

# Análisis estadístico de datos geográficos en geomarketing: el programa GeoDa

■ CORO CHASCO YRIGOYEN

Departamento de Economía Aplicada. Universidad Autónoma de Madrid

---

**E**l análisis estadístico de datos geográficos es denominado en la literatura científica como análisis exploratorio de datos espaciales (AEDE). Se trata de una disciplina de la estadística que ha sido diseñada para el tratamiento específico de los datos espaciales o geográficos. Este análisis se utiliza para identificar relaciones sistemáticas entre variables de distribución geográfica cuando no existen expectativas claras sobre la naturaleza de estas relaciones. Esto último resulta muy habitual en el ámbito de los estudios del mercado geográfico (geomarketing), donde se suele trabajar con grandes bases de datos cuya estructura no siempre es bien conocida. En este artículo se presentan algunas técnicas del AEDE, que combinan el análisis estadístico con el gráfico, haciendo posible el estudio de las variables, su tendencia, valores extremos y esquemas de asociación-dependencia espacial y concentración espacial o pun-

tos calientes/fríos de negocio ("hot spots"). Actualmente, existe en el mercado un nuevo programa informático, GeoDa, concebido como un producto "amigable", que no requiere del manejo de un GIS, y funciona en cualquier sistema operativo. Este programa ha sido desarrollado por el profesor Luc Anselin, de la Universidad de Illinois, y tiene la ventaja de ser, hasta el momento, un producto de libre acceso en Internet.

Es un hecho indiscutido que la información estadística constituye la base de todo estudio de mercado. Pero las bases de datos, muchas veces de gran tamaño, deben ser procesadas y analizadas convenientemente para extraer de ellas las conclusiones que hagan posible la toma de decisiones. En la actualidad, las técnicas estadísticas, desde la estadística descriptiva a los modernos métodos de minería de datos, se encuentran más fácilmente al alcance de los analistas, gracias

sobre todo al desarrollo, abaratamiento y "amigabilidad" de los programas informáticos. Sin embargo, en lo que se refiere a su conocimiento y aplicación, el análisis estadístico-informático de los datos geográficos –espaciales– se encuentra a gran distancia del desarrollo experimentalizado por el análisis de series temporales.

Este artículo pretende llamar la atención de los investigadores sociales sobre la importancia de llevar a cabo un adecuado análisis estadístico de los datos geográficos. En efecto, aunque la investigación de mercados suele trabajar con información de carácter espacial (referida a unidades geográficas), en pocas ocasiones estos datos suelen ser tratados de forma específica y diferente del análisis de series temporales (o de corte-transversal no espacial), utilizando técnicas adecuadas para el análisis estadístico-gráfico-cartográfico. Estas herramientas han sido bautizadas con el nombre de análisis

exploratorio de datos espaciales (AEDE) y se conciben como una disciplina dentro del más general análisis estadístico, diseñada para el tratamiento específico de los datos geográficos. El AEDE se utiliza para identificar relaciones sistemáticas entre variables, o dentro de una misma variable, cuando no existe un conocimiento claro sobre su distribución en el espacio geográfico.

Esto último resulta muy habitual en el ámbito de los estudios del mercado geográfico (geomarketing), donde se suele trabajar con grandes bases datos cuya estructura (tendencia, zonas calientes, puntos atípicos, etc.) no es siempre bien conocida. El geomarketing es una disciplina con carácter propio, cada vez más mencionada y conocida en la literatura científica (García, 1997; Moreno, 2001; Latour y Le Floc'H, 2001; Rosa, 2001; Völle, 2001; Chasco, 2003B, Gijsbrechts et al., 2003). El geomarketing podría definirse como el análisis geográfico de la realidad económico-social a través de instrumentos cartográficos y herramientas de la estadística espacial. La utilización de esta técnica en las estrategias de marke-

ting de las empresas comerciales se explica, sobre todo, por el aumento de la competencia, que exige una expansión continua de los negocios, lo que obliga a instalar nuevos puntos de venta y, por tanto, a elegir nuevas localizaciones. Además, la saturación de los establecimientos comerciales en las localizaciones más favorables dificulta la toma de decisiones respecto a futuras ubicaciones (centro de la ciudad, periferia, centros comerciales, localidades de mediano tamaño, etc.).

Con este artículo, se pretende llamar la atención sobre las capacidades del análisis exploratorio de datos espaciales (AEDE) como herramienta al servicio del geomarketing. Además, se presentará un "software" creado para el desarrollo del AEDE, GeoDa, concebido como un producto "amigable", que no requiere de la utilización de un sistema de información geográfica (GIS) y funciona en cualquier sistema operativo. GeoDa ha sido desarrollado por el profesor Luc Anselin, de la Universidad de Illinois, y tiene la ventaja de ser, hasta el momento, un producto gratuito, de libre acceso en Internet. A

continuación se presentan algunas técnicas del AEDE, aplicadas con el programa GeoDa, a algunos ejemplos ilustrativos.

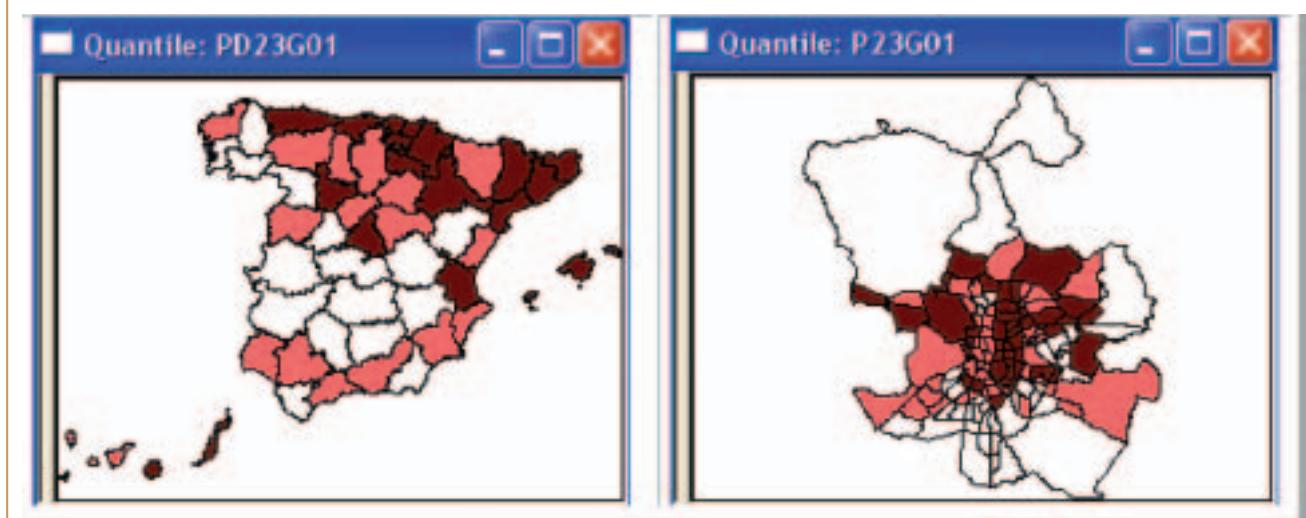
### EL ANÁLISIS EXPLORATORIO DE DATOS ESPACIALES (AEDE) Y GEODA

El AEDE ha sido definido como el conjunto de técnicas que describen y visualizan las distribuciones espaciales, identifican localizaciones atípicas o "atípicas espaciales" ("spatial outliers"), descubren esquemas de asociación espacial, agrupamientos ("clusters") o puntos calientes ("hot spots") y sugieren estructuras espaciales u otras formas de heterogeneidad espacial (Anselin, 1999). Por tanto, el AEDE se correspondería con los métodos de estadística descriptiva espacial. En los últimos años, el AEDE ha sido introducido en algunas aplicaciones de economía regional (por ejemplo, Moreno y Vayá, 2000), así como en los ejercicios de predicción-extrapolación de datos (Chasco, 2003A), pero nunca en el marketing.

Sin embargo, las técnicas del AEDE son muy eficaces en situaciones en las que no existe un marco formal o teoría previa acerca del fenómeno que se anali-

GRÁFICO N° 1

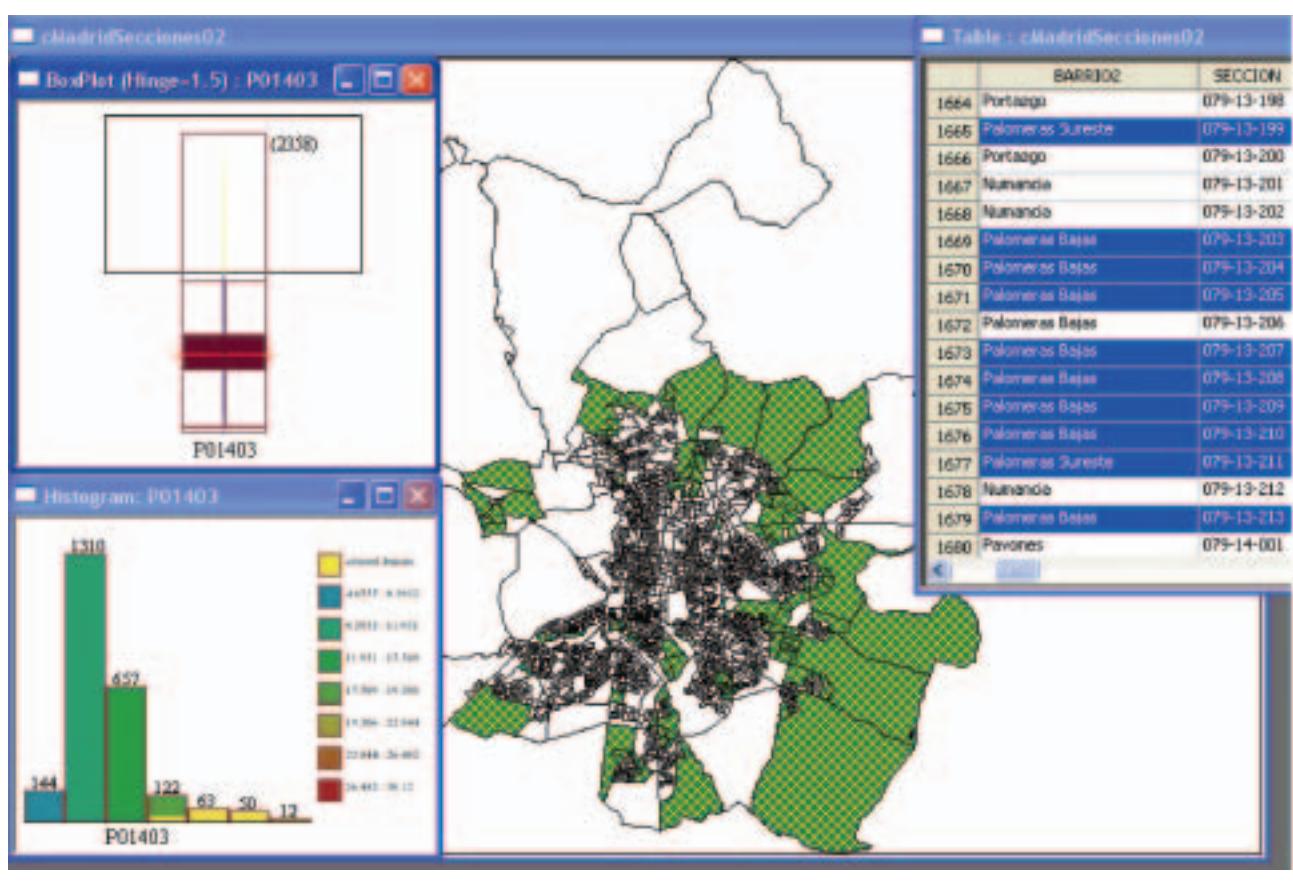
#### DISTRIBUCIÓN DE LA TASA DE INSTRUCCIÓN SUPERIOR\* EN LAS PROVINCIAS (IZDA.) Y BARRIOS DEL MUNICIPIO DE MADRID (DCHA.)



\* Número de personas con título universitario y de bachiller sobre la población de 16 y más años en 2001 (INE, Censo de Población, 2001).

GRÁFICO N° 2

**CONEXIÓN DINÁMICA DE DIVERSAS VISTAS (MAPA, DIAGRAMA DE CAJA, HISTOGRAMA DE FRECUENCIAS, TABLA DE DATOS) EN GEODA\***



\* Variable índice de infancia: personas menores de 15 años sobre población total en 2003 (INE, Padrón de Habitantes, 2003).

za. Estas situaciones se plantean muy a menudo en el campo de las ciencias sociales, cuando se analizan grandes bases de datos geográficos cuya distribución no tiene por qué ser conocida a priori. Por ejemplo, parece obvio que la distribución del nivel de instrucción en las provincias españolas sigue una tendencia general norte-sur (mayor nivel-menor nivel, respectivamente), tal y como se deriva del mapa de cuantiles (izquierdo) del gráfico n° 1. Sin embargo, esta distribución suele ser más desconocida cuando se analiza este fenómeno para ámbitos territoriales más desagregados (distritos, barrios, secciones censales), donde la tendencia espacial puede variar o simplemente no existir. Así, en el ejemplo

propuesto, la misma tasa de instrucción superior parece seguir una tendencia del centro a la periferia, en los barrios del municipio de Madrid.

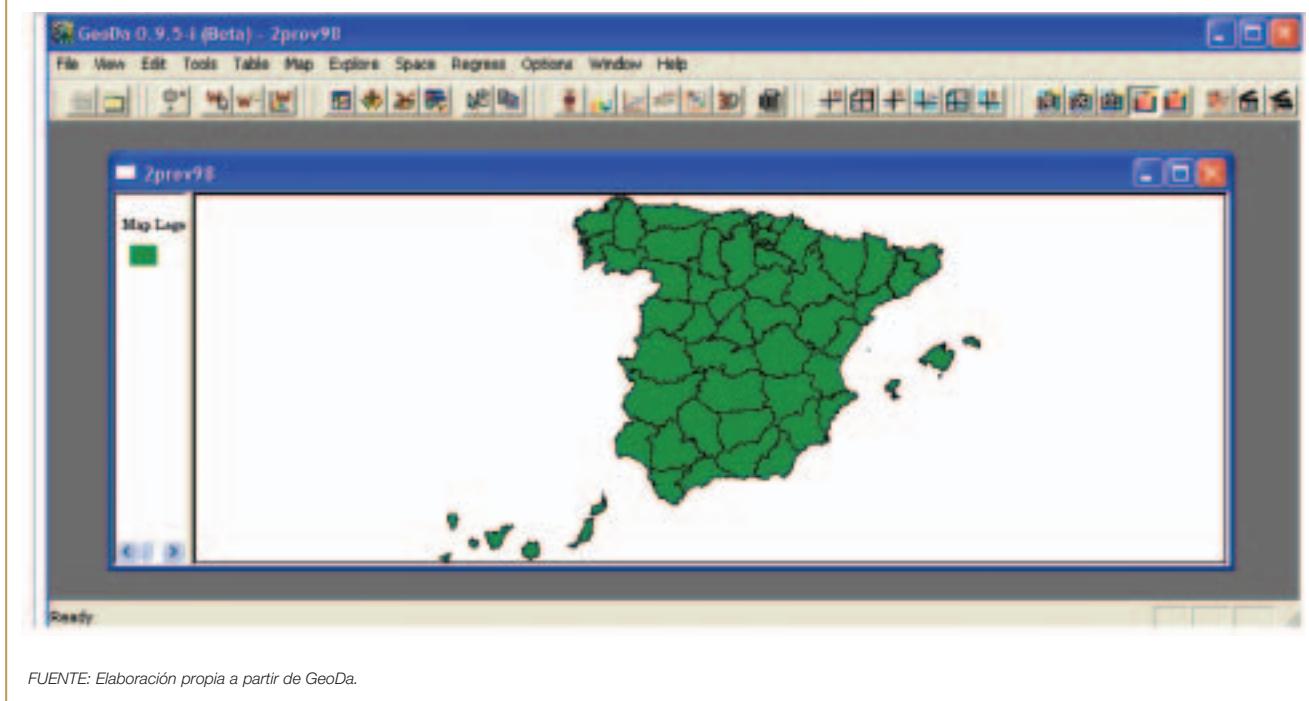
El AEDE combina el análisis estadístico con el gráfico, dando lugar a lo que podría denominarse una "visualización científica" (Haining et al., 2000) que, a los métodos más sencillos de estadística descriptiva (mapas de cuantiles, percentiles, etc.), une contrastes más complejos sobre los efectos espaciales de dependencia y heterogeneidad. Los programas informáticos propios del AEDE, como GeoDa, son además capaces de unir las distintas "vistas" (gráficos, tablas, mapas, etc.) de forma dinámica (Unwin, 2000). En el gráfico n° 2, se ilus-

tra esta potencialidad utilizando la distribución del índice de infancia (número de personas menores de 15 años sobre el total de la población) en las secciones censales del municipio de Madrid. En el diagrama de caja (superior izquierdo) se han seleccionado las secciones censales con valores atípicamente altos en el índice. Estos valores quedan automáticamente seleccionados en el resto de vistas: histograma de frecuencias (inferior izquierdo), mapa (fondo) y tabla de datos (derecha).

GeoDa es muy fácil de utilizar de forma visual e interactiva y su uso no exige un conocimiento previo de la tecnología GIS, mucho más compleja. En este artículo, nos valemos de este programa que ha si-

GRÁFICO N° 3

**GEODA: VISTA PRINCIPAL CON FUNCIONES DEL MENÚ Y BARRA DE HERRAMIENTAS**



FUENTE: Elaboración propia a partir de GeoDa.

desarrollado por el profesor Luc Anselin, de la Universidad de Illinois, para presentar la capacidad y posibilidades del AEDE. La versión más reciente del programa, 0.9.5-i5, data de agosto de 2004 y está siendo muy bien recibida por todo tipo de usuarios, especialmente los dedicados al mundo académico y de la investigación, en general (en septiembre de 2004, GeoDa tenía registrados a 3.500 usuarios, que se incrementan en una tasa de 150 nuevos cada mes). Esta versión está disponible en Internet en la dirección electrónica [http://sal.agecon.uiuc.edu/geo\\_da\\_main.php](http://sal.agecon.uiuc.edu/geo_da_main.php) y es de acceso libre. En esta página web hay diversos materiales de apoyo así como una película del programa "Quicktime" en la que se hace una demostración de las características principales de este programa. La versión actual de GeoDa utiliza la tecnología MapObjects LT2 de ESRI, que permite el acceso a datos geográficos, su representación geográfica y la obtención de nuevas variables mediante funciones de "query". Por este motivo, el único formato aceptado

para la cartografía digital es el proporcionado por ESRI, con la extensión "shp" ("shapefile").

En términos generales, las diferentes funciones de GeoDa podrían ser clasificadas en 6 categorías: tratamiento de datos geográficos, transformación de datos, representación gráfica en mapas, gráficos estadísticos, dependencia espacial y regresión espacial. Como se aprecia en el gráfico nº 3, estas funciones pueden ser activadas a través del menú superior o directamente, tecleando en los iconos de la barra de herramientas. En cada una de las secciones principales existen varias subaplicaciones relacionadas.

En una variable geográfica pueden identificarse tres elementos: tendencia, asociación-dependencia espacial y concentración-atípicos espaciales. Cada uno de ellos tiene sus propias técnicas de AEDE que permiten su estudio, como se expone a continuación. Para ello, se proponen algunos ejemplos ilustrativos de las capacidades de estas técnicas para el análisis socioeconómico. Los interesa-

dos en un análisis más en detalle de estas herramientas pueden consultar Anselin (2003, 2004) y Anselin et al. (2004A, 2004B).

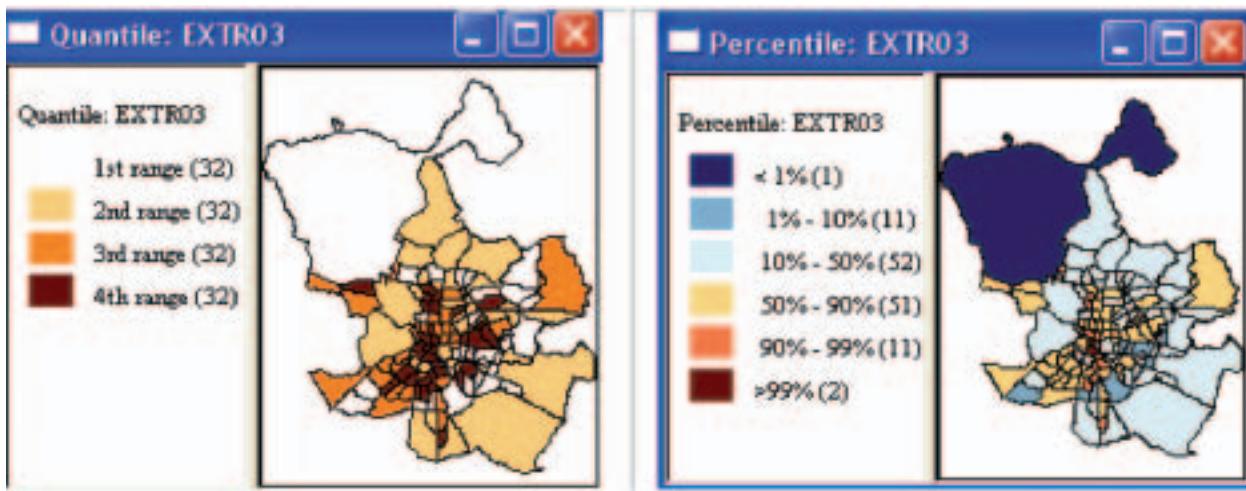
**MÉTODOS DE REPRESENTACIÓN DE LA TENDENCIA GEOGRÁFICA**

Dada la naturaleza multidireccional de los datos geográficos, la tendencia espacial no se puede representar mediante una línea recta (como en las series temporales) sino a través de un mapa o gráfico capaces de expresar la evolución de la variable en las distintas áreas o unidades geográficas (provincias, municipios, etc.). Por eso, las técnicas del AEDE que analizan este elemento tendencial son herramientas de representación cartográfica para las que lo fundamental no es tanto el mapa en sí mismo cuanto la representación gráfica de los estadísticos básicos.

Los métodos de análisis de la tendencia espacial más conocidos son los mapas temáticos, que consisten en la representación cartográfica (en un mapa digi-

GRÁFICO N° 4

## MAPAS DE CUANTILES: CUARTILES (IZDA.) Y PERCENTILES (DCHA.)\*



\* EXTR03 es la tasa de extranjeros, personas de nacionalidad no española, sobre total población en 2003 (INE, Padrón de Habitantes, 2003).

tal) de una variable geográfica. Esta representación en un mapa de la variable puede llevarse a cabo mediante símbolos y colores que pongan de manifiesto el valor de una variable en cada una de las unidades geográficas consideradas (países, regiones, etc.). Puede utilizarse un color/símbolo diferente para cada valor o para cada intervalo de valores de la variable. Hay dos tipos de mapas temáticos útiles para la representación de la tendencia espacial de una variable: los mapas de cuantiles y el mapa-diagrama de caja.

En los mapas de cuantiles, los datos se dividen y agrupan en una serie de categorías (cuantiles) con igual número de observaciones o unidades territoriales. Así, si la distribución se divide en 4 grupos, se tratará de un mapa de cuartiles, si se divide en 5 grupos, será un mapa de quintiles, y así sucesivamente. En el ejemplo del gráfico n° 4, el número de categorías se ha establecido en 4, para construir un mapa de cuartiles de la variable tasa de extranjeros para el conjunto de barrios del municipio de Madrid. Como puede observarse, en la leyenda se indica, entre paréntesis, el número de observaciones de cada categoría, 32. De la ob-

servación de este mapa puede deducirse que los valores más elevados de esta variable (4º cuartil) se localizan, por lo general, en barrios del centro-sur de la ciudad (distritos Centro, Arganzuela, Carabanchel, Usera, Puente de Vallecas), aunque se producen algunas discontinuidades espaciales motivadas por un número importante de barrios del centro urbano en los que la tasa de extranjeros es baja (1º cuartil).

Debe advertirse que este tipo de mapa no resulta de utilidad en casos en los que la variable en cuestión contenga un gran número de observaciones con valores parecidos (por ejemplo, en el caso de variables de sucesos raros, cuando muchas observaciones tienen valor cero). El motivo es obvio: muchos cuantiles no podrán ser definidos al no poder asignar un mismo número de observaciones a los diferentes grupos.

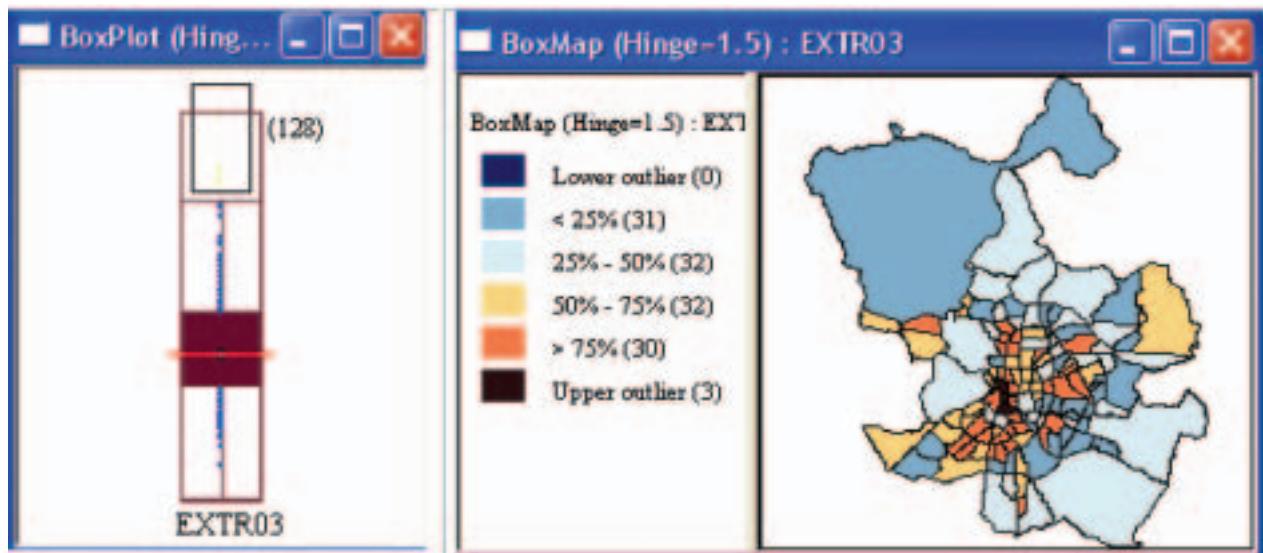
El mapa de percentiles es un caso particular del mapa de cuantiles que permite la identificación de los valores extremos de una distribución espacial. Para ello, las categorías se diseñan de modo que queden acentuados los valores más altos/bajos de una variable. En concreto, el programa GeoDa crea 6 categorías co-

rrespondientes a los siguientes percentiles (1): <1, [1,10), [10,50), [50,90), [90,99), >99. En el gráfico n° 4, se presenta el mapa de percentiles de la tasa de extranjeros, que identifica tres barrios con valores extremos: Sol y Embajadores (distrito Centro) con tasas especialmente elevadas, y El Pardo, el barrio con la tasa de extranjeros más baja de la capital.

Otra herramienta para el análisis de la tendencia de una variable geográfica es el diagrama-mapa de caja. El gráfico de caja constituye un método estándar de representación diseñado para todo tipo de variables (espaciales y no espaciales). Su construcción parte del cálculo de los cuartiles y la media de una variable, así como de la obtención de las llamadas cotas o valores adyacentes superior e inferior, que se obtienen, a su vez, como el producto de los valores del tercer (primer) cuartil por 1,5 veces el recorrido intercuartílico (2). De esta forma, se consideran como valores atípicos aquellos situados por encima (o por debajo) de dichas cotas (en el gráfico 5 se trata del diagrama de la variable tasa de extranjeros). En el mapa de caja se representan con distintos colores las unidades geográficas cuyos datos coinciden con la mediana, el

GRÁFICO N° 5

**GRÁFICOS DE CAJA DE LA TASA DE EXTRANJERÍA: DIAGRAMA (IZDA.) Y MAPA (DCHA.)**



\* EXTR03 es la tasa de extranjeros personas, de nacionalidad no española, sobre total población en 2003 (INE, Padrón de Habitantes, 2003).

recorrido intercuartílico y los valores extremos (altos y bajos).

En el ejemplo del gráfico nº 5 se han representado el diagrama y el mapa de caja de la tasa de extranjería de los barrios de Madrid. En este caso, sólo se han identificado como valores extremos tres barrios, que destacan por sus elevadas tasas de extranjeros: Centro, Embajadores y Universidad. Sin embargo, no existe ningún barrio en el otro extremo de la distribución caracterizado por tener unas tasas de extranjeros equivalentes.

Los métodos específicos del AEDE pueden complementarse con otras técnicas no específicamente diseñadas para el análisis de datos geográficos. Este es el caso del histograma de frecuencias, además del diagrama de caja. GeoDa calcula histogramas de frecuencias de las variables geográficas para distintas clasificaciones, aunque el número por defecto es 7. Cada una de las barras del histograma tiene un color y es posible realizar una selección en el histograma para ver sobre el mapa las observaciones a las que corresponde. Esto es lo que sucede en el gráfico nº 5, en donde se han seleccionado

las cuatro barras del histograma de frecuencias con mayor tasa de extranjería en 2003 (en amarillo), de forma que quedan destacados en el mapa los barrios de Madrid a los que corresponde (barrios del interior del municipio, por lo general).

Además de los presentados, GeoDa incluye otras herramientas del AEDE para la representación de tendencias espaciales (mapa dinámico, mapa de la desviación típica, cartograma...), que ofrecen información coincidente o, en todo caso, complementaria a la que se desprende de las representaciones anteriores.

#### **MÉTODOS DE REPRESENTACIÓN DE LA ASOCIACIÓN O DEPENDENCIA ESPACIAL**

La asociación espacial es, junto con la tendencia y la concentración espacial, un elemento caracterizador de las distribuciones espaciales. En concreto, la asociación espacial consiste en la coincidencia de valores (altos/bajos) de una variable en determinados lugares del espacio geográfico. Esta propiedad es también conocida con el nombre técnico de dependencia o autocorrelación espacial.

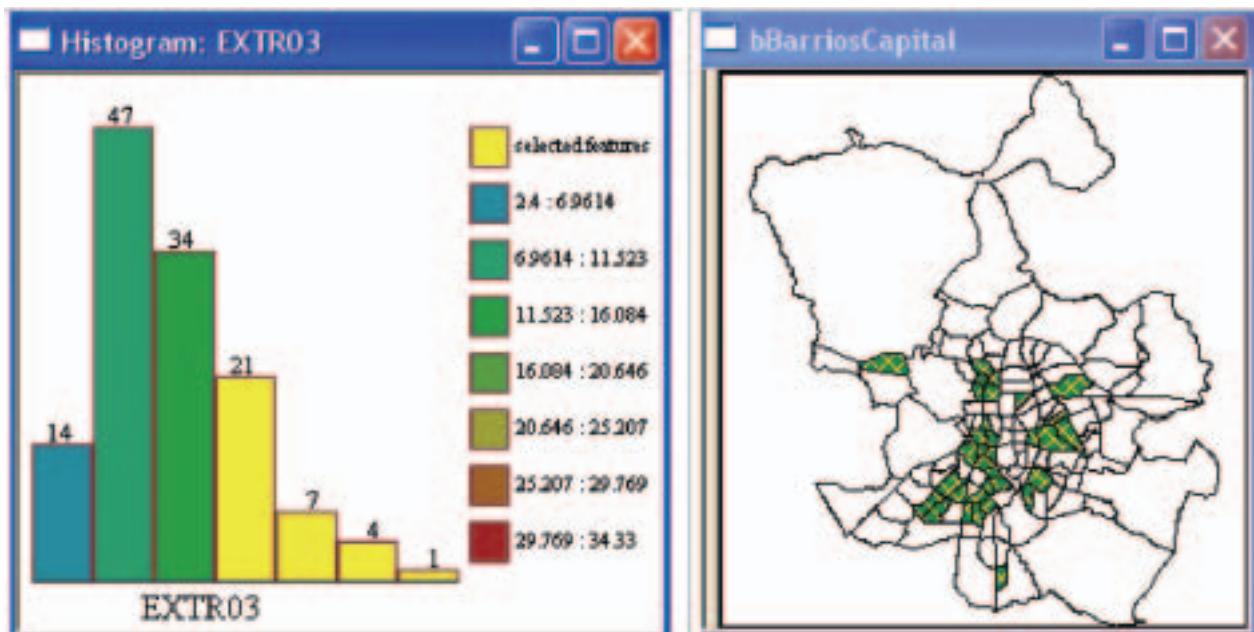
El fenómeno de asociación espacial en una variable puede expresarse matemáticamente como una relación funcional entre lo que ocurre en un punto determinado del espacio y lo que sucede en lugares cercanos –vecinos– al mismo. Es decir, en una variable se produce el fenómeno de la asociación o dependencia espacial cuando los valores observados en una observación (región, barrio, etc.) dependen de los valores observados en regiones vecinas. En este caso, la distribución de la variable sobre un mapa seguirá una continuidad geográfica.

Una de las más conocidas técnicas de AEDE de asociación espacial es el diagrama de dispersión de Moran. Se trata de un diagrama de dispersión que representa en el eje X la variable previamente estandarizada y en el eje Y se representa la variable espacialmente retardada de dicha variable estandarizada.

La variable espacialmente retardada de una dada o “retardo espacial” es un promedio ponderado de los valores que adopta una variable en el subconjunto de observaciones vecinas a una dada. Por ejemplo, el retardo espacial de la ta-

GRÁFICO N° 6

**HISTOGRAMA DE FRECUENCIAS DE LA TASA DE EXTRANJERÍA EN LOS BARRIOS DE MADRID**



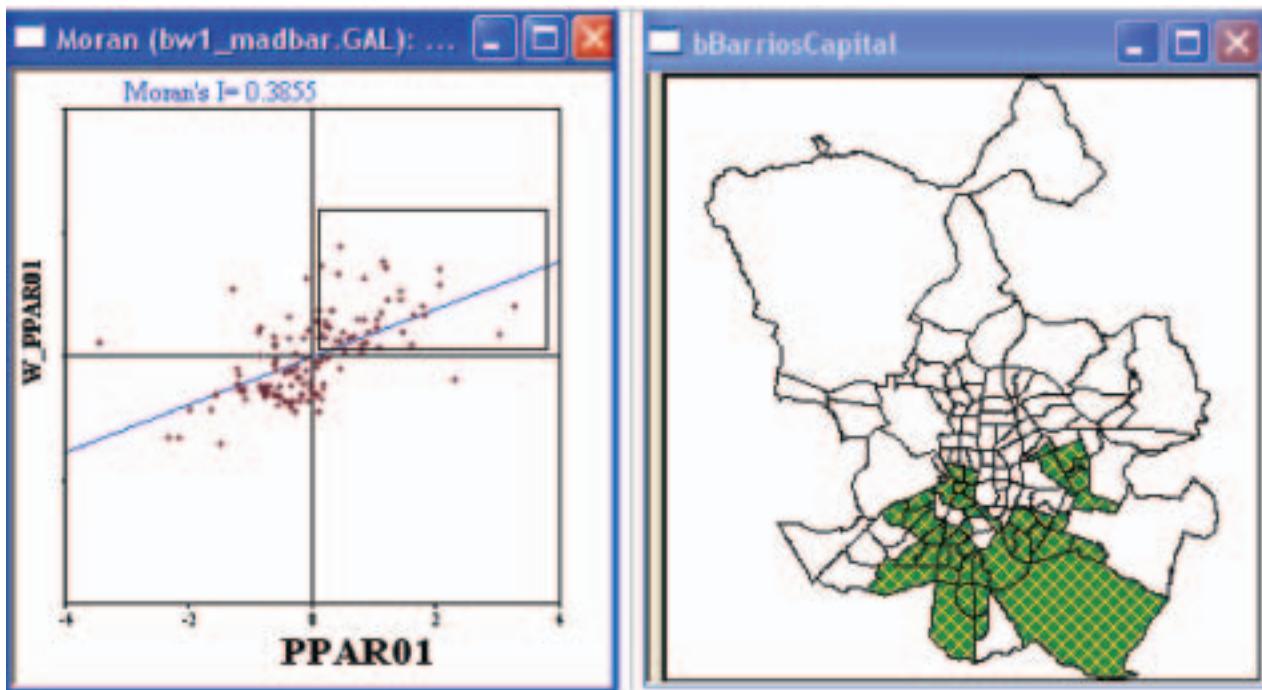
\* EXTR03 es la tasa de extranjeros, personas de nacionalidad no española, sobre total población en 2003 (INE, Padrón de Habitantes, 2003).

sa de paro de la provincia de Madrid sería la media aritmética de los valores de tasa de paro en las provincias limítrofes (Segovia, Ávila, Toledo, Cuenca y Guadalajara).

El diagrama de dispersión de Moran permite distinguir los dos signos propios del fenómeno de asociación espacial: asociación espacial positiva y negativa. El fenómeno de asociación o dependencia espacial positiva se produce en aquellas observaciones en las que una variable adopta valores altos/bajos que, al mismo tiempo, se encuentran rodeadas de observaciones con valores altos/bajos en dicha variable. Se trata, pues, de una coincidencia de valores de una variable en una zona determinada del espacio geográfico. Por su parte, la asociación o autocorrelación espacial negativa consiste en la coincidencia, en una zona, de valores altos de una variable rodeados por valores bajos de la misma, y viceversa). Las categorías de asociación espacial positiva se corresponden con los cuadran-



GRÁFICO N° 7

DIAGRAMA DE DISPERSIÓN DE MORAN Y MAPA CON SELECCIÓN DEL 1<sup>er</sup> CUADRANTE

\* PPAR01 es la variable estandarizada de la tasa de paro en 2001 (INE, Censo de Población, 2001). W\_PPAR01 es la variable espacialmente retardada de la tasa de paro estandarizada.

tes I y III, mientras que el fenómeno de asociación espacial negativa se encuentra en los cuadrantes II y IV.

Por ejemplo, en el diagrama de dispersión de Moran del gráfico n° 7 se han seleccionado las observaciones del cuadrante I, es decir, aquellos barrios de Madrid que con una tasa de paro superior a la media municipal (valor estandarizado mayor que cero) se encuentran rodeados por barrios que también disponen de tasas de paro superiores a la media municipal (valor de la variable espacialmente retardada mayor que cero). Estos barrios aparecen destacados en el mapa de la derecha y, como puede observarse, se localizan fundamentalmente en la zona sur del municipio de Madrid. Por su parte, en el cuadrante III se encuentra el fenómeno opuesto: barrios con valores bajos de tasa de paro rodeados de barrios con valores bajos.

Por el contrario, las categorías de aso-

ciación negativa vienen dadas por los cuadrantes II y IV de este diagrama, en los que se representan los barrios con valores bajos/altos de tasa de paro rodeados por barrios con valores altos/bajos de dicha variable, respectivamente. La asociación espacial negativa se produce de forma extrema en variables geográficas que adoptan una estructura idéntica a la que se produce en un tablero de ajedrez, donde las celdas blanca-negra se encuentran repartidas de forma totalmente alternante. Evidentemente, esta distribución espacial no suele producirse en ninguno de los sucesos económicos-sociales.

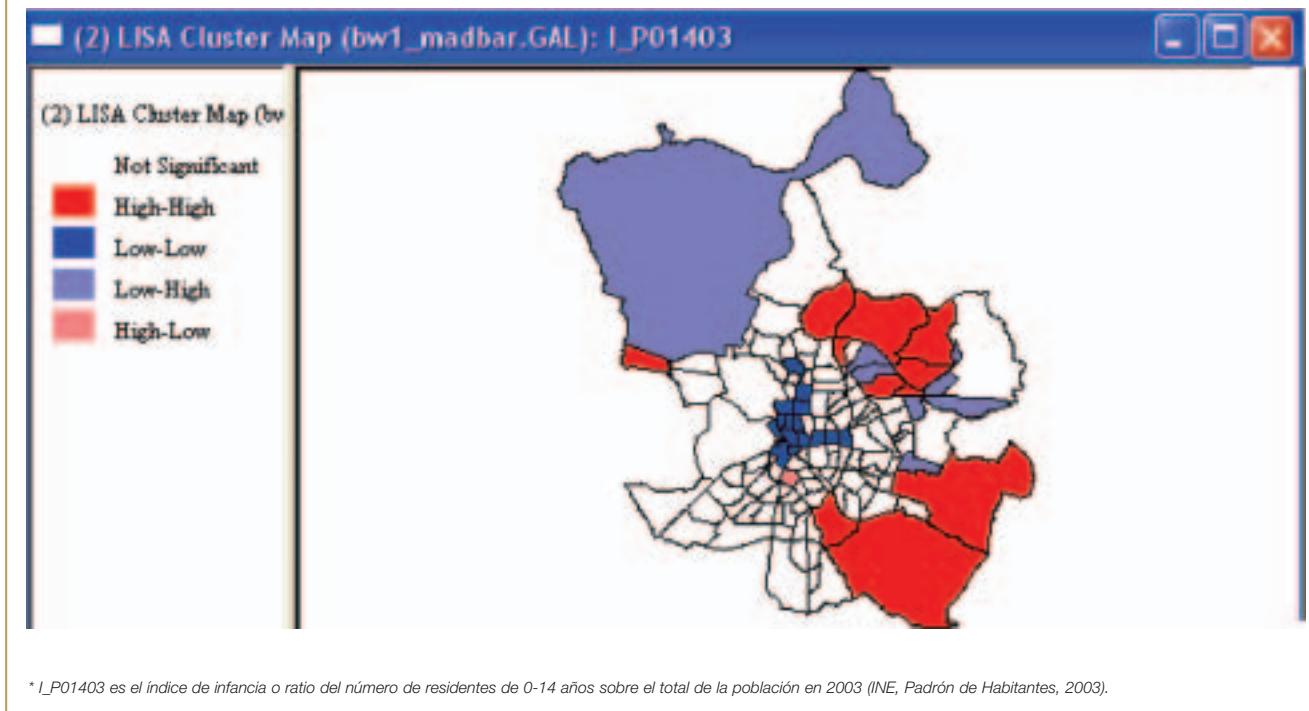
En este diagrama de dispersión, en el que se relacionan el valor de la variable en una observación geográfica y el valor promedio de dicha variable en las observaciones vecinas, la pendiente de la recta de regresión es coincidente con el valor del estadístico I de Moran de autocorrelación es-

pacial global (3). Cuanto mayor sea el valor de este estadístico, es decir, cuanto mayor sea el ángulo que forme la recta de regresión con el eje de abscisas, más fuerte será el grado de asociación o dependencia espacial en la variable, y viceversa. En el ejemplo del gráfico n° 7 el valor de la pendiente de la recta de regresión es 0,3855 que coincide con el valor del estadístico I de Moran. Este coeficiente expresa la intensidad de la relación causal existente entre los valores de una variable en una unidad territorial y los valores de la misma en las unidades vecinas. Es decir, en qué medida un cambio operado en los niveles de una variable en un lugar del espacio, por ejemplo, un barrio de Madrid, afectan a los niveles de dicha variable en los lugares vecinos al mismo.

Es por eso que este fenómeno de asociación, dependencia o autocorrelación espacial es también conocido como "interacción espacial", dado que expresa

GRÁFICO N° 8

**MAPA LISA DE CONCENTRACIÓN ESPACIAL**



las relaciones de interrelación existentes en los fenómenos, sobre todo aquéllos de carácter socio-económico, sobre el espacio geográfico.

**MÉTODOS DE REPRESENTACIÓN DE LA CONCENTRACIÓN ESPACIAL: ZONAS CALIENTES/FRÍAS**

El fenómeno de concentración espacial tiene lugar cuando se produce una especial concentración de valores extremos de una variable en torno a una o varias unidades territoriales cercanas en el espacio. Estas zonas de especial concentración de valores extremos de una variable se conocen también como zonas calientes/frías, según se trate de una concentración de valores especialmente altos/bajos de una variable, respectivamente. También se las conoce como "bolas" o agrupamientos ("clusters") espaciales.

Los gráficos LISA ("Local Indicator of Spatial Association") de concentración o dependencia espacial local se han dise-



ñado con el objetivo de determinar la existencia de concentración (asociación espacial local) espacial en una variable, así como el grado de significación estadística con la que se produce.

En concreto, los gráficos LISA incluidos en GeoDa se basan en el estadístico I local de Moran, que se trata de un estadístico que, a diferencia del estadístico I de Moran, no se calcula de forma global para todas las observaciones del mapa, sino que adquiere un valor diferente para todas y cada una de ellas. El estadístico I local de Moran mide el grado de concentración de valores extremos de una variable en el entorno geográfico de cada una de las observaciones de la muestra. Para cada valor del estadístico es posible realizar una inferencia para evaluar el nivel de significatividad estadística de rechazo de la hipótesis nula de ausencia de similitud o disimilitud de valores extremos de una variable en torno a una localización geográfica. De este modo, se pone de manifiesto la presencia de puntos calientes ("hot spots") o atípicos espaciales, cuya mayor o menor intensidad dependerá de la significatividad asociada de los citados estadísticos.

El mapa LISA es un mapa en el que se destacan con colores calientes/fríos aquellas observaciones en torno a las cuales se produce una concentración, estadísticamente significativa, de valores especialmente altos/bajos de una variable. En el gráfico nº 8 se presenta el mapa LISA del índice de infancia en los barrios de Madrid, que se trata de un indicador interesante para negocios dirigidos al público infantil. Como puede observarse, este mapa representa las "zonas calientes" de concentración de barrios con alto valor (color rojo) y bajo valor (azul) en el índice de infancia. La zona caliente de barrios con una significativa concentración de valores altos del índice de infancia está formada por 11 barrios de la periferia este de Madrid fundamentalmente. El fenómeno opuesto, es decir, la concentración de bajos niveles en el índice de infancia, se produce en 15 barrios del centro de la ciudad, con una población más envejecida.

### CONCLUSIÓN

El AEDE debe constituir la etapa previa a todo análisis del mercado destinado a la toma de decisiones en el campo de la investigación socioeconómica. En este artículo se han presentado las principales técnicas del AEDE, que combinan el análisis estadístico con el gráfico. Para ello, se ha utilizado un nuevo programa informático, GeoDa, que tiene como una de sus grandes ventajas el haber sido concebido como un producto "amigable", que no supone para el usuario el conocimiento de un sistema específico de GIS, funciona en cualquiera de los sistemas operativos existentes y es actualmente un producto gratuito de libre acceso en Internet. GeoDa es un programa que combina técnicas del análisis estadístico convencional con las más novedosas herramientas del AEDE, en un entorno dinámico que permite conexiones dinámicas entre las distintas vistas (gráficos, mapas, tablas, etc.).

Efectivamente, el análisis de las series geográficas requiere de herramientas propias, que van más allá de las convencionales técnicas estadísticas y, por tanto,

de un "software" específico. Estas herramientas deben estar dirigidas al análisis de tres elementos fundamentales: tendencia espacial y valores extremos, asociación o dependencia espacial y concentración espacial. Esta última cualidad de las distribuciones espaciales debe entenderse como una concentración significati-

va de valores extremos (altos/bajos) de una variable en torno a una unidad geográfica, lo que permite determinar zonas calientes/frías de negocio. ■

**CORO CHASCO YRIGOYEN**

Departamento de Economía Aplicada  
Universidad Autónoma de Madrid

### NOTAS

- (1) El paréntesis cuadrado indica que ese extremo se incluye en el intervalo y el curvo indica que no se incluye.
- (2) Un criterio un poco más estricto que el anterior consistiría en multiplicar por tres el recorrido intercuartílico para la fijación de las cotas.
- (3) Una revisión más extensa de las medidas estadísticas de autocorrelación espacial se encuentra, entre otras referencias, en Chasco (2003).

### BIBLIOGRAFÍA

- ANSELIN, L. (1999), "The future of spatial analysis in the social sciences". *Geographic Information Sciences*, 5 (2); pp. 67-76.
- ANSELIN, L. (2003), "GeoDaTM 0.9 user's guide". Página web del "Spatial Analysis Laboratory": [http://sal.agecon.uiuc.edu/stuff\\_main.php#tutorials](http://sal.agecon.uiuc.edu/stuff_main.php#tutorials) .
- ANSELIN, L. (2004), "GeoDaTM 0.9.5-i release notes". Página web del "Spatial Analysis Laboratory": [http://sal.agecon.uiuc.edu/stuff\\_main.php#tutorials](http://sal.agecon.uiuc.edu/stuff_main.php#tutorials) .
- ANSELIN, L.; Y.W. KIM y I. SYABRI (2004A), "Web-based analytical tools for the exploration of spatial data". *Journal of Geographical Systems* (próxima publicación).
- ANSELIN, L., I. SYABRI y Y. KHO (2004B), "GeoDa: An introduction to spatial data analysis". *Geographical Analysis* (próxima publicación).
- CHASCO, C. (2003A), "Econometría espacial aplicada a la predicción-extrapolación de datos microterritoriales". Consejería de Economía e Innovación Tecnológica de la Comunidad de Madrid (edición electrónica).
- CHASCO, C. (2003B), "El geomarketing y la distribución comercial". *Investigación y Marketing* 79, pp. 6-13.
- GARCÍA, J. (1997), "Geomarketing. Los sistemas de información geográfica aplicados a la planificación comercial". *Distribución y Consumo*, 31, págs. 99-107.
- GIJSBRECHTS, E., K. CAMPO y T. GOOSSENS (2003), "The impact of store flyers on store traffic and store sales: A geo-marketing approach". *Journal of Retailing* 79, pp. 1-16.
- HAINING, R., S. WISE y P. SIGNORETTA (2000), "Providing scientific visualization for spatial data analysis: Criteria and an assessment of SAGE". *Journal of Geographical Systems*, 2; pp.121-140.
- LATOUR, P. y J. LE FLOC'H (2001), "Géomarketing: Principes, méthodes et applications". Éditions d'Organisation.
- MORENO, A. (2001), "La geografía de los servicios aplicada al marketing como tema de estudio". En *Geomarketing con Sistemas de Información Geográfica*, Universidad Autónoma de Madrid, págs. 9-20.
- MORENO, R. y E. VAYÁ (2000), "Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial". Edicions Universitat de Barcelona, col·lecció UB 44, manuals.
- ROSA, I.M. (2001), "Marketing territorial: localización de puntos de venta". En *Distribución comercial*, dir. E. Díez de Castro; Mc. Graw Hill, 2<sup>a</sup> ed. Aravaca (Madrid).
- UNWIN, A. (2000), "Using your eyes- making statistics more visible with computers". *Computational Statistics & Data Analysis*, 32; pp. 303-312.
- VOLLE, P.P. (2001), "Produit et information géographique: le géomerchandising". Centre Lillois d'Analyse et de Recherche sur l'Evolution des Entreprises, UPRESA CNRS 8020.