,97	-2,31
	4	586,61	566,73	19,88
1996	1	658,02	575,26	82,76
	2	654,77	603,87	50,90
	3	660,20	594,39	65,81
	4	604,80	552,47	52,33
1997	1	666,11	601,29	64,82
	2	605,87	632,34	-26,47
	3	652,63	571,53	81,10
	4	611,22	567,00	44,22
1998	1	622,87	538,41	84,46
	2	624,54	605,98	18,56
	3	644,34	625,85	18,49
	4	702,14	654,11	48,03
1999	1	656,33	646,66	9,67
	2	694,45	650,80	43,65
	3	743,81	664,68	79,13
	4	726,64	649,43	77,21

Fuente: Elaboración propia a partir de la información suministrada por la Dirección General de Programación Económica y Presupuestaria del Ministerio de Fomento.

A partir de la información trimestral contenida en la Tabla 2 se obtienen las estimaciones anuales del precio medio de la vivienda en Toledo capital, que se exponen en la Tabla 3 junto con las estimaciones proporcionadas por el Ministerio de Fomento.

Tabla 3: Precio medio del metro cuadrado de la vivienda libre. Toledo capital.

Año	Estimación Krigeada	Estimación Ministerial	Diferencia
1995	632,28	582,10	50,18
1996	647,64	577,73	69,91
1997	639,76	586,42	53,34
1998	650,67	625,14	25,43
1999	706,79	652,89	53,90

Fuente: Elaboración propia a partir de la información suministrada por la Dirección General de Programación Económica y Presupuestaria del Ministerio de Fomento

Como puede apreciarse, en la ejemplificación llevada a cabo la metodología ministerial arroja precios medios inferiores a los aportados por la metodología del krigeado de la media, como consecuencia de no tener en cuenta la correlación espacial existente entre los precios de las viviendas e incurrir en «efecto pedanía». En concreto, en la simulación realizada la infraestimación media es de 50,55 euros (casi un 10% del precio medio del metro cuadrado). En consecuencia, como puede apreciarse en la Figura 1, la evolución anual del precio del m² de la vivienda que proporciona la estimación ministerial es muy diferente a la proporcionada por la estimación krigeada. No obstante, éste es tan solo un ejemplo, una simulación llevada a cabo a partir de una disposición concreta de las viviendas sobre la superficie de los códigos postales. Lo importante es que la presencia de correlación espacial, y por tanto la estimación krigeada de la media, junto con el «efecto pedanía», nos alejan de los resultados proporcionados por el estimador ministerial (la media aritmética) y, en consecuencia, genera una serie de precios de vivienda distinta.

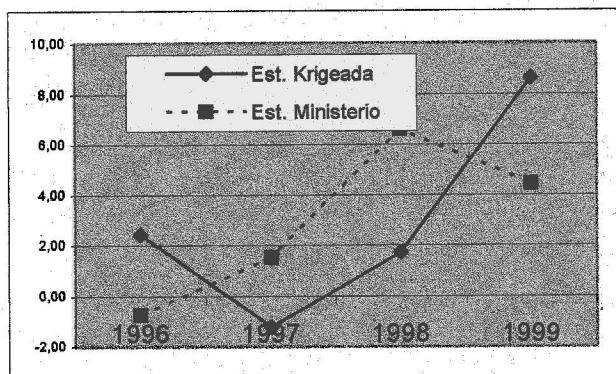


Figura 1: Incremento porcentual del precio medio del metro cuadrado de la vivienda libre en Toledo capital.

5. CONCLUSIONES

La determinación del precio medio de la vivienda y su evolución a lo largo del tiempo es una de las cuestiones que más preocupan a los ciudadanos y autoridades políticas y económicas de nuestro país. Sin embargo, pocos son los esfuerzos que se hacen para mejorar tanto las estadísticas como la metodología para llevar a cabo las estimaciones del mismo. Diversos autores han puesto de manifiesto la presencia de correlación espacial en los precios de la vivienda, razón por la cual la media aritmética no es un buen estimador de la media poblacional y se hace necesario el desarrollo de nuevos estimadores. En este trabajo se ha propuesto un estimador krigeado para la estimación del precio medio del metro cuadrado de la vivienda, estimador lineal insesgado y de varianza mínima en presencia de correlación espacial. Las estimaciones obtenidas a través de dicho estimador krigeado en la simulación realizada para Toledo capital en el periodo 1995-1999 difieren significativamente de las estimaciones ministeriales, tanto en nivel como evolución. En consecuencia, el no tener en cuenta la correlación espacial existente entre los precios de los bienes inmuebles, además del "efecto pedanía", puede llevarnos a graves distorsiones en una información de tanta relevancia económica y social como es la que se refiere a precios de vivienda.

Agradecimientos

Agradezco a Beatriz Larraz su inestimable ayuda en la parte computacional del artículo, y a Xavier Emery y Hans Wackernagel sus comentarios sobre las cuestiones particulares del mismo que les planteé.

REFERENCIAS BIBLIOGRÁFICAS

- CRESSIE, N. A. C. (1993): *Statistics for Spatial Data* (New York, John Wiley & Sons).
- CHICA, J. (1994): *Teoría de las Variables Regionalizadas. Aplicación en Economía Espacial y Valoración Inmobiliaria* (Granada, Universidad de Granada).
- CHILÈS, J.P.; DELFINER, P. (1999): *Geostatistics. Modelling Spatial Uncertainty* (New York, John Wiley).
- EMERY, X. (2000): *Geoestadística Lineal* (Santiago de Chile, Departamento de Ingeniería de Minas de la Universidad de Chile).
- FISCHER, M.M.; GETIS, A. (1997): *Recent Developments in Spatial Analysis* (Berlin, Springer).
- GÁMEZ, M. (1998): *Nuevas Técnicas de Estadística Espacial para la Economía: Modelización del precio de la vivienda libre en la ciudad de Albacete* (Universidad de Castilla-La Mancha, Tesis Doctoral).
- GÁMEZ, M.; MONTERO, J.M. (1999): *A spatial methodology for the estimation of housing price in Albacete*, Documentos de Trabajo de la Facultad de Ciencias Económicas y Empresariales de Albacete. UCLM, Doc. 2/1999/3.
- GÁMEZ, M.; MONTERO, J.M.; GARCÍA, N. (2000): *Kriging Methodology for Regional Economic Analysis*, *International Advances in Economic Research*, Vol.6, 3, pp. 438-450.
- GRANELLE, J.J. (1970): *Espacio urbano y precio del suelo* (Paris, Sirey).
- GRENANDER, U. (1954): *On the estimation of regression coefficients in the case of an autocorrelated disturbance*, *Annals of Mathematical Statistics*, 25, pp. 252-272.
- GRENANDER, U.; ROSEMBLATT, M. (1984): *Statistical Analysis of Stationary Time Series* (2ª ed.) (New York, Chelsea Publishing Co.).
- GUIGOU, J.L. (1982): *La rente foncière: les theories et leur evolution depuis 1650* (Paris, Economica).
- GUNST, R.F.; HARTFIELD, M.I. (1997): *Robust Semivariogram Estimation in the Presence of Influential Spatial Data Values*, in: Gregoire G. et al. (eds.): *Modeling Longitudinal and Spatially Correlated Data. Methods, Applications and Future Directions* (Berlín, Lecture Notes in Statistics, 122, Springer-Verlag).
- HAINING, R.P. (1988): *Estimating spatial means with an application to remotely sensed data*, *Communications in Statistics. Theory and Methods*, 17, pp. 573-597.
- JOURNEL, A.G.; HUIJBREGTS, C.J. (1978): *Mining Geostatistics* (London, Academic Press).
- 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, pp. 119 a 139.
- MATÉRN, B. (1986): *Spatial Variation* (2ª ed.) (Berlín, Lecture Notes in Statistics, 36, Springer-Verlag).

- MATHERON, G. (1965): Les variables régionalisées et leur estimation. Une application de la théorie des fonctions aleatoires aux sciences de la nature (París, Masson).
- MATHERON, G. (1971): La théorie des variables régionalisées et ses applications (Les Cahiers du Centre de Morphologie mathématique de Fontainebleau. Fascicule 5, ENMSP).
- MATHERON, G. (1989): Estimating and Choosing (Berlín, Springer-Verlag).
- MINISTERIO DE FOMENTO: Índice de Precios de las Viviendas. Estadística del precio medio del m². Datos obtenidos de tasaciones hipotecarias (Madrid, Dirección General de Programación Económica y Presupuestaria. Años 1995, 1996, 1997, 1998 y 1999).
- MINISTERIO DE FOMENTO (1995): Precio medio del m² de las viviendas (Madrid, Boletín Estadístico. Notas Metodológicas, 1, pp. 110 a 135)
- MONTERO, J.M.; BESCOS, J.; GÁMEZ, M. (1998): Consideraciones Metodológicas sobre la Configuración del Precio Medio del Metro Cuadrado de la Vivienda Libre, Actas de XII Reunión ASEPELT España, pp. 97, ETEA, Córdoba. (Versión completa disponible en CD-Rom).
- RICHARDSON, H.W. (1975): Economía del urbanismo (Madrid, Alianza Editorial).
- RICHARDSON, H.W. (1986): Economía regional y urbana (Madrid, Alianza Editorial).
- ROCA, J. (1988): La estructura de valores urbanos: un análisis teórico empírico (Madrid, Instituto de Estudios de Administración Local).
- SAMPER, F.J.; CARRERA, J. (1996): Geoestadística. Aplicaciones a la Hidrología Subterránea (2^a ed.) (Barcelona, Centro Internacional de Métodos Numéricos en Ingeniería).
- VENABLES, W.N.; RIPLEY, B.D. (1996): Modern Applied Statistics with S-Plus (3^a ed.) (New York, Springer Verlag).
- WACKERNAGEL, H. (2003): Multivariate Geostatistics. An Introduction with Applications (3^a Ed.) (Berlín, Springer-Verlag).