

Aglomeración espacial de establecimientos comerciales en un centro tradicional de negocios

Un análisis basado en las funciones K

Carlos Garrocho, José Antonio Álvarez-Lobato y Tania Chávez*

Fecha de recepción: 7 de enero de 2010; fecha de aceptación: 5 de octubre de 2010

Resumen: Comparamos dos de las estadísticas espaciales más robustas que reporta la literatura contemporánea para analizar patrones locacionales de firmas: las funciones K -plana y K -en red. Estas funciones generan valores que, en determinadas condiciones, permiten conocer la estructura espacial básica de la localización de las variables objeto de análisis, pudiéndose incorporar medidas de significancia estadística acerca de la aglomeración/dispersión/aleatoriedad locacional, tanto en espacios euclidianos como en espacios red, y permiten tratarlos como territorios continuos y no como la adición arbitraria de unidades geográficas discretas. Hasta donde sabemos, en la literatura internacional no existen estudios publicados que exploren las divergencias y coincidencias de las funciones K -plana y K -en red al caracterizar patrones espaciales básicos de firmas terciarias en el espacio intraurbano. Este trabajo busca subsanar este hueco en la literatura especializada.

Palabras clave: localización, aglomeración, medidas de concentración espacial, econometría espacial, funciones de densidad K .

Spatial Agglomeration of Firms in a Traditional Business Center: A K -function Analysis

Abstract: We compare two of the most robust spatial statistics reported in the contemporary literature to analyze locational patterns of firms: the K -planar and the K -network functions. These functions generate values that, under certain conditions, provide insights into the basic spatial structure of the location of the variables under analysis. Both functions are able to incorporate measures of statistical significance on the agglomeration/dispersion/randomness of locational patterns. The K -planar function does it in Euclidian spaces while the K -network function does it in network spaces. Both functions consider space as continuous territories, not as the arbitrary addition of discrete geographical units (*i.e.* municipalities, states). To our knowledge, there are no published studies in the inter-

*Carlos Garrocho Rangel, carlosgarrochorangel@yahoo.com.mx, profesor-investigador, El Colegio Mexiquense, y coordinador del Seminario de Estudios Estratégicos del Estado de México; José Antonio Álvarez-Lobato, jalvar@cmq.edu.mx, coordinador del Laboratorio de Análisis Espacial; Tania Chávez, tchavez@cmq.edu.mx, jefa de la Unidad de Informática, El Colegio Mexiquense, Zinacantepec, Estado de México.

national literature that explore the differences and similarities of the *K-planar* and *K-network* functions in characterizing basic spatial patterns of tertiary firms in the intra-urban space. This paper seeks to fill this gap in the literature.

Keywords: location, agglomeration, measures of spatial concentration, spatial econometrics, *K* density functions.

Clasificación JEL: C19, L81, R30.

Introducción

La tendencia de las actividades económicas a concentrarse en el territorio ha sido reconocida, al menos desde los trabajos clásicos de Alfred Marshall (Marshall, 1920). Este comportamiento espacial no es fortuito, sino que se debe a diversas ventajas que obtienen las firmas cuando se aglomeran en el territorio y que de manera genérica se les ha llamado *economías de aglomeración* (Fujita y Thisse, 2002; Fujita y Krugman, 2004; O'Sullivan, 2007). La naturaleza, el alcance y el origen de las economías de aglomeración son diversos (Rosenthal y Strange, 2004) e incluso pueden deberse a *ventajas naturales* (por ejemplo, en el espacio intraurbano podría ser la accesibilidad del Centro Tradicional de Negocios al resto de la ciudad) y aun al *azar* (Ellison y Glaeser, 1997; Krugman, 1991a; Mella, 2006).

El estudio de las economías que generan aglomeración espacial de firmas se ha revitalizado en el marco de la “nueva geografía económica”, derivada, principalmente, de los trabajos de Fujita (1988), Krugman (1991a; 1991b), Venables (1996), Fujita *et al.* (1999), y Ottaviano y Thisse (2003), entre otros. No obstante, se reconoce que existe, de inicio, un obstáculo importante para avanzar en el entendimiento de los procesos de aglomeración de las firmas: la carencia de un método que permita medir correctamente su concentración (dispersión) espacial en el territorio (Marcon y Puech, 2003; Quah y Simpson, 2003; Holmes y Stevens, 2004; Guillain y Le Gallo, 2007). Se acepta que si no se puede medir de manera confiable la concentración (dispersión) de las firmas, será complicado avanzar en el entendimiento de su comportamiento espacial (Duranton y Overman, 2005).

La literatura especializada reporta una gran diversidad de métodos e indicadores que se han utilizado para medir la aglomeración (dispersión) espacial de las firmas y las actividades económicas; sin embargo, no hay consenso sobre cuál o cuáles son los más apropiados (Guillain y Le Gallo, 2007). Duranton y Overman (2005) identifican tres *generaciones* de métodos que se han utilizado para estimar la concentración espacial de las actividades. La primera generación está encabezada, paradójicamente, por

indicadores *a-espaciales*, como los derivados del índice de Gini, que en realidad son ejercicios estadísticos desvinculados del territorio (aparentemente se asocian con el territorio porque organizan sus datos en unidades geográficas, pero en realidad no consideran ninguna variable espacial; *i.e.* Krugman, 1991a). La segunda generación de indicadores introduce algunos elementos espaciales y tiende a describir la concentración económica subyacente, pero, al igual que los métodos de la primera generación, considera el espacio de manera *discreta* (organizado en unidades espaciales administrativas delimitadas de manera arbitraria: países, estados, municipios) debido a que usualmente así está disponible la información estadística, pero no porque estas delimitaciones respondan a la conformación espacial de procesos económicos o porque sean las más adecuadas para investigar la aglomeración de firmas en el territorio (*i.e.* Glaeser *et al.*, 1992; Ellison y Glaeser, 1997; Devereux *et al.*, 2004). Finalmente, la tercera generación de métodos e indicadores, derivados en gran parte de la función *K* propuesta por Ripley (1976), se apoya en la concepción de espacios *continuos* (*i.e.* superficies euclidianas no fragmentadas por unidades político-administrativas artificiales y arbitrarias), lo que permite medir concentraciones de puntos (*i.e.* firmas) a diversas escalas espaciales de manera *simultánea* (*i.e.* Arbia, 2001; Arbia *et al.*, 2007; Marcon y Puech, 2003; Duranton y Overman, 2005, 2006).¹

Por lo tanto, desde una *perspectiva geográfica* las metodologías para estimar la aglomeración (dispersión) espacial de firmas en el territorio pueden dividirse en dos grandes categorías: las que consideran el territorio como un *espacio discreto* y las que lo consideran como un *espacio continuo*. Las primeras han sido ampliamente utilizadas (*i.e.* Maurel y Sédillot, 1999; Rysman y Greenstein, 2005; Mori *et al.*, 2005), pero se reconoce que tienen importantes limitaciones. Algunas de las principales son las siguientes: 1) se apoyan en datos artificialmente agregados en unidades espaciales *discretas* (*i.e.* países, estados, municipios), cuyos límites arbitrarios no corresponden necesariamente al comportamiento real de la economía

¹ Quizá los métodos e indicadores que Duranton y Overman (2005) llaman de tercera generación deberían en realidad considerarse como de una *cuarta generación*, ya que antes de ella aparecieron métodos que se derivan de la técnica del vecino más cercano (propuestos originalmente por Clark y Evans, 1954; *i.e.* Lee, 1979) y de otros métodos que permiten analizar patrones de puntos en espacios continuos (Anselin, 1992, 1995; Cressie, 1993). La diferencia de estos métodos con los derivados de la función *K* de Ripley es que no permiten, entre otras cosas, realizar análisis a diversas escalas espaciales de manera simultánea (análisis *multiescala*) o comparar patrones correspondientes a diferentes áreas de estudio, como sí lo hacen los métodos más sofisticados que Duranton y Overman llaman de tercera generación.

(“transforman puntos localizados en un mapa en unidades acumuladas en cajas”, Durantón y Overman, 2005, p. 1079); 2) las unidades espaciales discretas utilizadas se consideran homogéneas en su interior (lo que implica el grave problema de la *falacia ecológica*, Robinson, 1950); 3) la delimitación (forma) y escala (tamaño) de las unidades espaciales inciden directamente en los resultados de los análisis (*i.e.* lo que los geógrafos llaman el “problema de la unidad espacial modificable”, Openshaw, 1984); 4) Son metodologías *a-espaciales*, ya que no consideran la localización en el territorio de las unidades geográficas discretas en las que organizan la información (no se considera si las unidades espaciales son contiguas o si están más o menos distantes entre sí, *i.e.* Bertinelli y Decrop, 2005; Guillain y Le Gallo, 2007), y 5) sólo estiman un indicador general de concentración/dispersión, pero sin ninguna información sobre su *nivel de significancia*, de tal manera que no se puede saber con exactitud si los resultados obtenidos son producto del azar o son estadísticamente significativos. A cambio, estas metodologías reportan importantes ventajas: son relativamente fáciles de aplicar, se apoyan en datos ampliamente disponibles (*i.e.* censos) y su interpretación es usualmente sencilla (Albert *et al.*, 2007; *i.e.* Ellison y Glaeser, 1997; Maurel y Sédillot, 1999).

Las debilidades de las metodologías basadas en la concepción del espacio *discreto* son adecuadamente subsanadas por las que lo conciben de manera *continua* (Arbia *et al.*, 2007; Durantón y Overman, 2005; Marcon y Puech, 2003). Sin embargo, estos métodos enfrentan, cuando menos, dos problemas *operativos* importantes: uno relacionado con la información estadística que requieren, y el otro con los procedimientos de estimación. El primero se refiere a que son muy demandantes en términos de los datos necesarios para realizar los análisis, ya que se apoyan en la localización precisa (*i.e.* coordenadas geográficas) de las firmas. Esta información usualmente no está registrada en fuentes de amplio acceso, por lo que se tiene que levantar en campo, y cuando llega a estar disponible su utilización enfrenta, por lo regular, restricciones de confidencialidad (*i.e.* Guillain y Le Gallo, 2007; Kosfeld *et al.*, 2009). El segundo problema operativo se relaciona con lo laborioso y complicado de los cálculos numéricos, lo que hace que estas metodologías sean poco prácticas si no se dispone del *software* adecuado (Mitchell, 2005).

Las metodologías para medir la aglomeración (dispersión) de firmas en espacios continuos se han apoyado, fundamentalmente, en el cálculo de funciones derivadas de la función *K* de Ripley (Ripley, 1976), considerada una de las más robustas para analizar patrones de puntos a diversas escalas espaciales de manera simultánea (Bailey y Gatrell, 1995; Cressie,

1993; Mitchel, 2005). Dado que la función K no se ve afectada por los problemas inherentes a los métodos que utilizan unidades espaciales discretas, no es raro que se haya utilizado desde hace años para analizar patrones territoriales de diversos tipos de fenómenos naturales (*i.e.* Getis y Franklin, 1987), y que recientemente se esté probando para explorar el comportamiento espacial de fenómenos sociales y económicos (*i.e.* Yamada y Thill, 2004; Duranton y Overman, 2005).

Sin embargo, el método estándar de la función K (la llamada función K plana, Yamada y Thill, 2004) se basa en el supuesto de un espacio euclidiano e isotrópico (plano, infinito, bidimensional) donde los puntos pueden ocurrir en cualquier sitio y las distancias entre los puntos se miden en línea recta. Algunos autores han argumentado que este supuesto es insostenible cuando se intenta analizar patrones de puntos (*i.e.* firmas) cuya localización está restringida a territorios finitos y unidimensionales (Miller, 1999; Okabe y Yamada, 2001), como las redes de carreteras en espacios regionales (*i.e.* Lu y Chen, 2006) o la red de calles en espacios intraurbanos (Myint, 2008).

Como la medición de la distancia desempeña un papel clave en la estimación de la función K , medirla en un *espacio euclidiano* o en un *espacio red* es de vital importancia, ya que una u otra forma de medición puede conducir a conclusiones distintas respecto del patrón espacial (*i.e.* aglomerado, disperso o aleatorio) de una misma muestra de puntos (Lu y Chen, 2006).

Para estas situaciones, Okabe y Yamada (2001) han propuesto un método de análisis que expande las propiedades de la función K plana a un *espacio red* (*i.e.* una red de calles). En este método, la función K en red considera explícitamente que la localización de los eventos (*i.e.* firmas) está restringida a una red determinada, por lo que las distancias deben medirse no de manera euclidiana, sino a lo largo de la red en cuestión. La función K en red se ha aplicado recientemente para analizar el patrón espacial de accidentes en redes carreteras de Estados Unidos y Japón (Yamada y Thill, 2004; Xu, 2002; Okabe y Satoh, 2006), el robo de autos en áreas metropolitanas norteamericanas (Lu y Chen, 2006) y la localización de algunos tipos de servicios en ciudades pequeñas del medio oeste estadounidense (Myint, 2008).

Como las estimaciones de las funciones K plana y en red implican cálculos complejos y laboriosos, su aplicación en el campo de la geografía económica ha sido limitada (Lu y Chen, 2006), y los estudios publicados que utilizan las funciones K para analizar patrones de actividades terciarias en el *espacio intraurbano* son aún casos de excepción, que utilizan muestras

muy reducidas de negocios (Myint, 2008) o de usos del suelo (Cuthbert y Anderson, 2004). Esto es aún más notorio si nos referimos a ciudades de países en desarrollo (*i.e.* México), donde es poco lo que se conoce sobre el comportamiento espacial de las firmas (Duranton y Overman, 2006, p. 1).

El objetivo de este trabajo es identificar y examinar las coincidencias y divergencias de los resultados de la función K plana y la función K en red en el análisis del patrón de localización de firmas terciarias en el centro tradicional de negocios (CTN) del área metropolitana de Toluca (AMT, México), con el fin de verificar las ventajas que algunos autores le atribuyen a la función K en red cuando se estudian espacios intraurbanos de escala reducida (Okabe y Yamada, 2001). Los resultados de este trabajo permitirán orientar la correcta aplicación de las funciones K para analizar la distribución espacial de firmas terciarias al interior de ciudades mexicanas; es decir, para avanzar en el análisis a nivel micro del alcance geográfico de las economías de aglomeración en espacios intraurbanos (Rosenthal y Strange, 2004). Hasta donde sabemos, no existen antecedentes publicados en la literatura internacional que utilicen simultáneamente las funciones K plana y K en red para estimar la aglomeración de amplias muestras de negocios en centros tradicionales de grandes ciudades y que comparen sus resultados. Este trabajo busca subsanar este hueco en la literatura especializada. Sin embargo, la misma carencia de estudios como el que aquí se presenta impidió llevar a cabo análisis comparativos de nuestros resultados con otros que hubiesen podido servir de antecedentes.

En lo que sigue el artículo se divide en cinco secciones. En la primera se presentan de manera sucinta las funciones K plana y K en red, y se explica el método del análisis comparativo que articula este trabajo. En la siguiente sección se describen la zona de estudio y los datos utilizados, para situar el contexto espacial y estadístico del análisis. Posteriormente se estiman las funciones K plana y K en red, y se analizan comparativamente sus principales resultados. Finalmente, se presentan las conclusiones del trabajo y una lista de la bibliografía consultada.

I. Funciones K plana y K en red y método para su análisis comparativo

1.1. La función K plana

La función K plana es propuesta por Ripley (1976) con el propósito de probar la concentración (dispersión) de distribuciones espaciales de puntos.

El cálculo de esta función toma en cuenta la localización relativa de todos los puntos considerados en el análisis (*i.e.* la distancia entre todos ellos), lo que permite generar resultados simultáneos en un amplio rango de escalas espaciales (Yamada y Thill, 2004).²

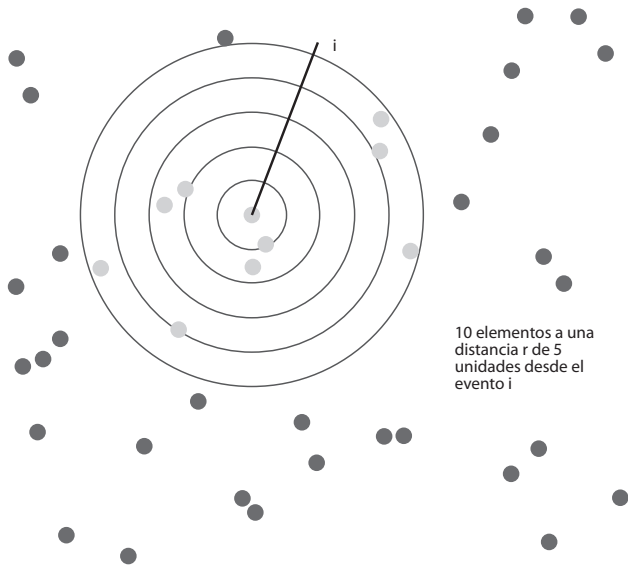
Los métodos espaciales que sintetizan distribuciones de puntos en un solo indicador global, como los derivados del cálculo de distancias al vecino más cercano, no tienen esta propiedad *multiescala*, ya que solamente consideran pares de puntos (*i.e.* la distancia de un punto al vecino más cercano) y una sola escala de análisis (*i.e.* la del área de estudio). En cambio, la función *K* plana ofrece, simultáneamente, resultados para muy diversas escalas espaciales dentro del área de estudio. Esta característica es particularmente importante, ya que los patrones de puntos pueden registrar comportamientos espaciales diferentes a escalas geográficas distintas (*i.e.* aglomerados a cierta escala, pero dispersos a otra) (Bailey y Gatrell, 1995; Dixon, 2002).

La idea básica detrás de la función *K* es examinar si una distribución espacial de puntos *observada* es significativamente distinta de una distribución espacial *aleatoria*. Es decir, si difiere en términos estadísticamente confiables de un proceso espacial homogéneo de tipo Poisson (también llamado de “aleatoriedad espacial completa” o *complete spatial randomness* [CSR], como se le conoce en inglés). En una distribución espacial aleatoria (o CSR) todos los puntos tienen la misma probabilidad de ocurrir en cualquier sitio del área de estudio, y la presencia de un punto en una cierta localización no afecta las probabilidades de ocurrencia de otros puntos. La función *K* permite comparar el número acumulado de puntos en bandas consecutivas de distancia de una distribución espacial *observada*, con el acumulado de puntos en las mismas bandas de distancia de una distribución espacial *aleatoria teórica* (Mitchell, 2005).

Para entender de manera intuitiva la función *K* plana, supóngase que se selecciona al azar cierto punto (*i.e.* firma) de una distribución cuyo patrón se quiere analizar. A partir de ese punto se define una serie de círculos concéntricos, cuyos radios se van incrementando de manera constante. Estos círculos serían *ventanas* que nos permiten observar el área de estudio en diferentes escalas. Luego se considera cada círculo de manera individual y se cuenta el número de puntos que se localizan dentro de cada

² Aunque en términos prácticos, la distancia máxima a la que se puede calcular la función *K* está definida por las dimensiones y características del área de estudio, como se describe más adelante. Sin embargo, a las diferentes escalas en que se puede calcular se considera que está libre del problema de la *unidad espacial modificable* (Lu y Chen, 2006).

Gráfica 1. Construcción de la función K plana



Fuente: Elaboración propia con base en Gatrell *et al.* (1996).

uno de ellos (gráfica 1). Se repite este procedimiento para todos los puntos de la distribución bajo análisis. A continuación se obtiene el promedio del número acumulado total de puntos en cada distancia considerada (cada *distancia considerada* es el radio de cada ventana o círculo concéntrico que se definió previamente alrededor de cada punto de la distribución analizada) dividido entre la densidad de puntos, y se grafica en el eje vertical de un plano cartesiano, mientras que los diferentes rangos de distancia (la medida de los radios de los círculos concéntricos previamente establecidos) se grafican en el eje horizontal. La curva resultante es la función K plana. Así, la función K plana a una cierta distancia se calcula como la media observada del número de puntos a esa determinada distancia, desde un punto cualquiera de la muestra, dividida por la intensidad total (densidad) de puntos (Yamada y Thill, 2004).

Efectuar pruebas estadísticas de significancia para la función K plana es ciertamente complicado, excepto cuando el área de estudio tiene una forma geométrica simple (*i.e.* rectángulo, círculo; Cressie, 1993). Pero cuando el área de estudio tiene una forma compleja, como usualmente ocurre en la realidad, normalmente se realizan simulaciones con el méto-

do Monte Carlo para generar funciones K aleatorias teóricas que sirvan como referente de comparación (*patrones contrafactuales*, Durantón y Overman, 2005) a la distribución observada, y así generar *seudoniveles* de significancia (Yamada y Thill, 2004; Manly, 1997).

Por lo tanto, lo que se hace es comparar gráficamente la función K plana observada con las funciones K teóricas aleatorias (derivadas para la misma cantidad de puntos analizados y para la misma área de estudio) que resultan de realizar suficientes simulaciones (*i.e.* 1000) mediante el método Monte Carlo (Baddeley y Turner, 2005). Por lo tanto, si a una determinada distancia el número de puntos acumulados en la función K plana observada es significativamente mayor que la que registra la función K teórica aleatoria (estimada con el método Monte Carlo), se puede concluir que la distribución de puntos observada sigue un patrón aglomerado a esa distancia (*i.e.* mayor que 99 por ciento de las funciones K teóricas aleatorias). Si, por el contrario, el número acumulado de puntos en la función K observada es significativamente menor que el de la función K teórica aleatoria, se puede concluir que la distribución de puntos observada sigue un patrón disperso (o de rechazo espacial). Si las diferencias entre las funciones K observada y teórica aleatoria no son significativas se puede concluir, lógicamente, que la distribución de puntos observada sigue un patrón espacial aleatorio. De esta manera, la función K plana permite identificar si una distribución espacial de puntos observada sigue un patrón *aglomerado*, *disperso* (de rechazo espacial) o *aleatorio* en un espacio bidimensional a diferentes escalas geográficas (Lu y Chen, 2006; Mitchell, 2005).

En análisis empíricos de patrones de puntos se ha detectado que la manera en que se definen las fronteras del área de estudio puede afectar las estimaciones de la función K . Una razón es que puede ocurrir que diversos puntos queden fuera del área de estudio (a veces muy cerca de sus límites) o porque se aplican a espacios finitos métodos estadísticos diseñados para espacios infinitos (Yamada y Thill, 2004). Dado que las fronteras del área de estudio no son usualmente extensas sino que por razones operativas deben circunscribirse a una determinada superficie (dejando necesariamente fuera del área de estudio puntos que pertenecen a la misma distribución), Ripley (1976) y otros (*i.e.* Cressie, 1993; Stoyan y Stoyan, 1994; Baddeley y Turner, 2005) han diseñado una gran variedad de correctores estadísticos de los llamados *efectos de borde* (*edge-effects*), derivados de la forma y extensión de las superficies consideradas como áreas de estudio, lo que otorga gran validez a los resultados de la función K plana y permite hacer comparaciones entre distintas áreas de estudio (Bailey y

Gatrell, 1995; Durantón y Overman, 2005). El software utilizado en este trabajo para estimar la función K plana (Spatstat) implementa automáticamente correcciones para eliminar los efectos de borde, si así se desea (Baddeley y Turner, 2005).

1.2. La función K plana: su expresión matemática

Los detalles matemáticos de la función K plana se reportan ampliamente en la literatura internacional y no vale la pena repetirlos en detalle en este espacio (*i.e.* Bailey y Gatrell, 1995; Cressie, 1993; Diggle, 2003; Lu y Chen, 2006; Ripley, 1976; Ripley, 1981; Yamada y Thill, 2004; entre muchos otros). Sin embargo, sí es necesario presentar sus principales rasgos para tener claro el método de cálculo que se siguió en este trabajo.

Dada una distribución de puntos $P = \{p_1, \dots, p_n\}$ en una determinada área de estudio, la función K plana compara el valor *observado* de K a una cierta distancia (K_{obs}) con el valor *esperado* de K a esa misma distancia. Así, la función K plana *esperada* se puede expresar como:

$$K(r) = \frac{1}{\lambda} E \left(\begin{array}{l} \text{número de puntos en } P \text{ dentro} \\ \text{de la distancia euclidiana } r \text{ de} \\ \text{un punto arbitrario en } P \end{array} \right) \quad (1)$$

Donde:

r : distancia a la cual se calcula K

λ : estimación de la distribución de la densidad de puntos en P (*i.e.* número de puntos por unidad de área)

$E(\cdot)$: valor esperado de puntos

Cuando el patrón de puntos sigue un proceso Poisson homogéneo, esto es, que tiene aleatoriedad espacial completa (CSR), la función puede expresarse (Smith, 2009; Okabe y Yamada, 2001; Dixon, 2002) como:

$$K(r) = \pi r^2 \quad (2)$$

Para un conjunto observado de n puntos, distribuidos sobre una región con área A , la función K *observada* toma la forma:

$$K_{obs}(r) = \frac{1}{\lambda' n} \sum \sum_{i \neq j} I_r(d_{ij}) \quad (3)$$

donde:

n : número de puntos observados

d_{ij} : distancia euclidiana entre los puntos p_i y p_j

$$I_r(d_{ij}) = \begin{cases} 1 & \text{si } d_{ij} \leq r \\ 0 & \text{otro caso} \end{cases} \quad (4)$$

$$\lambda' = \frac{n}{A} \quad (5)$$

Al sustituir λ' en la fórmula $K_{obs}(3)$, $K_{obs}(r)$ puede definirse como:

$$K_{obs}(r) = \frac{A}{n^2} \sum \sum_{i \neq j} I_r(d_{ij}) \quad (6)$$

Para simplificar los cálculos (Diggle, 2003), λ' puede redefinirse como:

$$\lambda' = \frac{(n-1)}{A} \quad (7)$$

que es prácticamente igual a la definición anterior de λ' para valores suficientemente grandes de n , por lo que la función $K_{obs}(r)$ puede representarse como:

$$K_{obs}(r) = \frac{A}{n(n-1)} \sum \sum_{i \neq j} I_r(d_{ij}) \quad (8)$$

Los factores de corrección de los *efectos de borde* se pueden incluir como término adicional dentro de las sumatorias de la ecuación (8) (Dixon, 2002).

Conforme se incrementa el valor de r hasta cubrir la distribución de puntos P en el área de estudio, se puede obtener una serie de valores $K_{obs}(r)$. Estos valores se grafican en el eje vertical de un plano cartesiano contra los valores r que se grafican en el eje horizontal. Al comparar la curva resultante de $K_{obs}(r)$ con la que resulta de $K(r)$ es posible distinguir si la distribución observada de puntos P sigue un patrón *aglomerado*, *disperso* o *aleatorio*, estadísticamente significativo a diversas escalas espaciales (Levine, 2004; Lu y Chen, 2006).

Debido a la alta variabilidad al estimar $K(r)$, cuando el valor de r es elevado, si se usan ciertas correcciones de efecto de frontera o el fenómeno

a evaluar es desconocido, la recomendación usual es tener como distancia máxima 0.5 de la dimensión más corta del área de estudio si ésta es rectangular, o $\sqrt{A/2}$ para un polígono de área A (Dixon, 2002), aunque un valor más conservador sugiere 0.25 de la dimensión más corta del rectángulo que engloba los puntos a estudiar, o de la distancia entre los pares de puntos más alejados (Kosfeld *et al.*, 2009).

I.3. La función K en red

La función K en red es una derivación de la función K plana que data apenas de la década de 1990 (Okabe *et al.*, 1995; Okabe y Kitamura, 1996; Okabe y Yamada, 2001). Su principal diferencia respecto a la función K plana es que utiliza un espacio-red *unidimensional* en lugar de un espacio isotrópico *bidimensional*. En ese espacio red se miden las distancias entre los puntos de la distribución que se quiere analizar.

De manera similar a la función K plana, el método de Okabe y Yamada (2001) utiliza simulaciones tipo Monte Carlo para generar una serie de funciones K en red derivadas de distribuciones teóricas aleatorias de puntos, donde cada punto tiene la misma probabilidad de ocurrir en cualquier sitio de la red (*i.e.* una red de calles), y donde la presencia de un punto en un determinado sitio no afecta la probabilidad de ocurrencia de otro punto sobre la red.

Al igual que se hace cuando se analiza la función K plana, la función K en red *observada* se compara con funciones K en red *teóricas aleatorias*. Si la función K en red observada es mayor o menor que cierto porcentaje de las funciones K en red derivadas de patrones aleatorios (*i.e.* 99 por ciento u otro valor, dependiendo del nivel de significancia seleccionado), se puede concluir que la muestra de puntos *observada* en el espacio red sigue un patrón espacial aglomerado (si es mayor que el porcentaje determinado), disperso (si es menor que el porcentaje determinado) o aleatorio (si no hay diferencias significativas con las funciones K *aleatorias teóricas*).

Si resumimos, la función K en red se aplica a un espacio red *unidimensional*, mientras que la función K plana opera en un espacio *bidimensional*. Por lo tanto, es de esperarse que los valores de las dos funciones sean distintos para muestras de puntos iguales, en función de la suposición de un espacio bidimensional (*i.e.* el centro tradicional de negocios de una ciudad) o un espacio red (*i.e.* la red de calles en el centro tradicional de negocios de la misma ciudad).

I.4. La función K en red: su expresión matemática

Dado un conjunto de puntos $P = \{p_1, \dots, p_n\}$ distribuidos en una red N conformada por enlaces y nodos (esto es, segmentos de calle y cruces de calles), se definen como L_T todos los enlaces que interconectan cada nodo, y como $|L_T|$ la longitud total de L_T .

Al suponer un proceso de puntos binomial, esto es, que los puntos estén uniforme e independientemente distribuidos en la red N la función K en red se puede definir como:

$$K(r) = \frac{1}{\omega} E \left(\begin{array}{l} \text{número de puntos en } P \\ \text{dentro de la distancia en red} \\ r \text{ de un punto } p, \text{ en } L_T \end{array} \right) \tag{9}$$

donde:

ω : estimador de la densidad de puntos a lo largo de la red (i.e. número de puntos por unidad de longitud).

$E(\)$: Valor esperado respecto a todas las posibles ubicaciones de p sobre L_T .

r : Distancia en red a la que se calcula la función K .

Si $L_p(r)$ es un subconjunto de L_T , donde la distancia en red entre p y cualquier punto en $L_p(r)$ es menor o igual que r , y $|L_p(r)|$ es la longitud de $L_p(r)$ ya que $\omega|L_p(r)|$ indica el número esperado de puntos de P en $|L_p(r)|$, la ecuación (9) se puede expresar como:

$$K(r) = \frac{1}{|L_T|} \int_{p \in L_T} |L_p(r)| dp \tag{10}$$

donde:

dp : Es la integración de p a lo largo de L_T .

Ahora bien, para un conjunto observado de n puntos de P , distribuidos a lo largo de la red definida por los enlaces L_T , la función K observada puede ser expresada como:

$$K_{obs}(r) = \frac{1}{\omega n} \sum_i \sum_{i \neq j} I_r(s_{ij}) \tag{11}$$

donde:

s_{ij} : Distancia más corta en red entre los puntos p_i y p_j .

$$I_r(s_{ij}) = \begin{cases} 1 & \text{si } s_{ij} \leq r \\ 0 & \text{otro caso} \end{cases} \quad (12)$$

$$\omega' = \frac{n}{|L_T|} \quad (13)$$

Así, la expresión de la función K_{obs} se puede definir como:

$$K_{obs}(r) = \frac{|L_T|}{n^2} \sum \sum_{i \neq j} I_r(s_{ij}) \quad (14)$$

Si se opta por el estimador de la densidad de puntos $\omega' = \frac{(n-1)}{|L_T|}$, la función K_{obs} se expresará como:

$$K_{obs}(r) = \frac{|L_T|}{n(n-1)} \sum \sum_{i \neq j} I_r(s_{ij}) \quad (15)$$

Mediante el método Monte Carlo se realizan simulaciones para generar una serie de distribuciones teóricas aleatorias de n puntos en N , con el fin de estimar el valor esperado de la función K en red, $K_{obs}(r)$, en condiciones de CSR. Las funciones *teóricas aleatorias* $K(r)$ se comparan con la función $K_{obs}(r)$ para establecer si el patrón de puntos analizado es *aglomerado*, *disperso* o *aleatorio* a una distancia r a lo largo de la red de calles N . Una explicación más detallada, incluyendo la prueba de significancia estadística de la función K en red, puede revisarse en Okabe y Yamada (2001).

1.5. Ejes del análisis comparativo de las funciones K

En este trabajo las funciones K plana y en red se aplicaron a los mismos conjuntos de firmas terciarias localizadas en una misma área de estudio (el Centro Tradicional de Negocios [CTN] del área metropolitana de Toluca [AMT]) y se compararon sus respectivos resultados en tres aspectos clave: 1) si coinciden o no en la *caracterización* del patrón espacial de las firmas (*i.e.* concentrado, disperso o aleatorio); 2) si coinciden o no en la localización del pico de mayor aglomeración de las firmas (lo que indica la *escala* de la aglomeración), y 3) si coinciden o no en la *intensidad* de la concentración de las firmas. Así, los ejes del análisis comparativo de los resultados de las funciones K plana y K en red que articulan la parte empírica de este trabajo son: *caracterización*, *escala* e *intensidad*. Estos tres ejes, y sus respectivos indicadores, permiten ahondar más en el análisis comparativo que la sola inspección visual de las gráficas de las funciones K , que ha sido

el principal método de comparación utilizado hasta ahora (Yamada y Thill, 2004; Lu y Chen, 2006).

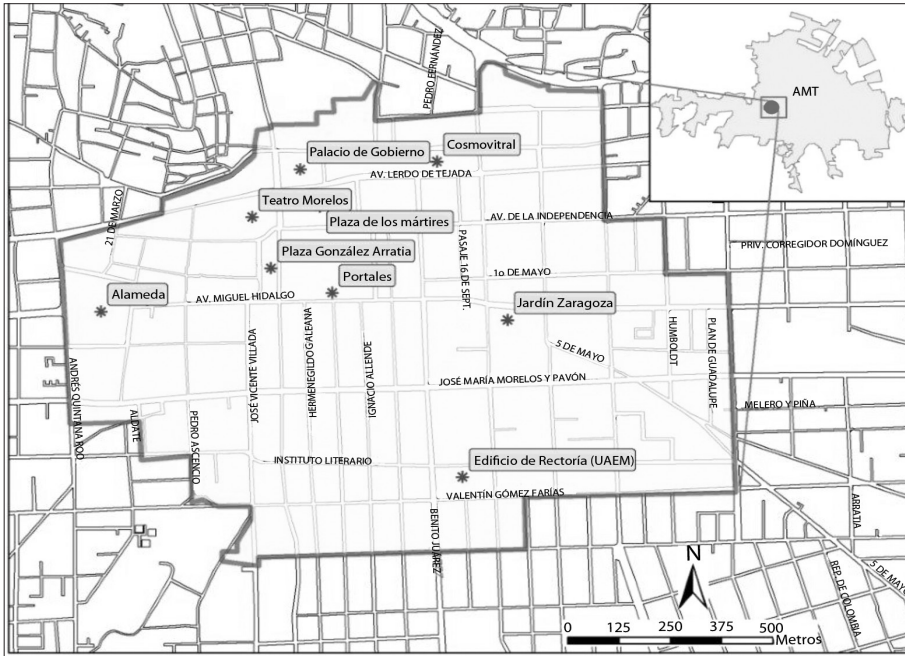
Dado el nivel de complejidad de los cálculos para la estimación de las funciones K plana y en red y de sus niveles de significancia estadística, se utilizaron dos paquetes de cómputo plenamente aceptados en la literatura internacional: *Spatstat* para estimar la función K plana (Baddeley y Turner, 2005) y *SANET* (Okabe *et al.*, 2007) para estimar la función K en red.

Para permitir la comparación entre ambas funciones se tomó como referente base el cálculo de la función K con *Spatstat* utilizando factores de corrección de efectos de frontera (el factor de corrección isotrópico de Ripley [1988], que puede utilizarse para áreas poligonales) y 512 intervalos de cálculo. Se consideró como distancia máxima de r el valor estimado por *Spatstat* de 338 metros (el valor conservador de 0.25 de la longitud de los límites del área de estudio). Para el cálculo de la función K en red se ajustaron las distancias máximas e intervalos a los mismos valores especificados para la función K plana. Para determinar las bandas superior e inferior de los niveles de confianza se realizó el procedimiento estándar de 999 simulaciones Monte Carlo (Yamada y Thill, 2004).

II. Zona de estudio y datos utilizados

II.1. Zona de estudio

El área metropolitana de Toluca (AMT) se localiza a media hora de la ciudad de México (40 kilómetros en dirección suroeste), y forma parte de la zona metropolitana del valle de Toluca, que está integrada por doce municipios del Estado de México. Tiene una superficie de 269.6 kilómetros cuadrados; su eje de longitud máxima Este-Oeste es de 31.6 kilómetros, y el Norte-Sur es de 21.1 kilómetros. Con 1.6 millones de habitantes, es una de las cinco ciudades más grandes de México. El empleo formal ha tenido un comportamiento ascendente, aunque afectado por las recurrentes crisis económicas del país; en 1994 contaba con 157 284 empleos formales que llegaron a 267 603 en 2004, lo cual indica que más allá de coyunturas económicas la zona metropolitana del valle de Toluca es una de las áreas urbanas más dinámicas a escala nacional y ha registrado un crecimiento metropolitano de alta intensidad. Esto ha modificado su estructura urbana y ha transformado la ciudad monocéntrica de principios de la década de 1980 en una enorme área metropolitana policéntrica al inicio del siglo XXI (Garrocho y Campos, 2007).

Mapa 1. Zona de estudio CTN

Fuente: Elaboración propia.

Para definir el centro tradicional de negocios del AMT se consideró como criterio preliminar la existencia de usos de suelo comercial continuos en las vialidades del CTN, y como segundo criterio la intensidad de flujos peatonales en la zona. Los detalles se explican en Garrocho y Flores (2009a; 2009b). El resultado es una zona continua y de forma irregular con una extensión de 154.8 hectáreas (poco más de 100 manzanas) que incluye alrededor de 4 300 unidades comerciales y de servicios, lo que arroja una densidad bruta de casi 28 unidades económicas por hectárea. Su extensión máxima en dirección Este-Oeste es de 1 770 metros y de 1 500 metros en dirección Norte-Sur. En total, la longitud de la red de calles en la zona de estudio alcanza 23.4 kilómetros (mapa 1).

II.2. Datos utilizados

En el AMT no existen registros automatizados de los negocios localizados en el CTN, por lo que fue necesario llevar a cabo un intenso trabajo de cam-

po dividido en dos etapas. En la primera se realizó un levantamiento exhaustivo de todas las firmas, viviendas, locales vacíos y usos del suelo diversos (*i.e.* parques, plazas) en un área deliberadamente más extensa que la zona de estudio; y en la segunda se midieron las afluencias peatonales en 120 puntos de conteo durante doce horas diarias en tres días consecutivos, para identificar con mayor precisión los límites del CTN (mapa 1) (Garrocho y Flores, 2009a; 2009b). El resultado del trabajo de campo se organizó en una base de datos georreferenciada vinculada a un sistema de información geográfica (ArcGis), lo que facilitó el manejo de la información estadística y cartográfica, y su interfaz con *Spatstat* y *SANET*.

III. Análisis comparativo de los resultados de las funciones K plana y K en red: caracterización, escala e intensidad

Con el propósito de analizar los resultados de las funciones K plana y K en red (en adelante: K_{plana} y K_{red}) en términos de la *caracterización* del patrón espacial de las firmas, así como de la *escala e intensidad* de su aglomeración en la zona de estudio, se seleccionaron doce giros comerciales y

Cuadro 1. Tipo de unidades económicas seleccionadas en el área de estudio

<i>Giro</i>	<i>Tipo</i>	<i>Total de unidades económicas</i>	<i>Descripción</i>
Abarrotes	Comercio	94	Abarrotes, dulcerías, perecederos, fruterías
Bancos	Servicios	37	Sucursales bancarias, agencias financieras, casas de préstamo y empeño
Alimentos preparados	Comercio	240	Alimentos preparados, pastelerías, panaderías, tortillerías
Electrónicos	Comercio	61	Audio y video, componentes electrónicos, computadoras, instrumentos musicales
Escuelas	Servicios	78	Preescolar, primaria, secundaria, media superior y superior (públicas y privadas)
Farmacias	Comercio	43	Farmacias (tanto de cadenas nacionales como de operación local)

Cuadro 1. Tipo de unidades económicas seleccionadas en el área de estudio (continuación)

<i>Giro</i>	<i>Tipo</i>	<i>Total de unidades económicas</i>	<i>Descripción</i>
Iglesias	Servicios	16	Templos religiosos
Mercerías	Comercio	167	Mercerías, regalos, papelerías, telas y bordados
Mueblerías	Comercio	66	Muebles, electrodomésticos, enseres para el hogar
Restaurantes	Servicios	71	Restaurantes, bares, cafés, comida rápida
Ropa	Comercio	310	Ropa, boutiques, artículos para novias, accesorios
Zapaterías	Comercio	85	Zapaterías

Fuente: Elaboración propia.

de servicios representativos del CTN del AMT: ocho de ellos correspondientes a comercio minorista y cuatro a servicios, que suman en total 1 268 unidades económicas (cuadro 1).

III.1. Caracterización de los patrones espaciales y variación en el territorio

Para caracterizar los patrones espaciales de las unidades económicas consideradas, se graficaron las funciones K_{plana} y K_{red} tanto *teórica* como *observada* (K_{obs} y $K_{teórica}$) para los doce giros comerciales seleccionados, incluyendo sus niveles de significancia. El patrón espacial bajo análisis se caracterizó como *aglomerado* si la función K_{obs} se ubicó por encima del límite superior del nivel de confianza de la $K_{teórica}$; como *aleatorio* si la K_{obs} se encontró dentro de los límites superior e inferior de los niveles de significancia de la $K_{teórica}$; y *disperso* si se localizó por debajo del límite inferior del nivel de confianza de la función $K_{teórica}$ (*i.e.* Yamada y Thill, 2004; Lu y Chen, 2006).

El análisis comparativo de la caracterización de los patrones espaciales derivados de las funciones K_{plana} y K_{red} debe considerar, al menos, dos elementos centrales: 1) el *tipo* de patrón espacial predominante (aglomerado, disperso/rechazo espacial, aleatorio) y 2) las *variaciones* territoriales del

Cuadro 2. Caracterización de los patrones espaciales predominantes de las unidades económicas seleccionadas

<i>Giro</i>	<i>Función K_{plana}</i>		<i>Función K_{red}</i>	
	<i>Patrón predominante</i>	<i>Variaciones territoriales del patrón predominante</i>	<i>Patrón predominante</i>	<i>Variaciones territoriales del patrón predominante</i>
Abarrotes	Aleatorio	Mixto: Ligera aglomeración entre 38 y 91 metros. Aleatorio en el resto de las distancias.	Aglomerado	Mixto: Ligera aleatoriedad a menos de 117 metros y una ligera aglomeración en el resto de las distancias.
Bancos	Aglomerado	Mixto: Ligera aleatoriedad entre 75 y 108 metros, claramente aglomerado a partir de 108 metros.	Aglomerado	Puro: Claramente aglomerado a cualquier distancia.
Alimentos preparados	Aglomerado	Puro: Ligeramente aglomerado a cualquier distancia.	Aglomerado	Puro: Aglomerado a cualquier distancia.
Electrónicos	Aglomerado	Puro: Aglomerado a cualquier distancia.	Aglomerado	Puro: Claramente aglomerado a cualquier distancia.
Escuelas	Aleatorio	Mixto: Mínima aglomeración entre 14 y 85 metros.	Aglomerado	Puro: Ligeramente aglomerado a cualquier distancia.
Farmacias	Aleatorio	Mixto: Ligera aglomeración a partir de 246 metros.	Aglomerado	Puro: Aglomerado a cualquier distancia.

Cuadro 2. Caracterización de los patrones espaciales predominantes de las unidades económicas seleccionadas (continuación)

<i>Giro</i>	<i>Función K_{plana}</i>		<i>Función K_{red}</i>	
	<i>Patrón predominante</i>	<i>Variaciones territoriales del patrón predominante</i>	<i>Patrón predominante</i>	<i>Variaciones territoriales del patrón predominante</i>
Iglesias	Aleatorio	Puro: Aleatorio a cualquier distancia.	Aleatorio	Mixto: Ligera aglomeración entre 45 y 127 metros.
Mercerías	Agglomerado	Puro: Fuertemente aglomerado a cualquier distancia.	Agglomerado	Puro: Claramente aglomerado a cualquier distancia.
Mueblerías	Agglomerado	Puro: Fuertemente aglomerado a cualquier distancia.	Agglomerado	Puro: Claramente aglomerado a cualquier distancia.
Restaurantes	Agglomerado	Puro: Agglomerado a cualquier distancia.	Agglomerado	Puro: Claramente aglomerado a partir de 180 metros.
Ropa	Agglomerado	Puro: Fuertemente aglomerado a cualquier distancia.	Agglomerado	Puro: Claramente aglomerado a cualquier distancia.
Zapaterías	Agglomerado	Puro: Fuertemente aglomerado a cualquier distancia.	Agglomerado	Puro: Claramente aglomerado a cualquier distancia.

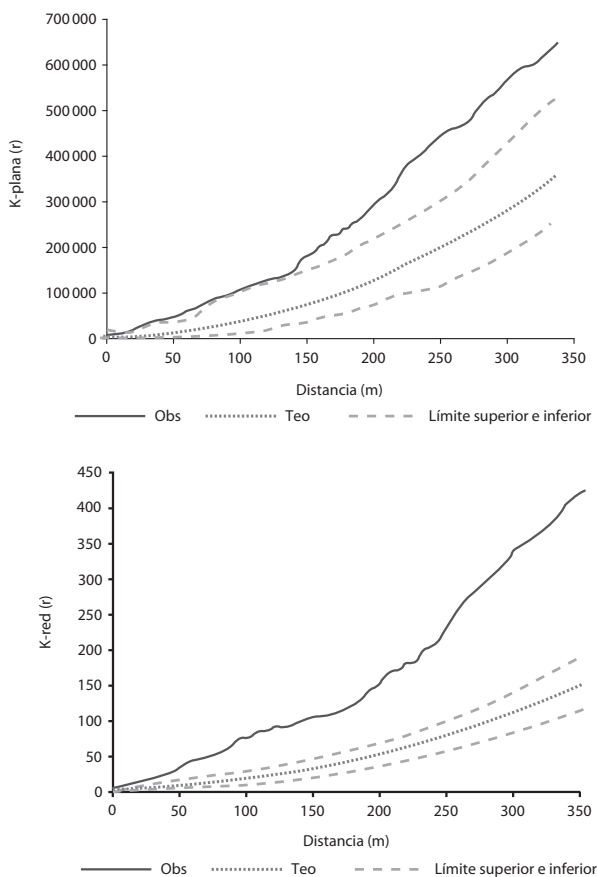
Fuente: Elaboración propia.

patrón espacial predominante (los cambios del patrón predominante a diferentes escalas y distancias: si el patrón predominante pasa de aglomerado a aleatorio o de aleatorio a disperso, por ejemplo) (cuadro 2).

III.1.1. Caracterización del tipo de patrón espacial predominante

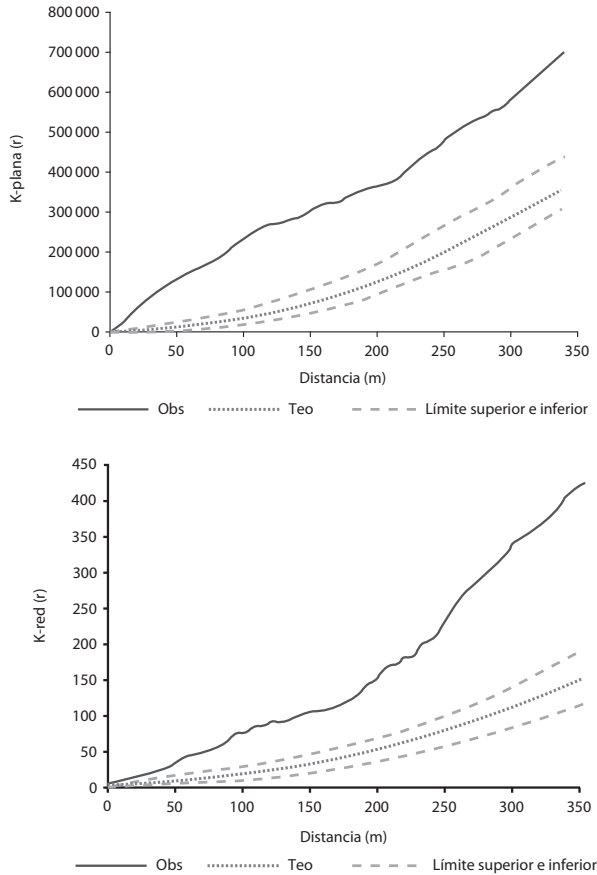
En términos de la caracterización del tipo de patrón espacial predominante, los resultados de las funciones K_{plana} y K_{red} coinciden en nueve de los

Gráfica 2. Comparación de la función K plana contra la función K red, giro seleccionado: bancos



Fuente: Elaboración propia.

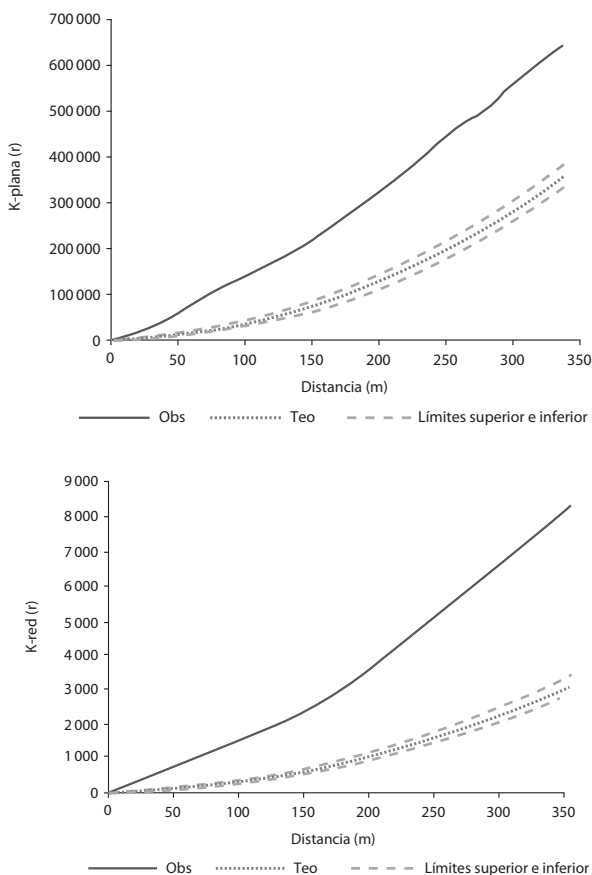
Gráfica 3. Comparación de la función K plana contra la función K red, giro seleccionado: mueblerías



Fuente: Elaboración propia.

doce casos (bancos: aglomerado; alimentos preparados: aglomerado; electrónico: aglomerado; iglesias: aleatorio; mercerías: aglomerado; mueblerías: aglomerado; restaurantes: aglomerado; ropa: aglomerado; zapaterías: aglomerado) (gráficas 2, 3 y 4). Sólo en tres casos arrojan resultados diferentes: farmacias, escuelas y abarrotes (gráficas 5, 6 y 7). En los tres últimos casos la función K_{plana} caracteriza los patrones como *aleatorios*, mientras que la función K_{red} los califica como *aglomerados*, ya que permite una estimación más fina de los patrones espaciales en espacios intraurba-

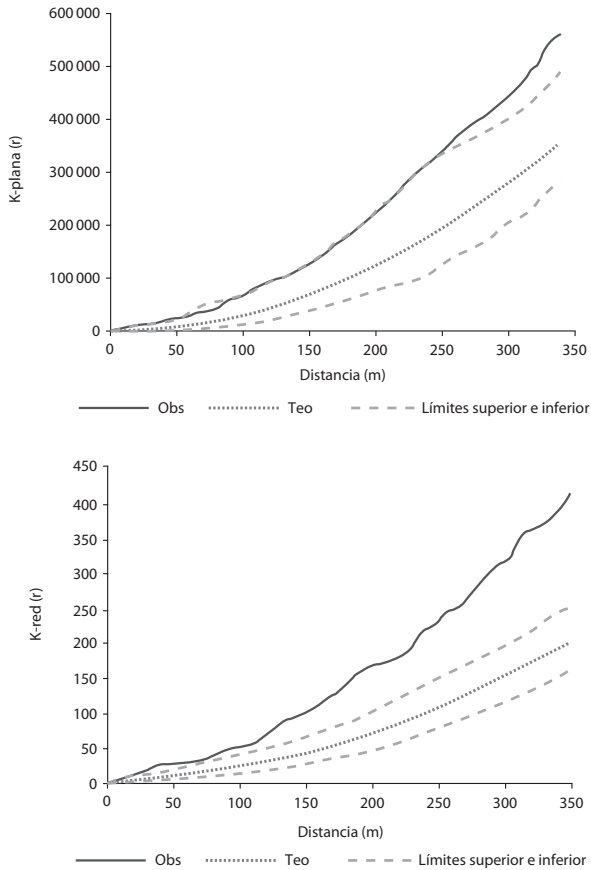
Gráfica 4. Comparación de la función K plana contra la función K red, giro seleccionado: mercerías



Fuente: Elaboración propia.

nos, donde la distancia en red es más significativa que la distancia euclidiana. Esto es particularmente importante en los *casos límite*, donde la función K_{plana} observada tiende a ubicarse cerca de los límites de los niveles de confianza. Para el caso de las escuelas, la función K_{plana} caracteriza su patrón como claramente aleatorio, pues prácticamente a lo largo de toda la distancia calculada está por debajo del límite superior del nivel de confianza, mientras que la función K_{red} lo identifica como ligeramente aglomerado a prácticamente todas las distancias. La estimación más fina

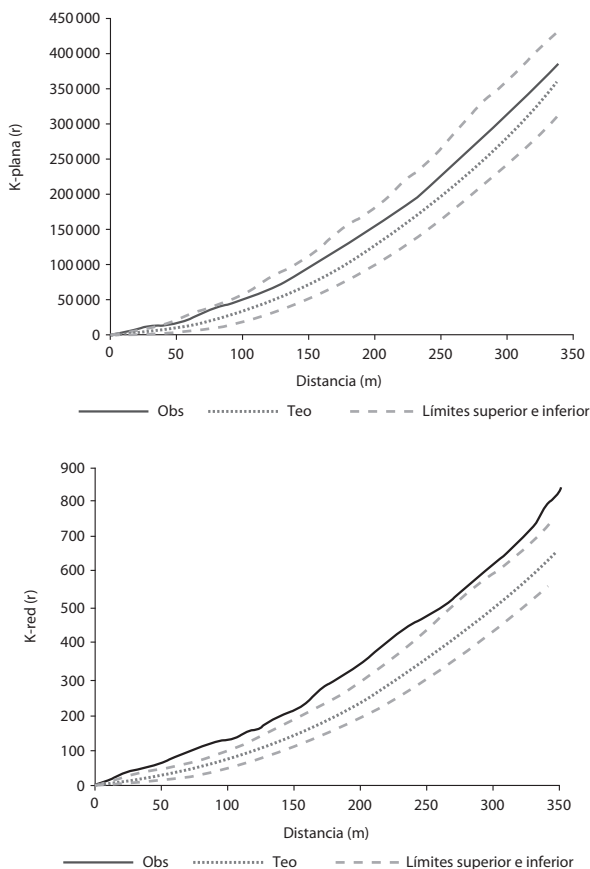
Gráfica 5. Comparación de la función K plana contra la función K red, giro seleccionado: farmacias



Fuente: Elaboración propia.

de la función K_{red} es aún más clara en el caso de las farmacias, ya que mientras la función K_{plana} observada se localiza muy cerca del límite superior de confianza, la función K_{red} se ubica claramente por encima del límite superior, con lo que caracteriza el patrón espacial como evidentemente aglomerado. La situación es distinta en el caso de los abarrotes, donde la función K_{plana} oscila por debajo del nivel superior, pero en K_{red} oscila por encima del nivel superior de confianza, con lo que declara una ligera aglomeración (gráficas 2 a 7).

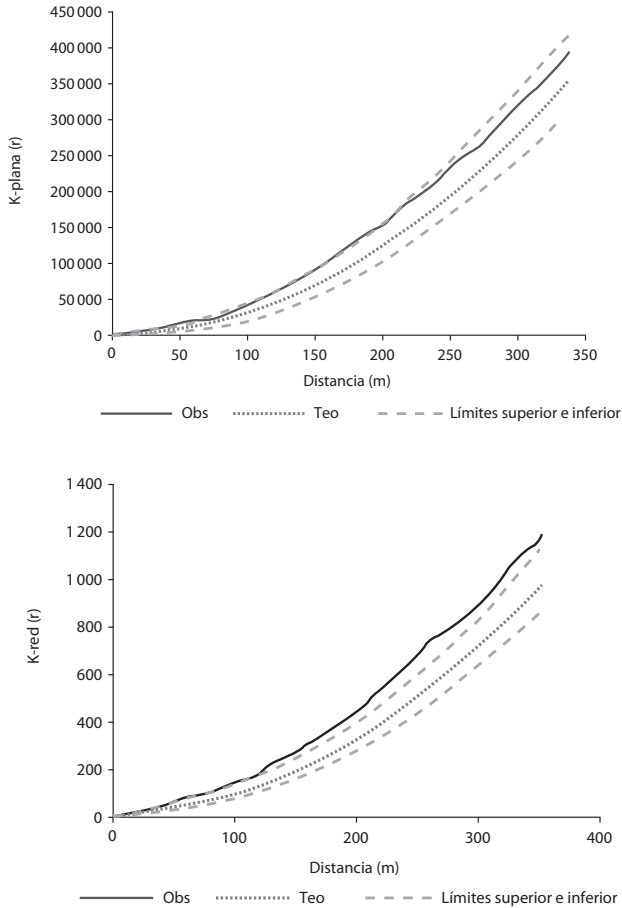
Gráfica 6. Comparación de la función K plana contra la función K red, giro seleccionado: escuelas



Fuente: Elaboración propia.

Por lo tanto, si bien en la mayoría de los casos las funciones K_{plana} y K_{red} coinciden en la caracterización de los patrones espaciales de las unidades económicas en la zona de estudio (aglomerado, disperso, aleatorio), en los *casos límite* (donde la función K_{plana} observada se localiza cerca de los límites de confianza) la función K_{red} parece ofrecer una caracterización más evidente y precisa de los patrones espaciales. Vale subrayar que en estos *casos límite* la diferencia en la caracterización de los patrones radica en calificarlos como aleatorios o como aglomerados/dispersos; es decir, se tra-

Gráfica 7. Comparación de la función K plana contra la función K red, giro seleccionado: abarrotos



Fuente: Elaboración propia.

ta de una diferencia entre los patrones topológicamente más próximos (*aleatorio-aglomerado*; *aleatorio-disperso*), pero no se detectaron casos en los que la caracterización de los patrones de las funciones K fueran diametralmente opuestas: *concentrado-disperso* o viceversa. En los *casos límite* de la zona de estudio, las diferencias de caracterización de los patrones espaciales son casi de matiz, y son mejor detectadas por la función K_{red} .

III.1.2. Variaciones territoriales de los patrones espaciales predominantes

Los patrones espaciales de las unidades económicas pueden ser aglomerados, dispersos o aleatorios a prácticamente cualquier escala (*i.e.* distancia), o pueden registrar variaciones en el territorio (*i.e.* ser aglomerados a una cierta escala, pero evidentemente dispersos o aleatorios a otra). A los primeros los denominamos patrones espaciales *puros*, y a los segundos, patrones espaciales *mixtos*. En la zona de estudio con la función K_{plana} se detectaron cuatro patrones espaciales mixtos (tres predominantemente aleatorios: abarrotes, escuelas y farmacias, y uno predominantemente aglomerado: bancos), mientras que en la función K_{red} se identificaron dos (uno predominantemente aleatorio: iglesias, y uno predominantemente aglomerado: abarrotes) (cuadro 2; gráficas 2 a 7).

Los patrones aleatorios mixtos detectados por la función K_{plana} son declarados como patrones puros por la función K_{red} , excepto el caso de los abarrotes, que se encuentra como un caso límite y en el que ambas funciones lo declararon como mixto, sólo que en la función K_{red} se presenta una conclusión de aglomeración y en la función K_{plana} de aleatoriedad. El caso de las iglesias es distinto: la función K_{red} lo declara como aleatorio mixto dado que existe una ligera aglomeración a corta distancia, situación que no se presenta en la función K_{plana} ; esto se debe al número relativamente pequeño de unidades localizadas, que obliga a tener una banda más amplia en los niveles de confianza. Por otro lado, los patrones *puros* detectados por la función K_{plana} son confirmados más evidentemente por la función K_{red} , por lo que esta última pareciera delinear una imagen más clara de las variaciones territoriales de los patrones espaciales de las unidades económicas en espacios intraurbanos con una estructura vial de alta densidad (algo similar concluyen Lu y Chen, 2006, cuando analizan el patrón espacial de robo de autos en San Antonio, Texas), al contrario de lo que se ha reportado para espacios regionales con baja densidad de carreteras (Yamada y Thill, 2004).

Aun cuando en la mayoría de los casos las funciones K_{plana} y K_{red} llegan a las mismas conclusiones acerca de la presencia de aglomeración espacial de unidades económicas en nuestra área de estudio, es evidente que ambas formas de cálculo no indican la misma *magnitud de concentración*, ni su presencia a las mismas *escalas* (distancias). Justamente, esto es lo que se explora en la siguiente sección.

III.2. Escala de la aglomeración / dispersión

Para determinar la distancia a la que se establece la máxima aglomeración (dispersión) de las unidades económicas consideradas, se calcularon dos indicadores que involucran relaciones entre el valor observado $K_{obs}(r)$ y el límite superior de su nivel de confianza $K_{ls}(r)$. El primero de ellos se obtiene de la ecuación 16, y permite identificar la distancia donde la aglomeración alcanza su valor máximo absoluto (Duranton y Overman, 2006).

$$M(r) = K_{obs}(r) - K_{ls}(r) \quad (16)$$

El cuadro 3 muestra los resultados del cálculo de la ecuación M cuando ésta llega a su máximo valor, tanto en la función K_{plana} como en la función K_{red} , y se identifican las coincidencias y divergencias.

Cuadro 3. Distancia de máxima aglomeración: ecuación M

a) Función K_{plana}		b) Función K_{red}	
Giro	Distancia	Giro	Distancia
Iglesias*	—	Iglesias	63
Escuelas	27	Abarrotes	258
Abarrotes	160	Farmacias	316
Alimentos preparados	251	Escuelas	337
Bancos	302	Alimentos preparados	337
Mercerías	309	Bancos	337
Farmacias	327	Mercerías	337
Electrónicos	334	Electrónicos	337
Restaurantes	335	Restaurantes	337
Mueblerías	337	Mueblerías	337
Ropa	337	Ropa	337
Zapaterías	337	Zapaterías	337

Fuente: Elaboración propia. *El patrón predominante es aleatorio; por lo tanto, no hay distancia de máxima aglomeración.

Como puede observarse, la función K_{plana} tiende a subestimar la distancia de máxima aglomeración absoluta, debido, sobre todo, a que estima la distancia entre las unidades económicas en línea recta (“a vuelo de pájaro”), mientras que la función K_{red} la estima a través de la red de vialidades. La excepción es el caso límite de las farmacias, donde la función K_{plana} registra la máxima aglomeración a una distancia mayor que la calculada en la función K_{red} .

Es importante resaltar que las distancias de máxima aglomeración se dan a la distancia máxima calculada por la función K , lo que es muy evidente sobre todo en la función K_{red} y sugiere que muy probablemente la aglomeración continúa más a allá de la distancia máxima calculada.

III.3. Intensidad de la concentración / dispersión de los patrones espaciales

Para determinar la intensidad del fenómeno de aglomeración (dispersión o rechazo espacial), ya sea en términos absolutos o relativos, se graficó el indicador $M(r)$ que muestra las magnitudes absolutas de la aglomeración. Sin embargo, $M(r)$ sólo permite la comparación entre las diversas distribuciones dentro de la función K_{plana} o K_{red} , pero no entre ellas; es por ello que se proponen dos índices de concentración comparables entre los dos métodos (el índice IC y el índice \overline{IC}).

IC calcula la proporción de exceso de unidades económicas (aglomeración) respecto al límite superior del nivel de confianza a una distancia determinada (Dixon, 2002), y está dada por la ecuación (17).

$$IC(r) = \frac{K_{obs}(r)}{K_{ls}(r)} \tag{17}$$

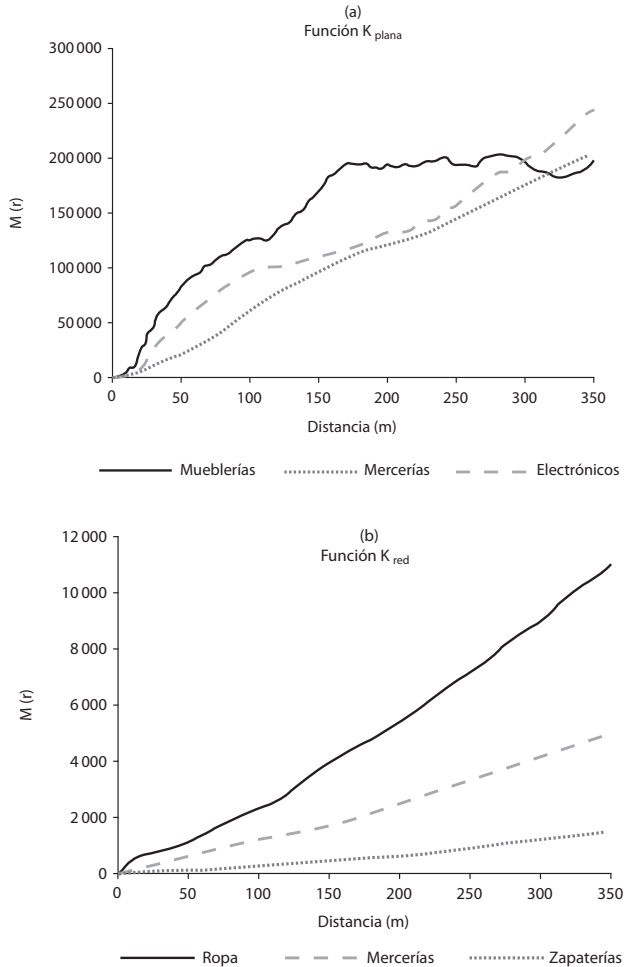
\overline{IC} es un índice de concentración promedio que relativiza la magnitud del área entre las curvas definidas por $K_{obs}(r)$ y $K_{ls}(r)$ respecto a $K_{ls}(r)$, y está dado por la ecuación (18).

$$\overline{IC}(r) = \frac{\sum_i (K_{obs}(r_i) - K_{ls}(r_i))}{\sum_i K_{ls}(r_i)} \tag{18}$$

Donde:

i : representa cada uno de los intervalos en los que se calculó la función $K(r)$.

Estos índices serán mayores que 1.0 si existe aglomeración, y menores que 1.0 si existe aleatoriedad o dispersión.

Gráfica 8. Intensidad de la concentración/dispersión: función M 

Fuente: Elaboración propia.

Así, al considerar la magnitud absoluta [indicador $M(r)$] para la función K_{plana} , son las unidades económicas de electrónicos, mercerías y mueblerías las que muestran mayor intensidad en su concentración espacial, mientras que para la función K_{red} la máxima aglomeración absoluta se observa en ropa, mercerías y zapaterías (gráficas 8a y 8b).

La función K_{red} resalta las aglomeraciones con un mayor número de establecimientos (310 para ropa, 167 para mercerías) con corredores espe-

cializados de mayor extensión que abarcan varias calles y manzanas, mientras que la función K_{plana} resalta las concentraciones que tienen una distribución espacial de corredor especializado más pequeño pero más concentrado (como el caso de electrónicos y mueblerías, cuyo corredor especializado en el área de estudio abarca apenas una calle que concentra cerca de 50 por ciento de los establecimientos). Esto pareciera indicar que la función K_{red} puede detectar de mejor manera las aglomeraciones a nivel de toda la zona de estudio, y la K_{plana} detecta las aglomeraciones a distancias más cortas (donde afecta menos considerar la distancia euclidiana).

Si se calcula el $IC(r)$ para las distancias de máxima aglomeración absoluta se obtienen los resultados que se muestran en el cuadro 4. Para el caso de la función K_{plana} , este índice no es mayor que 1.0 en el caso de las iglesias, mientras que los giros de mercerías, electrónicos, escuelas y zapaterías registran los índices más altos. Así, por ejemplo, para el caso de las

Cuadro 4. IC : Índice de concentración

a) Función K_{plana}		b) Función K_{red}	
Giro	$IC(r)$	Giro	$IC(r)$
Mercerías	1.79	Zapaterías	2.63
Electrónicos	1.69	Mercerías	2.48
Escuelas	1.63	Mueblerías	2.44
Zapaterías	1.63	Bancos	2.24
Mueblerías	1.60	Iglesias	2.00
Ropa	1.42	Ropa	1.99
Bancos	1.34	Farmacias	1.73
Restaurantes	1.20	Electrónicos	1.64
Farmacias	1.18	Restaurantes	1.59
Alimentos preparados	1.10	Alimentos preparados	1.18
Abarrotes	1.07	Abarrotes	1.15
Iglesias	—	Escuelas	1.07

Fuente: Elaboración propia. *El patrón predominante es aleatorio; por lo tanto, no hay distancia de máxima aglomeración.

mercerías un IC de 1.79 implica que existe 79 por ciento más unidades económicas de este giro que lo que se esperaría de una distribución derivada de un proceso de localización completamente aleatorio, si se consideran los niveles de confianza. La función K_{red} , nuevamente debido a la forma del cálculo de la distancia, estima mejor el valor de IC prácticamente en todos los giros, excepto en el de electrónicos y en el caso límite de las escuelas, donde la función K_{plana} detectó un pico de máxima aglomeración alto, cuando en realidad su patrón fue declarado como aleatorio o como de mínima aglomeración en el caso de K_{red} , y que como se ha resaltado, esta variación es explicable en razón del número relativamente pequeño de unidades localizadas en el área de estudio.

El $\overline{IC}(r)$ ofrece un panorama más claro y concluyente (cuadro 5); por un lado, declara como aleatorios/dispersos para el caso de la función K_{plana} los giros de iglesias, escuelas y abarrotes, mientras que en la función K_{red} sólo el de iglesias cae en esta categoría. En todos los casos, el indicador es ma-

Cuadro 5. \overline{IC} : Índice de concentración promedio

a) Función K_{plana}		b) Función K_{red}	
Giro	$\overline{IC}(r)$	Giro	$\overline{IC}(r)$
Iglesias	-0.6032	Iglesias	-0.0383
Escuelas	-0.1449	Abarrotes	0.0944
Abarrotes	-0.0457	Escuelas	0.1013
Farmacias	0.0411	Alimentos preparados	0.2333
Alimentos preparados	0.0847	Farmacias	0.5766
Restaurantes	0.2044	Restaurantes	0.5899
Bancos	0.2919	Bancos	1.1569
Ropa	0.6621	Ropa	1.2387
Zapaterías	0.8341	Zapaterías	1.7993
Electrónicos	1.0432	Mercerías	1.8839
Mueblerías	1.0825	Electrónicos	2.3516
Mercerías	1.1058	Mueblerías	2.4218

Fuente: Elaboración propia.

yor en la K_{red} que en la K_{plana} , con lo que se confirma que la función K_{plana} subestima los niveles de aglomeración respecto a K_{red} . En su jerarquía relativa la K_{plana} definió como más aglomerados los giros de mercerías, mueblerías y electrónicos, mientras que la función K_{red} define un orden distinto: mueblerías, electrónicos y mercerías.

Dado que la estructura vial del centro del AMT registra una alta densidad de calles, la diferencia relativa entre la distancia en red y la distancia euclidiana es menor que si el análisis se realizara a escala regional usando la estructura carretera (Lu y Chen, 2006). No obstante, aun en un espacio intraurbano con alta densidad de vialidades, la función K_{plana} tiende a subestimar la caracterización de aglomeración respecto a la función K_{red} , como ocurre en nuestra área de estudio. Adicionalmente, el hecho de que la distancia de máxima aglomeración se obtenga a menores valores en la K_{plana} que en la K_{red} confirma que la distancia euclidiana es significativamente más corta que la distancia en red, por lo que la función K_{red} refleja más fielmente los patrones observados de aglomeración de los establecimientos. Es así, sobre todo, para el centro del AMT, donde la actividad económica se observa principalmente en corredores comerciales de forma irregular (*i.e.* cruces de calles o pequeñas calles especializadas).

Las posiciones relativas del $\bar{IC}(r)$ calculado con ambos métodos muestran pocas diferencias, siendo el giro de mercerías el que presenta, en todo caso, la mayor diferencia debido al elevado número de unidades económicas concentradas en un corredor especializado muy definido del CTN del AMT.

IV. Conclusiones

El propósito de este trabajo fue analizar las coincidencias y divergencias de los resultados de las funciones K_{plana} y K_{red} en la *caracterización* (aglomeración/dispersión o rechazo espacial/aleatoriedad), *escala* (localización del pico de máxima intensidad de la aglomeración/dispersión) e *intensidad* (más o menos unidades económicas respecto a una distribución en condiciones de CSR) de los patrones de localización de unidades económicas terciarias en el CTN del AMT (México). Para el cálculo de las funciones K_{plana} y K_{red} se utilizaron como herramientas computacionales *Spatstat* y *SANET*, respectivamente. Con estas herramientas fue posible manejar sin problema el área de estudio considerada, organizar y procesar la información, y determinar la caracterización, escala e intensidad de la aglomeración/dispersión de los patrones espaciales de las unidades económicas en el espacio intraurbano.

Contra lo que reporta la literatura especializada (Yamada y Thill, 2004), para la zona de estudio seleccionada la función K_{plana} no sobreestimó el patrón de aglomeración de las actividades terciarias, debido a la alta densidad de la estructura vial en la zona y a su estructura comercial. Más bien ocurrió lo contrario: aun en un espacio intraurbano con alta densidad de vialidades, como el CTN del AMT, la función K_{plana} subestimó los patrones de aglomeración. Cabe mencionar que la estructura comercial del CTN del AMT se articula básicamente a través de corredores especializados para los giros que resultaron ser más aglomerados. En estas circunstancias la función K_{red} los detectó más claramente, por lo que es más recomendable su uso a escala intraurbana y para este tipo de estructura comercial. Sin duda, la función K_{red} fue más concluyente y más fina en sus apreciaciones de aglomeración, mientras que la función K_{plana} tuvo resultados mixtos o en los límites de los niveles de confianza.

Se debe subrayar que el número de establecimientos considerados para cada giro tiene un impacto en la generación de las simulaciones para determinar los niveles de confianza, y es por ello que, en el caso específico de este trabajo, para giros como iglesias, farmacias y bancos, en los que se ubicó un número relativamente pequeño de establecimientos en el área de estudio, fue más difícil declarar un patrón de aglomeración o aleatoriedad en cada uno de los métodos considerados. Así, en el espacio intraurbano, mientras mayor sea el número de puntos considerados, mayor será la coincidencia en la caracterización de los patrones de las funciones K_{plana} y K_{red} , porque las bandas de los niveles de confianza tienden a acercarse a la función K teórica.

El hecho de que en la mayoría de los casos el máximo de aglomeración absoluto se obtenga a la distancia máxima a la que se calculó la función K sugiere que la aglomeración continúa más allá de la distancia calculada, por lo que es posible, para caracterizar patrones espaciales básicos de firmas terciarias en el espacio intraurbano, incrementar el límite más allá del valor sugerido por el software utilizado para la función K_{plana} .

En esta investigación se propusieron y calcularon los indicadores M , IC e \overline{IC} para determinar la *escala* e *intensidad* de las aglomeraciones de una manera más precisa y evaluable que la sola interpretación de las gráficas resultantes de las funciones K . Estos indicadores son complementarios, pero sólo los índices IC e \overline{IC} son comparables entre K_{plana} y K_{red} , ya que el indicador M está dado en unidades bidimensionales para el caso de K_{plana} y unidimensionales para K_{red} . En contraparte, los índices IC e \overline{IC} relativizan las funciones K , lo que facilita notablemente su comparación. El indicador

\overline{IC} , en particular, puede considerarse como un indicador cuantitativo más concluyente para determinar el patrón espacial predominante que la simple inspección visual.

El trabajo empírico realizado con ambos métodos tendrá que complementarse para otros giros comerciales y de servicios, y para otros espacios intraurbanos en los que se presente una estructura vial distinta o bien una estructura comercial más nucleada; por ejemplo, plazas y centros comerciales; pero en cualquier caso, a una escala intraurbana la función K_{red} parece ser el método más adecuado para medir los niveles de aglomeración, su escala y su intensidad. Las herramientas computacionales disponibles permiten realizar todos los cálculos asociados sin grandes complicaciones, lo que le facilita al analista concentrarse en el estudio de los resultados y no en el desarrollo matemático.

Referencias bibliográficas

- Albert, J. M., J. Mateu y V. Orts (2007), *Distribución espacial de la actividad económica en la Unión Europea*, WP-EC 2007-02, Valencia, Instituto Valenciano de Investigaciones Económicas.
- Anselin, L. (1992), *Spatial Data Analysis with GIS: An Introduction to Application in the Social Sciences*, technical report 92-10, agosto, Santa Barbara, University of California, National Center for Geographic Information and Analysis.
- _____ (1995), "Local Indicators of Spatial Association (LISA)", *Geographical Analysis*, 27 (1), pp. 93-115.
- Arbia, G. (2001), "Modelling the Geography of Economic Activities in a Continuous Space", *Papers in Regional Sciences*, 80 (4), pp. 411-424.
- Arbia, G., G. Espa y D. Quah (2007), "A Class of Spatial Econometric Methods in the Empirical Analysis of Clusters of Firms in the Space", discussion paper 5, Universidad de Trento, Departamento de Economía.
- Baddeley, A. y R. Turner (2005), "Spatstat: An R Package for Analyzing Spatial Point Patterns", *Journal of Statistical Software*, 12 (6), pp. 1-42.
- Bailey, T. C. y A. C. Gatrell (1995), *Interactive Spatial Data Analysis*, Longman, Harlow, Reino Unido.
- Bertinelli, L. y J. Decrop (2005), "Geographical Agglomeration: Ellison and Glaeser's Index Applied to the Case of Belgian Manufacturing Industry", *Regional Studies*, 39 (5), pp. 567-583.
- Clark, P. J. y F. C. Evans (1954), "Distance to Nearest Neighbor as a Measure of Spatial Relationships in Populations", *Ecology*, 35 (4), pp. 445-453.

- Cressie, N. (1993), *Statistics for Spatial Data*, Nueva York, John Wiley & Sons.
- Cuthbert, A. L. y W. P. Anderson (2004), "Using Spatial Statistics to Examine the Pattern of Urban Land Development", *The Professional Geographer*, 54 (4), pp. 521-532.
- Devereux, M. P., R. Griffith y H. Simpson (2004), "The Geographic Distribution of Production Activity in the UK", *Regional Science and Urban Economics*, 34 (5), pp. 533-564.
- Diggle, P. J. (2003), *Statistical Analysis of Spatial Point Patterns*, 2a. ed., Londres y Nueva York, Oxford University Press.
- Dixon, P. M. (2002), "Ripley's K Function", en A. H. El-Shaarawi y W. W. Piegorsch, *Encyclopedia of Environmetrics*, 3, Chichester, John Wiley & Sons, pp. 1796-1803.
- Duranton, G. y H. G. Overman (2005), "Testing for Localization Using Micro-Geographic Data", *Review of Economic Studies*, 72 (4), pp. 1077-1106.
- _____ (2006), "Exploring the Detailed Location Patterns of UK Manufacturing Industries Using Microgeographic Data", discussion paper 756, octubre, Londres, Centre for Economic Performance, London School for Economics and Political Science.
- Ellison, G. y E. L. Glaeser (1997), "Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach", *Journal of Political Economy*, 105 (5), pp. 889-927.
- Fujita, M. (1988), "A Monopolistic Competition Model of Spatial Agglomeration: A Differentiated Product Approach", *Regional Science and Urban Economics*, 18 (1), pp. 87-124.
- Fujita, M. y J.-F. Thisse (2002), *The Economics of Agglomeration*, Cambridge, Cambridge University Press.
- Fujita, M. y P. Krugman (2004), "The New Economic Geography: Past, Present and the Future", *Papers in Regional Science*, 83 (1), pp. 139-164.
- Fujita, M., P. Krugman y A. Venables (1999), *The Spatial Economy*, Boston, The MIT Press.
- Garrocho, C. y J. Campos (2007), "Dinámica de la estructura policéntrica del empleo terciario en el área metropolitana de Toluca, 1994-2004", *Papeles de Población*, 13 (52), abril-junio, pp. 109-136.
- Garrocho, C. y Z. Flores (2009a), "Delimitación del centro tradicional de comercio y servicios del área metropolitana de Toluca", *Papeles de Población*, 15 (61), pp. 233-274.

- _____ (2009b), "Metodología para delimitar centros tradicionales de negocios en ciudades mexicanas a partir de la intensidad de flujos peatonales", en D. Velázquez, M. E. Orozco y L. Castillo (coords.), *Ciudades, urbanización y metropolización*, México, Universidad Autónoma de Quintan Roo-Universidad Autónoma del Estado de México-Plaza y Valdés, pp. 11-46.
- Gatrell, A. C., T. C. Bailey, P. J. Diggle y B. S. Rowlingson (1996), "Spatial Point Pattern Analysis and its Application in Geographical Epidemiology", *Transactions of the Institute of British Geographers*, 21 (1), pp. 256-274.
- Getis, A. y J. Franklin (1987), "Second-order Neighborhood Analysis of Mapped Point Patterns", *Ecology*, 68 (3), pp. 473-477.
- Glaeser, E. L., H. D. Kallal, J. A. Scheinkman y A. Shleifer (1992), "Growth in Cities", *The Journal of Political Economy*, 100 (6), pp. 1126-1152.
- Guillain, R. y J. Le Gallo (2007), *Agglomeration and Dispersion of Economic Activities in Paris and its Surroundings: An Exploratory Spatial Data Analysis*, Regional Economics Applications Laboratory Discussion Paper 06-T-10.
- Holmes, T. J. y J. J. Stevens (2004), "Spatial Distribution of Economic Activities in North America", en V. Henderson y J.-F. Thisse (eds.), *Handbook of Urban and Regional Economics: Cities and Geography*, Amsterdam, Elsevier.
- Kosfeld, R., H.-F. Eckey y J. Lauridsen (2009), *Spatial Point Pattern Analysis and Industry Concentration*, MAGKS Papers on Economics, Philipps-Universität Marburg, Marburgo, Alemania, Facultad de Administración de Negocios, Departamento de Economía.
- Krugman, P. (1991a), "Increasing Returns and Economic Geography", *The Journal of Political Economy*, 99 (3), pp. 483-499.
- _____ (1991b), *Geography and Trade*, Cambridge, The MIT Press.
- Lee, Y. (1979), "A Nearest-Neighbor Spatial-Association Measure for the Analysis of Firm Interdependence", *Environment and Planning A*, 11 (2), pp. 169-176.
- Levine, N. (2004), *CrimeStat: A Spatial Statistics Program for the Analysis of Crime Incident Locations*, v.3.0, Houston y Washington, Ned Levine & Associates y The National Institute of Justice.
- Lu, Y. y X. Chen (2006), "On the False Alarm of Planar K-function when Analyzing Urban Crime Distributed along Streets", *Social Science Research*, 36 (2), pp. 611-632.
- Manly, B. F. J. (1997), *Randomization, Bootstrap and Monte Carlo Methods in Biology*, 2a. ed., Londres, Chapman.

- Marcon, E. y F. Puech (2003), "Evaluating the Geographic Concentration of Industries Using Distance-Based Methods", *Journal of Economic Geography*, 3 (4), pp. 409-428.
- Marshall, A. (1920), *Principles of Economics: An Introductory Volume*, 8a. ed., Londres, MacMillan.
- Maurel, F. y B. Sédillot (1999), "A Measure of the Geographic Concentration in French Manufacturing Industries", *Regional Science and Urban Economics*, 29 (5), pp. 575-604.
- Mella, P. (2006), "Spatial Co-Location of Firms and Entrepreneurial Dynamics: The Combinatory Systems View", *Entrepreneurship Management*, 2 (3), pp. 391-412.
- Miller, H. J. (1999), "Potential Contribution of Spatial Analysis to Geographic Information Systems for Transportation (GIS-T)", *Geographical Analysis*, 31 (4), pp. 373-399.
- Mitchell, A. (2005), *The ESRI Guide to GIS Analysis, vol. 2: Spatial Measurement & Statistics*, Redlands, California, ESRI Press.
- Mori, T., K. Nishikimi y T. E. Smith (2005), "A Divergence Statistic for Industrial Localization", *The Review of Economics and Statistics*, 87 (4), pp. 635-651.
- Myint, S. W. (2008), "An Exploration of Spatial Dispersion, Pattern, and Association of Socio-Economic Functional Units in an Urban System", *Applied Geography*, 28 (3), pp. 168-188.
- Okabe, A., H. Yomono y M. Kitamura (1995), "Statistical Analysis of the Distribution of Points on a Network", *Geographical Analysis*, 27 (2), pp. 152-175.
- Okabe, A. e I. Yamada (2001), "The K-Function Method on a Network and its Computational Implementation", *Geographical Analysis*, 33 (3), pp. 271-290.
- Okabe, A., K. Okunuki y S. Zooide (2007), *SANET: A Toolbox for Spatial Analysis on a Network*, Version 3.3-030907, Tokio, Universidad de Tokio, Center for Spatial Information Science.
- Okabe, A. y M. Kitamura (1996), "A Computational Methods for Optimizing the Location of Store on a Network", *Geographical Analysis*, 28 (4), pp. 330-349.
- Okabe, A. y T. Satoh (2006), "Uniform Network Transformation for Points Pattern Analysis on a Non-Uniform Network", *Journal of Geographical Systems*, 8 (1), pp. 25-37.
- Openshaw, S. (1984), *The Modifiable Areal Unit Problem: Concepts and Techniques in Modern Geography*, Norwich, Reino Unido, Geo Books.

- O'Sullivan, A. (2007), *Urban economics*, Boston, McGraw-Hill.
- Ottaviano, G. y J.-F. Thisse (2003), "Agglomeration and Economic Geography", en J. Vernon Henderson y J.-F. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4, Ámsterdam, Elsevier, pp. 2563-2608.
- Quah, D. y H. Simpson (2003), *Spatial Cluster Empirics*, trabajo en curso, London School of Economics, Economics Department e Institute for Fiscal Studies, junio.
- Ripley, B. D. (1976), "The Second-Order Analysis of Stationary Point process", *Journal of Applied Probability*, 13 (1), pp. 255-266.
- _____ (1981), *Spatial Statistics*, Chichester, Brisbane, Toronto, Nueva York, John Wiley & Sons.
- _____ (1988), *Statistical Inference for Spatial Processes*, Cambridge, Cambridge University Press.
- Robinson, W. S. (1950), "Ecological Correlations and the Behavior of Individuals", *American Sociological Review*, 15 (3), pp. 351-357.
- Rosenthal, Stuart S. y William C. Strange (2004), "Evidence on the Nature and Sources of Agglomeration Economies", en J. V. Henderson y J. Thisse (eds.), *Handbook of Regional and Urban Economics*, vol. 4, Ámsterdam, disponible en: <http://www.econ.brown.edu/faculty/henderson/willandstuart.pdf> [fecha de consulta: septiembre de 2011].
- Rysman, M. y S. Greenstein (2005), "Testing for Agglomeration and Dispersion", *Economics Letters*, 86 (3), pp. 405-411.
- Smith, T. E. (2009), Notebook on Spatial Data Analysis, libro electrónico, pp. I.4.9-I.4.11, disponible en: <http://www.seas.upenn.edu/~ese502/#notebook> [fecha de consulta: septiembre de 2009].
- Stoyan, D. y H. Stoyan (1994), *Fractals, Random Shapes and Point Fields: Methods of Geometrical Statistics*, Chichester, Wiley.
- Venables, A. J. (1996), "Equilibrium Locations of Vertically Linked Industries", *International Economic Review*, 37 (2), pp. 341-359.
- Xu, J. (2002), "Detect Seasonal Pattern Change of DVAs in the Town of Amherst Using Network K-function", documento y ponencia presentados en el 2002 Summer Assembly of the University Consortium of Geographic Information Science, 26-30 de junio, Athens, Georgia.
- Yamada, I. y J.-C. Thill (2004), "Comparison of Planar and Network K-Functions in Traffic Accident Analysis", *Journal of Transport Geography*, 12 (2), pp. 149-158.