

USOS DE INFORMACIÓN GEORREFERENCIADA PARA PRESTACIÓN DE SERVICIOS A LA POBLACIÓN: UNA REVISIÓN DE LITERATURA

SANDRA LILIANA MORENO MAYORGA, JAVIER MAURICIO JÁCOME MOLINA¹

¹ Departamento Administrativo Nacional de Estadística -DANE

Cra. 59 #26-60, Bogotá, Colombia.

s.lilianamoreno@gmail.com, javierjacome1@gmail.com

RESUMEN

El creciente volumen de información estadística georreferenciada es una oportunidad para mejorar la prestación de servicios a la población desde una perspectiva territorial. Sin embargo, el aprovechamiento de los datos requiere la apropiación de metodologías adecuadas. Este documento sintetiza algunas de ellas, alrededor de los cinco conceptos fundamentales del análisis espacial enunciados por Buzai (2010): localización, distribución, asociación, interacción y evolución. Para cada una de estos se presentan varias técnicas que permiten su aplicación para mejorar la prestación de servicios a la población.

Palabras clave: análisis espacial, SIG, asociación espacial, interacción espacial, análisis espacio - temporal.

USING GEOREFERENCED INFORMATION TO SUPPORT THE SERVICES PROVISION FOR POPULATION: A LITERATURE REVIEW

ABSTRACT

The growing volume of geo-referenced statistical data it is an opportunity to improve the provision of services to the population considering a territorial perspective. However, to leverage data is required the use of adequate methodologies. This paper summarizes some of them around of the five basic concepts of the spatial analysis stated by Buzai (2010): localization, distribution, association, interaction and evolution. For each one, are presented techniques that allow the improvement in the services provision for population.

Keywords: Spatial analysis, GIS, spatial association, spatial interaction, spatial-temporal analysis.

Recibido: 02/06/2017

Aceptada versión definitiva: 16/11/2017

Editor al cargo: Dr. Ismael Vallejo

Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0)

© Los autores
www.geo-focus.org

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

1. Introducción

Desde hace varias décadas la geografía ha desarrollado metodologías para mejorar la prestación de servicios a la población. Obtener los insumos para implementarlas es más fácil ahora que hace algún tiempo dado que la geografía, igual que otras disciplinas, se ha movido de un ambiente pobre en datos y capacidad de cómputo a otro rico en estos. Como lo expone Goodchild, (2015), las tecnologías de la información han permitido como nunca antes que más personas tengan acceso a sistemas de posicionamiento global desde sus celulares, puedan consultar datos de observación de la tierra de alta resolución espacial y temporal y manipulen robustos sistemas de información geográfica vía web; de tal manera que los usuarios no solo consultan información espacial sino que también generan datos geoespaciales, lo que brinda oportunidades sin precedentes para el aprovechamiento del análisis espacial.

A pesar de todo, muchas organizaciones y gobiernos no han aprovechado los beneficios de este incremento exponencial en el volumen de datos geoespaciales, debido a la carencia de recursos, conocimientos, capacidad y oportunidad (IEAG, 2014). Por tanto, se debe propender por maximizar las oportunidades de las fuentes de información geoespacial para estudiar, interpretar y difundir la información geoespacial y apoyar la toma de decisiones y la presentación de servicios a la población.

Moreno Jiménez (1995) realiza un listado de problemas que habría que superar para realizar una transferencia efectiva de innovaciones geográficas por parte de las organizaciones productoras de información. Estos son: obtener productos de conocimiento útil para la toma de decisiones, preparar dichos productos de tal forma que sean más asimilables por sus destinatarios, identificar y localizar los agentes y organizaciones capaces de asumir e impulsar acciones acordes con este conocimiento, difundir dicho conocimiento por los canales más pertinentes, disponer de un marco normativo adecuado, e implementar las soluciones geográficas mediante mecanismos de gestión eficaces.

Desde el punto de vista de la generación de información estadística georreferenciada el problema de la lenta difusión tiene una connotación especial, ya que genera una precaria apropiación del conocimiento geográfico, lo que conlleva a que la demanda de información estadística sea más baja y menos cualificada de su potencial. Esto no solo implica una subutilización de las estadísticas georreferenciadas disponibles, sino la pérdida de un impulso externo necesario para avanzar en nuevos desarrollos.

El presente trabajo tiene la intención de ofrecer algunas ideas de cómo puede ser utilizada la información estadística georreferenciada para mejorar la prestación de servicios a la población. Para hacerlo se ha organizado alrededor de los cinco conceptos fundamentales del análisis espacial descritos por Buzai (2010): Localización, Distribución, Interacción, Asociación y Evolución, los cuales concuerdan con los abordados por (Nyerges, 1991). Para cada una de estos conceptos se presenta un conjunto de técnicas que permiten su aplicación para mejorar la prestación de servicios a la población.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, n° 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

2. Localización

De acuerdo a la Academia Nacional Ciencias de los Estados Unidos, los objetos científicos de diferentes dominios del conocimiento tienen unas propiedades comunes: 1. la identidad o nombre; 2. la localización en el espacio; 3. la magnitud; 4. la duración (National Academy of Sciences, 2006). Entre estas cuatro propiedades, únicamente la localización es directamente espacial, lo que significa que para aplicarles un tratamiento geográfico previamente los datos deben ser localizados.

Los objetos científicos con propiedades de localización son conocidos como geodatos. Moreno, Rodríguez, Jiménez y Martínez (1986) conceptualizan el geodato como “Un dato en el que un componente describe el fenómeno (componente temático), y otro componente lo localiza en el territorio (componente geométrico)”. Es importante precisar que estos autores consideran que dicha localización puede ser implícita (nombre del municipio, número de la dirección censal, calle y número de casa, etc.) o explícita como en las coordenadas.

Diversas Instituciones públicas y privadas han impulsado la divulgación y el acceso libre a geodatos a nivel global sobre diversas temáticas. Un referente internacional importante son las recomendaciones del Comité de Expertos sobre Gestión Global de Información Geoespacial de Naciones Unidas a los Institutos Estadísticos Nacionales (United Nations, 2015).

Entre los muchos ejemplos de iniciativas de acceso a geodatos puede mencionarse la Administración Nacional de la Aeronáutica y del Espacio -NASA- que través del Servicio Geológico de los Estados Unidos permite la descarga de imágenes satelitales Landsat1, con una serie histórica desde 1972. También la Agencia Espacial Europea -ESA- la cual permite la descarga libre de imágenes ópticas y de radar². Estos datos han sido utilizados en diversos estudios y han permitido generar productos geoespaciales derivados, tales como mapas de cobertura del suelo, estudios de cambios en la cobertura, monitoreo de clima, recursos marinos, entre otros. Finalmente, la iniciativa colaborativa de Natural Earth apoyada por la Sociedad de Información Cartográfica de América del Norte -NACIS dispone, de manera pública y para cualquier uso, de un conjunto de datos geográficos globales, sobre características físicas, culturales y topográficas a escala 1:100 000, 1:500 000 y 1:1 000 000.

En términos de información social, destacan las iniciativas globales que buscan brindar información sobre conteo de población y densidad poblacional en superficies continuas a escalas detalladas. NASA dispone al público la Grilla de Población del Mundo -GPW-, que en su cuarta versión incluye los años 1995, 2000, 2005, 2015 y 2020 a una resolución espacial de 1 km². En Europa, el proyecto GEOSTAT ha publicado la grilla de población de Europa 2011 a una resolución de 1 km² la cual puede ser utilizada por el público en general para propósitos no comerciales. World Pop también ha generado estimaciones de la distribución de la población globales en grillas de 100 m² utilizando como datos de referencia censos nacionales, imágenes satelitales y datos geográficos globales.

¹ <https://earthexplorer.usgs.gov/>

² <https://scihub.copernicus.eu/s3/#/home>

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, n° 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

En Colombia vale la pena resaltar la política de datos abiertos enmarcada dentro de la estrategia “Gobierno en Línea” (Presidencia de la República, 2000); la infraestructura de datos espaciales colombiana, que ha supuesto un avance significativo en la definición de políticas, estándares, metadatos y mecanismos de acceso a la información (Consejo Nacional de Política Económica y Social, 2009); los esfuerzos de la Infraestructura de Datos Espaciales de Bogotá – IDECA- que dispone al público de manera gratuita la información geográfica oficial de Bogotá ; y la publicación, por parte del DANE, de la cartografía del Marco Geoestadístico Nacional. A continuación se exponen otros mecanismos con los que se puede obtener información estadística georreferenciada.

2.1. Georreferenciación

La mayoría de sistemas de recolección y organización de datos no definen un componente espacial explícito. Por el contrario, los datos son recogidos en función de metodologías previamente definidas en cada campo de conocimiento, sin considerar explícitamente la propiedad de localización. Por otra parte, es común que estas bases de datos tengan una localización implícita, ya sea el nombre de un lugar o una dirección. La conversión de estos datos en geodatos es relativamente sencilla para el caso de unidades administrativas como los municipios, pero se hace más compleja si se requiere un mayor nivel de detalle. En este sentido los sistemas geocodificadores de direcciones son de gran utilidad para convertir datos en geodatos.

Machado (2008) define la geocodificación o geolocalización como la acción de atribuir coordenadas geográficas a uno o varios eventos, utilizando como referencia su dirección postal o una lista de referencias toponímicas o *gazetteers*. Por su parte, Google Code (2009) la define como la traducción de información disponible en un lenguaje natural a coordenadas geográficas, de la misma manera, el proceso de geocodificación inversa se referiría a la traducción de una coordenada geográfica en una dirección interpretable por humanos.

De acuerdo a López (2005) una dirección define una localización de la misma forma en que lo hace una coordenada geográfica, sin embargo, a diferencia de estas últimas que emplean sistemas numéricos, las direcciones emplean cadenas de texto que contienen información sobre el número de calle, el número del predio, el código postal, entre otros. Cuando la georreferenciación se realiza a través de *gazetteers*, estos aportan al menos tres elementos para cada entidad: un nombre, una localización y una clasificación. Esta última es un criterio de orden que permite desempatar cuando hay entidades con nombres iguales (Mostern, Southall, & Berman, 2016).

Por otro lado, los geocodificadores de direcciones se componen de tres grandes elementos: el geocodificador de direcciones, un software SIG y la base de datos cartográfica digital de ejes viales. Las direcciones que se introducen en el software deben estar previamente estandarizadas. Como resultado se obtienen las coordenadas de un subconjunto de las direcciones, y valores que indican la probabilidad de que un punto en particular esté correctamente georreferenciado. Resulta evidente que en los casos en que la probabilidad arrojada por el software sea baja es necesario realizar una revisión manual. Aunque una parte de los resultados deben revisarse manualmente, la geocodificación es una herramienta valiosa para obtener geodatos; los registros administrativos pueden ser susceptibles a ser geocodificados si se dispone de un código postal o dirección, lo que

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

permitiría disponer de datos geográficos de diversas temáticas a un alto nivel de desagregación espacial.

Las aplicaciones de la georreferenciación son múltiples. Moreno Jiménez et al., (1986) resaltan sus ventajas en términos de: 1. Relocalización de la información territorial entendida como agregación y síntesis de datos; 2. Elaboración de cartografía automatizada; 3. Análisis estadístico. A lo anterior habría que agregar la posibilidad de realizar tratamiento masivo a los datos.

2.2. Recolección de datos por medio de cartografía colaborativa

La cartografía colaborativa es una tendencia en la que ciudadanos corrientes consultan, crean y comparten información espacial (Turner, 2006). Esto ha sido posible gracias a los avances tecnológicos en comunicaciones, el surgimiento de la geoweb 2.0 con los denominados globos digitales (Google Earth, Microsoft Virtual Earth, Google Maps, Bing Maps, OpenStreetMap – OSM-), y el aumento en el acceso a dispositivos móviles con sistemas de posicionamiento global (Dodge, McDerby, & Turner, 2008; Mateos, 2013).

En la base del funcionamiento de esta herramienta está un concepto llamado *crowdsourcing*. Este término guarda cierta analogía con el de *outsourcing* en el que las tareas de una empresa son transferidas a localizaciones remotas. En el *crowdsourcing* la idea es que grandes grupos de personas puedan realizar tareas que son difíciles de automatizar o de implementar (Haklay & Weber, 2008).

Uno de los ejemplos más exitosos de este enfoque es el de openstreetmap –OSM-. Similar al de colaboración por pares de Wikipedia, cada persona registrada puede editar la información disponible o crear información nueva. OSM soporta la edición de la cartografía disponible por parte de usuarios registrados. Adicionalmente, el sitio ha agregado información cartográfica donada o que previamente era de acceso público. Como resultado, se ha obtenido una base de datos que cubre buena parte del planeta.

De acuerdo a Barron, Neis, y Zipf (2014), las evaluaciones de calidad de la información de esta fuente han mostrado, en términos generales, “una alta precisión posicional y una amplia cantidad de detalle alrededor de áreas urbanas con un alto número de contribuidores. En contraste, la mayoría de las áreas rurales a menudo tienen un bajo nivel de calidad de los datos para OSM”. Esto lleva a pensar que el éxito de este enfoque depende, en buena medida, del número y el nivel de actividad de los contribuidores.

También vale la pena considerar la experiencia del DENU. Esta sigla hace referencia al Directorio Estadístico Nacional de Unidades Económicas, construido por el Instituto de Estadística y Geografía de México –INEGI-. En el DENU la actualización no está abierta a todos los colaboradores. Los datos de una unidad económica únicamente pueden ser modificados por sus propietarios, representantes legales o personal autorizado por el negocio, previo registro y validación de identidad.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, n° 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

En este punto es necesario mencionar los incentivos de los informantes para registrarse y actualizar. El principal es que el DENUe funciona como un directorio comercial georreferenciado, con más de 5 millones de establecimientos, de libre acceso al público. Para actualizar la información comercial previamente deben actualizarse las variables básicas del Directorio. Como mecanismo para garantizar la calidad de los datos, la información reportada por los informantes es verificada por funcionarios del INEGI (Instituto Nacional de Estadística y Geografía -INEGI, s.f.)

Otro mecanismo para incentivar el uso del directorio es que funciona como un aplicativo de código abierto. De esta forma puede integrarse con otras aplicaciones, lo que aumenta la relevancia del registro. Lo anterior, sin perjuicio de que los datos también pueden ser descargados de forma masiva.

Finalmente, vale la pena considerar el concepto de geointeligencia colectiva. Se trata de la recolección de información situacional por medio de redes sociales. Este enfoque fue implementado en la ciudad de Peta Yakarta, Indonesia. Este proyecto, desarrollado por la Universidad de Wollongong, en Colaboración con la Agencia de Gestión de Emergencias de Jakarta y Twiter Inc, busca mejorar la resiliencia de esta ciudad a las inundaciones. Vale la pena resaltar que está desarrollado con base en la plataforma abierta "Cognicity".

En este contexto, la inteligencia colectiva está orientada a extraer valor de la información producida desde las redes sociales, filtrando, incentivando reportes valiosos y desplegando información de una manera que permita a los usuarios, gobiernos y otras agencias tomar decisiones en momentos críticos (Holderness & Turpin, 2015).

El aplicativo detecta, en primer lugar, tweets con las palabras clave *flood* y *banjir*. Un segundo paso es enviar un mensaje a estos usuarios, pidiendo que envíen un mensaje georreferenciado confirmando si hay una inundación en su posición. Esta información se publica en la página web del proyecto y también, parcialmente y en forma de mapa, en la cuenta de Twitter del informante. Esta aproximación podría usarse en el ámbito público con diversos propósitos, por ejemplo para georreferenciar la ocurrencia de delitos, o la calidad de prestadores de servicios públicos.

2.3. Tecnologías conscientes de la localización

Gracias a los dispositivos de recolección de datos vinculados a tecnologías "conscientes de la localización" (TCL) tales como receptores GPS, se están recolectando cantidades significativas de geodatos mediante investigaciones de campo que antes no producían un componente espacial explícito. Esto es cierto no sólo para investigaciones estadísticas, sino también para registros administrativos e información de empresas privadas. Un ejemplo de lo último son los dispositivos que utilizan los distribuidores de alimentos en pequeños establecimientos para registrar los pedidos. Por otra parte, una característica distintiva de la época actual es que los geodatos se producen por diversos tipos de sensores que monitorean actividades. Por ejemplo TCLs vinculadas o embebidas en dispositivos como celulares, sistemas de navegación de vehículos, y clientes de internet inalámbricos proveen contenido específico de localización a cambio de rastrear a los individuos en el espacio y en el tiempo (H. J. Miller & Han, 2009). También se pueden mencionar datos transaccionales, tales como facturación o movimientos entre cuentas bancarias, información

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

generada directamente por las máquinas como por ejemplo sensores de localización vinculados a vehículos de transporte y la información de redes sociales tales como Twitter y Facebook.

2.4 Big data

El big data puede definirse como “bases de datos estructuradas y no estructuradas con volúmenes masivos de datos que no pueden ser fácilmente capturadas, almacenadas, manipuladas, analizadas, gestionadas y presentadas por tecnologías tradicionales de software, hardware y bases de datos” (Li et al., 2016). El big data también puede caracterizarse como la gestión de datos caracterizados por su variedad, velocidad y volumen. A estos rasgos se ha agregado el reto de la veracidad en el sentido de que debe garantizarse la calidad e integridad de los datos (Laney, 2001).

La gestión de datos geográficos está cada vez más inmersa en esta nueva realidad. Por ejemplo, los sensores remotos generan continuamente volúmenes considerables de información, lo que también sucede con dispositivos conscientes de la localización e información georreferenciada proveniente de redes sociales. Desde el punto de vista de la velocidad existen cada vez más sistemas que transmiten datos en streaming con localización explícita o implícita. Si se hace énfasis en la variedad pueden verse cada vez más datos geográficos no estructurados en los tradicionales formatos ráster y vector, por ejemplo fotos, videos e información de redes sociales con etiquetas (Li et al., 2016).

De acuerdo a Mayer-Schönberger & Cukier (2014) el big data conlleva tres grandes desafíos: 1. Poblaciones no muestras; 2. Datos desordenados, no limpios; 3. correlaciones, no causalidad. Por su parte Miller & Goodchild (2015) analizan estos retos desde el punto de vista de la geografía, concluyendo que al interior de este campo son evolucionarios más que revolucionarios porque la geografía ha tratado con problemas similares a lo largo de su historia.

Cuando se dice “poblaciones no muestras” se está hablando del paso de una situación dónde era costoso recolectar datos y en muchas ocasiones se dependía de muestras a una en la que se trabaja con poblaciones completas. El primer estado conducía a dos problemas: la falta de representatividad para segmentos pequeños de población y las dificultades de hacer uso secundario de estos datos. En la segunda situación las poblaciones tienden a estar autoseleccionadas. Esto invierte el problema del muestreo dónde se tiene una pregunta y se recolectan datos para responderla, a uno en el que se tiene un conjunto de datos y se determina qué preguntas pueden responder (Miller & Goodchild, 2015). Un ejemplo de este último problema son los sesgos en la recolección de información geográfica voluntaria ya mencionados.

Sobre los datos no estructurados Miller & Goodchild (2015) presentan dos alternativas: 1. usarlos únicamente para tareas en las que no se busque hacer generalizaciones ni requieran supuestos sobre la calidad; 2. Limpiarlos y verificarlos. Para este último caso presentan tres alternativas: la solución de la multitud; la solución social y la solución del conocimiento. La primera se basa en la idea de que con suficientes ojos todos los errores se hacen evidentes, con el problema de que no para todas las zonas geográficas puede haber suficientes ojos. La solución social consiste en establecer una jerarquía con moderadores, como en el caso de Openstreetmaps o Wikipedia, lo que mejora la calidad al costo de reducir la velocidad. Finalmente, la solución de la

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

calidad es una forma de detectar inconsistencias examinando la coherencia de la información con conocimiento apriorístico de la forma en que está organizado el mundo.

Por otra parte, la disponibilidad de datos ha producido un cambio de enfoque en la ciencia pasando de los intentos de establecer leyes universales a la búsqueda de descripciones más profundas de lo que está sucediendo en un lugar y momento determinado. De acuerdo a Miller & Goodchild (2015), estos virajes no son nuevos para la geografía que ha mantenido a lo largo de su historia una discusión entre los enfoques ideográficos y los nomotéticos. Para la geografía es importante encontrar explicaciones generales de los fenómenos, pero la heterogeneidad y la dependencia espacial también crean contextos locales que son importantes para entender el comportamiento espacial de diferentes tipos de sistemas (Miller & Goodchild, 2015; Starman, 2016).

Este nuevo enfoque epistemológico ha sido acompañado de un conjunto de técnicas específicas orientadas al “descubrimiento del conocimiento”. Cuando se utiliza este término se parte de la idea de que la información está escondida en bases de datos muy grandes en forma de patrones interesantes que no puede ser revelada mediante consultas tradicionales a bases de datos (Miller & Han, 2009).

Se busca entonces propiedades y relaciones no aleatorias que son válidas, originales, útiles, y comprensibles. Válido significa que el patrón es suficientemente general para aplicarlo a datos nuevos; original significa que el patrón es no trivial o esperado; útil implica que el patrón debe liderar alguna acción efectiva, decisión exitosa e investigación científica; comprensible en última instancia significa que el patrón debe ser simple e interpretable por los seres humanos (Miller & Han, 2009).

Dentro del proceso global de descubrimiento del conocimiento la minería de datos geográfica ocupa un lugar especial. En la geografía, como en otros dominios, los datos son altamente dimensionales, pero en esta se tiene la propiedad de cuatro dimensiones interrelacionadas –tres espaciales y una temporal-, que proveen un marco de medida para todas las demás. Además, la información geográfica puede recoger fenómenos de dependencia y heterogeneidad espacial (Miller & Han, 2009).

Un ejemplo de la utilidad de estas nuevas fuentes de información es el estudio de Louail et al (2014), en el cual se infiere la estructura espacial de 33 ciudades españolas utilizando registros de llamadas de telefonía celular. La base de datos con la que parte el estudio tiene conteos de usuarios únicos que llaman desde una determinada antena, la cual está georreferenciada. Los autores identifican un patrón similar de cambio en la distancia promedio entre las llamadas a lo largo del día, aunque con algunas ciudades que se comportan ligeramente diferente. También identifican puntos con alta densidad de llamadas, sus patrones de cambio y relaciones de distancia entre ellos. Logran así vislumbrar diferentes estructuras urbanas en función de los patrones de movilidad y conectividad de la población. El conjunto del estudio se realizó utilizando R.

Otro ejemplo del potencial del uso de redes sociales es el estudio de Compton et al (2014) en el que utilizan información pública proveniente de Twitter y Tumblr con miras a predecir disturbios en Latinoamérica. El sistema busca permanente combinaciones de palabras que puedan

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, n° 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

significar disturbios, lo que se refina filtrando cuentas con alta probabilidad de anunciar protestas. Estas últimas se identifican mediante una clasificación hecha a través de una regresión logística. La georreferenciación se realiza utilizando las coordenadas del usuario, cuando están disponibles, o de lo contrario con un algoritmo que la infiere de sus contactos. El sistema logró la georreferenciación de más de 91 millones de usuarios de Twitter y un éxito para un 45.9% de las predicciones.

3. Distribución

Buzai (2010) define la distribución como la forma en que “el conjunto de entidades de un mismo tipo se reparten de una determinada manera en el espacio”. La observación de la distribución sirve para evidenciar desigualdades espaciales susceptibles de intervención y aproximarse a la adecuación espacial entre oferta y demanda. En el presente trabajo se expondrán dos conjuntos de herramientas que sirven a este fin: las medidas centrográficas y el análisis exploratorio de datos espaciales.

3.1. Medidas centrográficas

Las medidas centrográficas fueron pensadas como una estadística descriptiva de datos geográficos que consiste en técnicas de resumen y visualización de distribuciones. Son especialmente adecuadas para hacerse una idea de la tendencia central, dispersión y dirección de las entidades consideradas (Kellerman, 1981; Mitchell, 2005).

3.1.1. Medidas de tendencia central.

Entre las medidas de tendencia central se encuentra el centro medio, el centro mediano y la entidad central. Entre otras cosas, sirven para identificar el cambio en una distribución, también para identificar una localización óptima para un objeto que requiere ser ubicado centralmente.

El centro mediano, se usa cuando se quiere encontrar la mejor ubicación de algo que debe estar centralmente localizado, siempre y cuando las distancias se midan como líneas rectas. Se define como el lugar con la mínima distancia a un conjunto de entidades geográficas. Para su cálculo primero se mide la distancia de un conjunto de localizaciones candidatas a los objetos geográficos evaluados, y luego se escoge la que representa la menor suma a todas las demás. Igual que con el centro medio, puede calcularse con ponderación.

Por otra parte, la “entidad central” es el objeto que tiene una menor distancia a todos los demás objetos. Se diferencia del centro mediano porque considera tipos de distancias diferentes a la euclidiana – por ejemplo una red vial-. Para determinar la entidad central se requiere: 1) obtener las distancias de todos los objetos geográficos a todos los demás; 2) sumar las distancias de cada objeto, individualmente considerado, a los demás; 3) escoger el objeto que tenga una suma menor. Se usa cuando se quiere determinar una ubicación óptima, por ejemplo una estación de bomberos o de policía, siempre que se tengan restricciones como que las entidades se conectan por determinado medio de transporte.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

3.1.2. Medidas de dispersión

Las medidas de dispersión sirven para determinar qué tan alejados están los valores de una distribución de una medida de tendencia central como la media o la mediana. En el caso de una distribución espacial se mide que tan compacta es la agrupación de los objetos geográficos, así como la dirección en que se distribuye el fenómeno.

Una primera medida a considerar es la distancia estándar, que consiste en el cálculo de la distancia media de todos los puntos al centro medio de la distribución. Mide qué tan compacta es una distribución geográfica. En ocasiones se representa gráficamente utilizándola como radio de un círculo cuyo centro es el centro medio. La distancia puede ponderarse por variables tales como la población de un lugar o el número de eventos que ocurren en el mismo. La elipse de desviación estándar, por otra parte, considera al mismo tiempo lo compacta de la distribución y su dirección.

Estas herramientas han encontrado un gran potencial de aplicación en el estudio del crimen. Por ejemplo: LeBeau (1987) las utilizó para estudiar dinámicas de crimen aplicándolas a una base de datos de San Diego, California entre 1971 y 1975. El autor encontró que los patrones espaciales varían: por tipos de delincuentes y en el tiempo, y que la dinámica del cambio tenía relación con el tipo de delincuentes. Por su parte Harries y LeBeau (2007) mencionan el uso de tres de estas medidas: centro de mínima distancia, centro medio y centro mediano, entre las herramientas para realizar perfilamiento geográfico de criminales seriales. El software Crimestat, por su parte, implementa la mayoría de estas técnicas como parte de sus funciones (Smith & Burce, 2008). La utilidad de las medidas centrográficas no termina, sin embargo, en el estudio del crimen. Miller menciona aplicaciones en áreas tan diversas como estudio de epicentros de volcanes, movimientos de la vida silvestres y estudio de mercados (Miller, 2007).

3.2. Análisis exploratorio de datos espaciales

El análisis exploratorio de datos espaciales –AEDE- es un tipo de análisis derivado del análisis exploratorio de datos –AED-, que toma en cuenta la naturaleza espacial de los datos georreferenciados (Bivand, 2010). De acuerdo a Yrigoyen & Calderón (2009) el AEDE permite tener una primera aproximación a los datos, examinando la existencia de alisado y discontinuidad. Por alisado se entiende una tendencia general de cambio de una variable en el espacio de análisis. La discontinuidad es local en su naturaleza y se refiere a los cambios abruptos en el valor una variable, por ejemplo valores atípicos en una localización dada.

De acuerdo con Anselin (1999) el AEDE:

“Puede ser definido ampliamente como un conjunto de técnicas para describir y visualizar las distribuciones espaciales, identificar localizaciones atípicas, descubrir patrones de asociación espacial y sugerir diferentes regímenes espaciales y otras formas de inestabilidad espacial o no estacionariedad territorial”(p. 258).

3.2.1. Vinculación interactiva de gráficos y mapas

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

El AEDE es interactivo. Los programas que lo implementan combinan gráficos con mapas. Al seleccionar un conjunto de datos en un gráfico, se ven resaltados inmediatamente sobre el mapa y a la inversa (Laurent, Ruiz-Gazen, & Thomas-Agnan, 2009). Por ejemplo, al seleccionar en un box plot, un histograma o un diagrama de dispersión los valores atípicos, se resaltan en el mapa las localizaciones geográficas correspondientes.

Este enfoque no está limitado a unos tipos de gráficos en particular. En principio una o más variables que se representan en un mapa también pueden ser representadas en un gráfico, lo que ofrece una amplia gama de posibilidades de exploración interactiva. Es el caso del paquete GeoXP del lenguaje de programación R, que calcula el valor del Coeficiente de Gini, genera la curva de Lorenz asociada al mismo y presenta un mapa interactivo. En el artículo que presenta el paquete se presenta un ejemplo que compara el número de estudiantes y el índice de ruralidad en Francia. (Laurent et al., 2009).

En términos generales, el AEDE permite realizar análisis univariantes y multivariantes. Entre los primeros se encuentra la representación de localizaciones, la representación de variables, y la detección de atípicos. Entre los segundos se encuentra la representación cartográfica de diagramas de dispersión, gráficos de coordenadas paralelas y gráficos condicionales (Anselin, 2005; Yrigoyen & Calderón, 2009). Además del paquete GeoXP ya mencionado, estas funcionalidades se encuentran en el programa Geoda y algunas de ellas en ArcMap de ESRI.

3.2.2. Dependencia espacial

Este concepto, de gran importancia en el análisis espacial, puede definirse como la tendencia para desviaciones en valores del proceso desde su media para "seguir" los de otros sitios vecinos (Bailey & Gatrell, 1996). También como la coincidencia de los valores de una variable en determinados lugares (Yrigoyen & Calderón, 2009).

Medidas comúnmente usadas para medirlo son el Indicador Global de Moran y el Indicador Getis-Ord General G, los cuales se interpretan dentro del contexto de la hipótesis nula de aleatoriedad espacial completa. Cuando el indicador Global de Moran toma un valor positivo, indica la presencia de un patrón espacial tipo conglomerado, en donde valores altos de la variable de estudio tienden a estar rodeados por valores altos mientras que valores bajos tienden a estar rodeados de valores bajos. Un valor negativo en el indicador indica la presencia de un patrón espacial disperso o uniforme, que puede reflejar un tipo de proceso competitivo. En este caso, una entidad con un valor alto rechaza a otras entidades con valores altos; del mismo modo, una entidad con un valor bajo rechaza a otras entidades con valores bajos. Cuando el indicador toma un valor de 0 se presenta ausencia de autocorrelación espacial.

En el caso del estadístico Getis-Ord General G cuando el valor P de la prueba no es estadísticamente significativo indica que la variable de estudio está distribuida aleatoriamente en el espacio, es decir, hay ausencia de autocorrelación espacial. Entre tanto, cuando el valor P es estadísticamente significativo y la puntuación z es positiva, indica que los valores altos de la variable tienden a estar rodeados por valores altos. Si el valor P es estadísticamente significativo y la puntuación z es negativa se revela que la distribución espacial de los valores bajos de la variable es agrupada.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

Sin embargo pueden existir dos casos que requieren de la introducción de indicadores locales (Yrigoyen & Calderón, 2009):

- En primer lugar, es posible que el estadístico I de Moran no detecte la presencia de dependencia espacial en la distribución de una variable sobre un espacio dado, pero que existan agrupamientos de unidades en los que dicha variable experimenta una concentración.
- En segundo lugar, también es posible que el estadístico I de Moran detecte la presencia de dependencia espacial en una variable, pero que no todas las unidades del espacio considerado contribuyan con igual peso en el indicador global.

Para estos casos se utilizan indicadores que detectan conglomerados de valores altos o bajos –Moran Local y Getis Ord G_i^* –, así como atípicos espaciales, los cuales permiten realizar cartografía de conglomerados espaciales. También vale la pena mencionar el diagrama de dispersión de Moran, que permite observar de manera interactiva, en cuatro cuadrantes, las zonas donde la variable tiene valores altos rodeados de altos, bajos rodeados de bajos, altos rodeados de bajos y bajos rodeados de altos (Laurent et al., 2009).

Entre los estudios que utilizan esta aproximación puede citarse a Estrada & Moreno (2013), con un estudio sobre la pobreza multidimensional. El trabajo, realizado a nivel municipal para Colombia, calculó el índice de pobreza multidimensional a partir de datos del Censo 2005, los resultados mostraron autocorrelación espacial, con municipios pobres rodeados de municipios pobres y otros con poca pobreza rodeados de municipios similares. Las autoras logran identificar conglomerados geográficos claramente diferenciados y una marcada diferencia entre valores urbanos y rurales.

Por su parte, Rivero (2008) aboga por el uso del AEDE para estudiar datos turísticos. Además de una clara conceptualización, el documento presenta un estudio de caso para Extremadura en España. Como variables utiliza el número de plazas en hoteles, alojamientos rurales y restaurantes. Encuentra conglomerados para las tres variables. También identifica una correlación espacial negativa entre restaurantes de primera y segunda categoría que atribuye a la dinámica competitiva.

Para la ciudad de Bogotá, Jácome (2015) realiza una aproximación orientada a la política pública mediante el análisis de sedes educativas. El autor logra establecer la existencia de conglomerados espaciales en indicadores educativos tales como los resultados en pruebas estandarizadas, deserción y docentes con postgrado. El mismo autor efectuó un estudio de las elecciones locales en Bogotá en el que, con el apoyo de esta técnica, logró determinar una relación entre grupos poblacionales y preferencias electorales (Jácome, 2013).

Han surgido también estudios de epidemiología espacial o geografía médica. Con ellos se busca visualizar los patrones espaciales de la enfermedad y la mortalidad, profundizar en sus causas (mecanismos biológicos de transmisión, identificación de agentes infecciosos) y evaluar la probabilidad de potencial exposición (Waller & Gotway, 2004).

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, n° 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

En Colombia, la mortalidad infantil ha sido estudiada desde un enfoque espacial por Ruiz-Santacruz & Durán (2013), mostrando que a nivel municipal presenta autocorrelación espacial positiva. Este estudio identificó focos críticos (aglomerados espaciales con valores altos de mortalidad) en las zonas periféricas del territorio nacional, y conglomerados espaciales con valores bajos de mortalidad en la región centro-oriental del país.

4. Asociación espacial

La Asociación espacial estudia la forma en que se relacionan las distribuciones espaciales de las variables, ya sea desde un punto de vista descriptivo o analizando la forma en que su combinación produce determinados efectos. En términos generales Buzai describe este término como: “el estudio de las semejanzas encontradas al comparar distintas distribuciones espaciales” (Buzai, 2010). Al respecto se abordará tanto la geodemografía como la Regresión Geográficamente Ponderada.

4.1. Geodemografía

La geodemografía ha sido definida como “el análisis de datos socioeconómicos y conductuales sobre las personas, para investigar los patrones geográficos que estructuran y son estructurados por las formas y funciones de los asentamientos” (Harris, Sleight, & Webber, 2005, p. 14). Comúnmente se asocia con el estudio de las personas en relación al lugar en donde viven (Sleight, 2004), basado en el principio de que la estructura socio-espacial se encuentra altamente correlacionada con comportamientos, actitudes y preferencias (Singleton & Spielman, 2013).

En un estudio geodemográfico se generan clasificaciones por área, esto es: “agrupamiento de áreas dentro de grupos sobre la base de la similitud de las características dentro de ellas” (Vickers & Rees, 2007, p.380). La metodología para su generación puede resumirse en siete pasos: 1. Definir los elementos a reunir o agrupar; 2. Definir las variables a utilizar; 3. Estandarizar las variables; 4. Medir la asociación entre las variables; 5. Definir un método de agrupación; 6. Definir el número de agrupaciones; 7. Interpretar, probar y replicar (Lopes, 2011; Vickers & Rees, 2007). Un ejemplo de su aplicación se puede ver en el trabajo de Ojo, Vickers, & Ballas (2013).

Pensando en su aplicación al sector de planeación urbana Song & Knaap (2007) enuncian cuatro ventajas de este tipo de clasificaciones: 1. Permiten la comprensión: “el conocimiento humano se mejora cuando ideas complejas y atributos múltiples son organizados y estructurados en unos pocos constructos bien definidos”; 2. Facilitan un análisis cuantitativo riguroso, en la medida en que permite pasar de muchas variables correlacionadas a pocas no correlacionadas; 3. Apoyan el diseño de políticas públicas; 4. Son útiles para la evaluación de políticas públicas.

Por su parte, Singleton & Spielman (2013) afirman que el valor de la aproximación contextual, en tanto esté embebido en las modernas clasificaciones geodemográficas, es la capacidad de proveer un cuadro matizado de grandes áreas, para dibujar una analogía entre áreas geográficas dispares, y para habilitar la provisión de servicios focalizados (públicos y privados).

Siguiendo con lo anterior, Harris et al (2005) consideran que al crear unas clasificaciones de áreas pequeñas con cobertura nacional es posible evaluar poblaciones y sus características por tipo

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

de áreas. Posteriormente se pueden introducir encuestas en los segmentos identificados, para generar insumos sobre las características fundamentales de quienes las responden.

4.1.1. Aplicaciones de la geodemografía

Inicialmente, la geodemografía fue empleada en el área del comercio como herramienta para la localización de negocios y el marketing, luego, en la década de los noventa se incrementó su uso para el sector público. De acuerdo a un estudio realizado por Singleton & Spielman (2013), las aplicaciones académicas de geodemografía en el Reino Unido y Estados Unidos, han abordado temas de salud, educación, gestión de recursos ambientales, crimen, política urbana, transporte turismo, segregación, migración, ventas, estructura local urbana, vivienda y privaciones.

En temas de salud, la geodemografía ofrece la posibilidad de integrar, modelar y cartografiar necesidades en atención de salud y otros indicadores, lo que permite orientar programas y campañas de salud en cada vecindario (Petersen et al., 2011). La geodemografía se ha utilizado para estudiar las condiciones de morbilidad o mortalidad por vecindario, y la influencia de las características de cada zona en las condiciones de salud de la población.

Autores como Grigsby-Toussaint, Chi, & Fiese (2011) estudiaron si los residentes, en edad preescolar, de vecindarios con alta presencia de zonas verdes registran mayores tasas de actividad física al aire libre. El estudio utilizó el índice de vegetación de diferencia normalizada -NDVI, obtenido a partir de imágenes satelitales Landsat, para medir la presencia de zonas verdes en los vecindarios. Se encontró que un aumento de una unidad en el NDVI aumentó el tiempo de juego al aire libre en un niño en aproximadamente 3 minutos, por tanto se concluyó que la presencia de zonas verdes en los vecindarios influye en el comportamiento de la actividad física.

Dewulf, Neutens, Van Dyck, De Bourdeaudhuij, & Van de Weghe (2012) estudiaron la influencia de los vecindarios y la socio-demografía en los tiempos de caminata realizados y percibidos de la población. El estudio encuestó a 1164 personas de 24 vecindarios diferentes en la ciudad de Ghent y comparó el tiempo de caminata reportado con el medido a través de herramientas de Sistemas de Información Geográfica. La peatonalidad de los vecindarios fue evaluada teniendo en cuenta la densidad residencial, la conectividad y el uso del suelo. Como resultado se encontró que las personas que residen en un vecindario con baja peatonalidad subestimaron los tiempos de caminata con más frecuencia que los residentes de barrios con alta peatonalidad.

Por otro lado, Webber (2004) estudió las desigualdades por vecindario en los patrones de ingreso al hospital. El autor utilizó registros del Departamento de Salud de Inglaterra, que contenía información de edad, sexo, diagnóstico médico y código postal del paciente. La segmentación de la población se realizó a partir de los 61 perfiles geodemográficos elaborados por la empresa Experian, con los datos del Censo 2001. Como resultado, se encontró altas tasas de ingreso al hospital en vecindarios con mayor número de personas de la tercera edad o de familias de bajos ingresos. Analizando las variaciones por vecindarios según el diagnóstico, se identificó que ciertas enfermedades están concentradas en vecindarios específicos, el caso más común se presenta en enfermedades pulmonares, seguido de esquizofrenia, lesiones y envenenamiento. En contraste, diagnósticos asociados a cáncer, están distribuidos de manera dispersa en todos los vecindarios. El

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

estudio menciona que esta información geográfica a escala local sirve para la asignación de recursos y la implementación de campañas de salud específicas.

Los estudios de segmentación geodemográfica también han permitido entender los patrones de crimen en relación con las características de los vecindarios. Como lo expresa Brown, Hirschfield, & Batey (2000) el análisis geodemográfico puede contribuir a un mejor entendimiento del grado en el que un tipo particular de crimen tiende a concentrarse en ciertos tipos de vecindarios o sectores residenciales.

El estudio de Webber & Ashby (2006), por su parte, tuvo como objeto identificar el alcance del análisis geodemográfico para su uso en seguridad urbana. Para ello, vincularon la información georreferenciada de crimen con los perfiles geodemográficos definidos en el Censo de Reino Unido del 2001, identificando los tipos de vecindarios que tienen mayor proporción de víctimas y mayores tasas de incidencia de un crimen en particular. Como resultado se generó un reporte para cada vecindario, indicando el tipo más común de crimen, el nivel de incidencia y la estrategia de seguridad propuesta de acuerdo a las características sociodemográficas del entorno.

Por otro lado, la geodemografía brinda un contexto estructural para el análisis en la provisión de servicios públicos, por ejemplo vecindarios con una población predominantemente joven demandará servicios educativos, mientras que vecindarios con mayor proporción de población en la tercera edad demandará servicios de salud. Un conocimiento robusto de la geodemografía local soportará la toma de decisiones para la localización de equipamientos y servicios públicos.

El trabajo de Hindle, Spollen, & Dixon (2004) aplicó técnicas de geodemografía para apoyar técnicamente la localización de servicios públicos, teniendo en cuenta la relación costo-beneficio y el impacto de las áreas de servicio en términos de sus características. Para lograrlo, se seleccionaron 26 áreas geográficas que diferían ampliamente en términos demográficos y se estudió un conjunto de servicios para estas áreas (escuelas, enfermería, respuesta policial a incidentes urgentes, respuesta a llamadas de emergencia). Se calcularon indicadores en términos de tasas de demanda, accesibilidad y tiempos de desplazamiento, con el fin de poder predecir la relación costo beneficio de cada servicio en el área geográfica. De esta manera se obtuvo para todas las áreas geográficas el servicio con mayor beneficio para la población.

La aplicación de geodemografía para localización de servicios ha sido ampliamente desarrollada por el sector privado en el área del geomarketing, siendo sus aplicaciones más frecuentes: adaptar el surtido de una tienda; trabajar en micromercados; reestructurar una red; determinar una nueva localización; ajustar los objetivos en relación al potencial; identificar los mejores consumidores; optimizar las inversiones en marketing directo; soportar las ventas; y analizar la competencia (Douard, 2006).

Vale la pena considerar también los trabajos de Gustavo Buzai quien ha utilizado análisis multivariado para variables sociales en combinación con técnicas SIG con el fin de mejorar la comprensión del espacio urbano en algunas ciudades argentinas. Buzai no sólo utiliza técnicas de agrupamiento de unidades de análisis sino también otras de agrupamiento de variables, como el

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

análisis de componentes principales y el análisis factorial, para producir variables sintéticas que son representadas cartográficamente (Buzai, 2003; Buzai, Baxendale, Cruz, & González, 2003).

La utilidad de esta aproximación se puede observar en el uso de la técnica de componentes principales para analizar la segregación espacial. Es el caso del estudio de Díaz & Acosta (2011) en Puerto Vallarta, México quienes la utilizan para reducir la dimensionalidad de un conjunto de variables que miden características del hogar y la vivienda, obtenidas a partir del Censo de Población y Vivienda. En su análisis los autores caracterizan cuatro componentes principales y pasan a generar con ellos mapas en los que se observan los extremos y la dirección de la segregación residencial en la ciudad.

4.2. Regresión Ponderada Geográficamente para detectar heterogeneidad especial

La heterogeneidad espacial es definida por Yrigoyen (2004) como: "la ausencia de estabilidad en el espacio del comportamiento humano o de otras relaciones en estudio" (p. 1). Para los modelos espaciales esto implica que las formas funcionales y los parámetros variarán con la localización geográfica no siendo homogéneos para toda la matriz de datos.

Por otra parte, la regresión lineal debe interpretarse geográficamente como parámetros promedio para el conjunto del área de estudio o una caracterización global de la relación prevaleciente, lo que es válido si se presenta estacionariedad espacial, es decir si las relaciones especificadas en el área de estudio se mantienen constantes (Longley & Tobón, 2004).

Sin embargo, la regresión así planteada no da cuenta de variaciones locales en los datos y no provee indicación de si cada variable es especificada en la escala apropiada (Longley & Tobón, 2004). Por otra parte, en la mayoría de los análisis geográficos, el supuesto de errores con distribución normal no correlacionados y con varianza constante de los modelos lineales no es realista. La razón es la existencia de dependencia espacial en los datos que puede llevar a una dependencia espacial en los residuales, así como la existencia de heterogeneidad espacial, es decir relaciones entre variables que cambian en función del lugar (Serrano & Valcarce, 2000).

La regresión Ponderada Geográficamente –GWR - es uno de los métodos con los que se puede tratar el problema de la heterogeneidad espacial; (Yrigoyen, 2004; Brunson, Fotheringham, & Charlton, 2008). El objetivo del GWR es generar coeficientes locales, más que globales, generando así un modelo para cada entidad. La ecuación puede definirse como sigue:

$$y_i = a_{i0} + \sum_{k=1,m} a_{ik}x_{ik} + \epsilon_i$$

Dónde y_i es la variable a ser explicada o variable dependiente; x_{ik} es la observación i de la variable independiente k , a_k representa un coeficiente que mide la influencia de las variables independientes sobre la dependiente, siendo a_0 el intercepto o valor esperado para la variable dependiente cuando los coeficientes de todas las variables explicativas es cero. ϵ_i es el término de error resultante de comparar los valores predichos por el modelo con los observados en y_i .

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, n° 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

El término que diferencia esta formulación de la de un modelo lineal simple es a_{ik} que representa el valor del parámetro k en la localización i. Para calcularlo se realiza una calibración en la que se penalizan más los errores que están más cerca de cada entidad geográfica. El resultado es que las entidades más cercanas pesan más en el modelo, aunque su peso específico depende de la función de decaimiento escogida (Fotheringham, Brunson, & Charlton, 2002).

Gutiérrez, Cardozo, & García (2013) utilizan el GWR para estimar la demanda de transporte a nivel de estación en el metro de Madrid. Le asignan a esta metodología la función de mejorar los resultados de modelos lineales simples que suponen estabilidad en los parámetros. Los modelos lineales se han venido usando como alternativa a modelos de escala regional, de eficacia probada pero sofisticados, costosos y orientados a escalas menos detalladas.

Estos autores escogieron un modelo OLS con cuatro variables independientes: cantidad de empleo, líneas de autobús interurbano, población ocupada y líneas de metro, al que se encontró autocorrelación en los residuales. Mediante la realización de un modelo GWR se aumentó el R^2 ajustado (porcentaje de la varianza explicada por las variables independientes) de 0,56 a 0,70. Como resultado se generaron mapas con los coeficientes locales para las variables independientes, mostrando que su influencia en la demanda del metro varía en función del lugar, lo que es de gran utilidad para la planificación del transporte público.

Por otra parte, el Centro de Excelencia para Información Estadística de Gobierno, Seguridad Pública, Victimización y Justicia, de México (2015), realizó un trabajo orientado a comparar los patrones de delitos violentos en varias ciudades de América Latina y generar recomendaciones para combatirlos. Para la escritura del presente artículo se tuvo acceso a los resultados de la investigación para la ciudad de Zapopan en México.

Los autores del estudio realizan una afirmación que es una justificación de la necesidad de incluir la heterogeneidad espacial en modelos sobre crimen: "En la actualidad, existe prácticamente consenso entre investigadores, diseñadores de políticas públicas y demás interesados en el estudio del delito en el reconocimiento de una asociación significativa entre el crimen y el espacio geográfico, es decir, el delito no es un fenómeno que se distribuye uniformemente en el espacio, sino que tiende a ocurrir en ciertos lugares con características particulares" (2015, p.1).

El estudio buscó determinar la relación entre el delito y sus determinantes económicos y demográficos en el área de estudio. Bajo el esquema de la teoría de la desorganización social se definieron unas variables que fueron operacionalizadas con información disponible a nivel de áreas geoestadísticas básicas.

Se realizó un modelo de regresión Poisson inflado en cero para tres tipos de delito: robo, lesiones y homicidio, que permitió identificar un patrón espacial característico para cada uno de ellos. Por ejemplo: los homicidios tienden a ocurrir en zonas con altos porcentajes de población joven y de población divorciada; mientras que los robos sucedieron principalmente en zonas con alta proporción de población joven, alta proporción de población divorciada y población con alta escolaridad.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

4. Interacción espacial

Para Buzai (2010) el concepto de interacción espacial “considera la estructuración de un espacio relacional en el cual las localizaciones (sitios) distancias (ideales o reales) y vínculos (flujos) resultan fundamentales en la definición de espacios funcionales” (p. 9). Por su parte, Yrigoyen (1996) lo define, de modo general, como todo movimiento o comunicación sobre el espacio resultante de un proceso de decisión, esta interacción implica un origen, un destino y el movimiento resultante del desplazamiento de acuerdo a la elección efectuada. Los movimientos pueden darse por migración, trabajo, acceso a servicios, compras, entre otros.

Los modelos de interacción espacial buscan identificar los flujos entre proveedores/receptores (oferta) y usuarios/consumidores (demanda) brindando una aproximación al entendimiento de la dinámica espacial que hace que las personas visiten un equipamiento o establecimiento en particular. Desde esta perspectiva también son de interés los denominados modelos de localización – asignación, que buscan optimizar la relación espacial entre oferta y demanda.

5.1. Modelos de interacción espacial

Los modelos de interacción espacial se caracterizan por tomar en cuenta, simultáneamente tres componentes: la capacidad de atracción de una entidad geográfica –que puede ser por ejemplo una ciudad o una tienda-; la capacidad de repulsión de otra, por ejemplo el poder adquisitivo, desempleo; y la interacción entre las dos (Bailey & Gatrell, 1996). En estos modelos las unidades de análisis no son ni el establecimiento, ni las áreas de donde proviene la población, sino las relaciones entre ellas. Entre los modelos más conocidos se encuentran el Modelo Gravitacional Determinístico de Reilly, conocido también, como ley de la gravitación del comercio minorista, el Gravitacional Probabilístico de Huff y el Modelo Multiplicativo de Interacción Espacial de Natanishi y Cooper (Cliquet, 2006; Huff, 1963; Hernando, 1990). Es importante anotar que el paquete MCI de R implementa tanto el modelo de Huff como el MCI (Wieland, 2017).

Sus aplicaciones abarcan estudios sobre flujos de transporte, mercancía, acceso a servicios, migración, viajes del lugar de residencia al sitio de trabajo, planificación territorial, entre otros. Aunque su desarrollo tiene una historia de varias décadas, estas metodologías tienen una oportunidad de expandir sus alcances con los datos de interacción provenientes del Big Data. Es el caso de los datos de movilidad que pueden ser derivados de la telefonía móvil o simplemente por la mejora en los registros administrativos que permiten realizar una trazabilidad de la movilidad de personas durante distintas etapas de la vida.

El modelo gravitacional, nace de una analogía con la ley de la gravitación universal de Isaac Newton. Formulado originalmente por Reilly en 1931, ofrece una explicación de la atracción comercial entre dos ciudades considerando una variable “masa” o de atracción y otra variable “fricción” o de frenado asociada a la distancia. De acuerdo a Cliquet (2006) este modelo supone que “dos ciudades atraen comercio de una ciudad intermedia o pueblo en la vecindad de un punto de corte, aproximadamente en proporción directa a las poblaciones de las dos ciudades y en proporción inversa al cuadrado de las distancias al pueblo intermedio”. El modelo gravitacional de Reilly considera que la aglomeración tiende a incrementar el atractivo de un establecimiento, es decir,

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

aquellos localizados en centros con alta población atraen más clientes de distancias lejanas que equipamientos en centros pequeños (Yrigoyen, 1997). Un ejemplo práctico de aplicación de este modelo fue realizado por Martín Hernando (Hernando, 1990) quien lo utilizó para delimitar áreas comerciales en la región de Murcia, en España.

Si bien la propuesta de Reilly ha brindado resultados aceptables en el estudio de las dinámicas de flujos de consumidores, se le ha criticado que sea determinística y que no permite estimar la población en las áreas urbanas. Por ejemplo, no fundamenta la propuesta de que los costos de interacción entre dos puntos sean exactamente iguales a la distancia que los separa al cuadrado.

Una solución a estos inconvenientes proviene del modelo gravitacional probabilístico de Huff, (1963). Huff argumenta que cuando los consumidores tienen diferentes alternativas para adquirir un servicio, pueden visitar diferentes establecimientos en lugar de restringirse a uno, de tal manera que cada establecimiento dentro del área geográfica tiene una probabilidad de ser frecuentado. Dicha probabilidad, que está relacionada positivamente con el tamaño y negativamente con la distancia, puede calcularse de varias formas de acuerdo a las necesidades de la investigación y la disponibilidad de datos. Sin embargo este modelo adolece de algunas fallas: el poder explicativo es restringido por el número insuficiente de variables; hay varias propuestas para estimar el coeficiente de determinación pero ninguna es realmente satisfactoria (Cliquet, 2006).

Una solución surgió de la mano de Natanishi y Cooper quienes formularon el Modelo Multiplicativo de Interacción Competitiva –MCI- por sus siglas en Inglés- De acuerdo a Cliquet (2006), una ventaja de este modelo es que puede ser abordado como una regresión con un procedimiento de resolución por mínimos cuadrados ordinarios –OLS-, esto implica que sólo soporta variables cuantitativas.

Por otra parte, existen dos formulaciones de este modelo en relación con las variables incluidas, el objetivo y el subjetivo. Los modelos objetivos se caracterizan por el tipo de variables utilizadas: área de parqueaderos, tamaño de la tienda, accesibilidad, etc. Los subjetivos indagan por la percepción de los clientes sobre determinados aspectos, que son conceptualizados como factores de atracción (Cliquet, 2006).

Sobre la selección de variables más relevantes para incluir en el modelo, una posible aproximación consiste en realizar regresiones que indiquen la sensibilidad de variables candidatas para diferentes segmentos de la población (González-Benito, Greatorex, & Muñoz-Gallego, 2000). Cliquet por su parte utiliza un enfoque similar, con variables subjetivas derivadas de una encuesta, calibrando los efectos de atracción de diferentes tipos de promociones en tiendas de muebles (Cliquet, 1995).

5.2. Modelos de localización y asignación

Los modelos de localización asignación tienen un doble objetivo: 1. Definir de manera óptima la ubicación de servicios o establecimientos para que se atienda de manera adecuada la población demandante y 2. Asignar la demanda a los puntos de servicio, obteniendo áreas de servicio o de mercado.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

En el marketing el primer objetivo se transforma en optimizar el número y localización de las tiendas de tal forma que maximicen el valor de los consumidores atendidos (Cliquet, 2006). La aplicación de estos modelos puede ampliarse con el fin de evaluar la formulación de políticas y planes, ampliar una red de equipamientos o cerrar óptimamente algunos de ellos (Moreno & Bosque, 2010).

En los modelos de localización asignación se consideran los siguientes criterios: justicia espacial, eficiencia económica espacial y tipo de externalidad que genera el equipamiento.

- Las externalidades pueden ser positivas, para equipamientos deseables de instalar cerca a la población (escuelas, centros culturales, cines, etc); o pueden ser negativas cuando el equipamiento genera un efecto molesto en la población (vertederos de residuos sólidos, cárceles, entre otros).
- La eficiencia espacial se refiere al coste, en tiempos de recorrido o distancias, que la población se verá obligada a recorrer para poder utilizar los servicios (Moreno & Bosque, 2010).
- La justicia espacial se refiere al grado de igualdad en la distribución de los equipamientos en relación a la población demandante, de esta manera se busca optimizar la posición de las instalaciones de modo que se repartan equitativamente las externalidades que generan (Moreno & Bosque, 2010).

Los componentes de los modelos de localización y asignación son como siguen: una función objetiva a ser optimizada acorde a los diferentes sitios posibles; las áreas de demanda reducidas a un punto central o centroide donde la demanda por bienes y servicios está concentrada; posibles sitios con localización, accesibilidad e infraestructura; la matriz de distancia o tiempo de acceso; la regla de asignación, o la forma en que los consumidores escogen entre las ofertas propuestas (Cliquet, 2006).

6. Evolución espacial

De acuerdo a Buzai, (2010), la evolución espacial investiga el camino de transición de un estado a otro, considerando la incorporación de la dimensión temporal. En otras palabras, centrarse en el donde corresponde a una visión basada en la localización y el cuándo en una visión basada en el tiempo. El cómo, desde un punto de vista espacial, presentaría el camino evolutivo espacio-temporal.

Para que un evento sea analizado en un contexto espacio-temporal se requiere que los datos sean recolectados a lo largo del tiempo y del espacio. Los ejemplos son abundantes y están relacionados con el uso de imágenes satelitales para la detección del cambio en la cobertura del suelo, análisis de variables meteorológicas medidas en diferentes estaciones, trayectorias de desplazamiento de personas o animales, patrones de distribución de enfermedades, entre otras.

Sin embargo, a pesar de la abundancia de estos datos, de acuerdo a Schabenberger & Gotway (2004) frecuentemente cuando se analizan datos espacio-temporales se suele estudiar primero el aspecto espacial y posteriormente el aspecto temporal o viceversa, pero no se utiliza un enfoque integrado donde el espacio y el tiempo no están separados.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

El integrar el espacio y el tiempo, permite acudir al pasado para ver las génesis de las configuraciones espaciales actuales y también permite obtener resultados de simulación sobre posibles situaciones futuras (Buzai, 2010). De esta manera, una modelación espacio-temporal podría tener como propósitos la inferencia de parámetros que expliquen la etiología de un proceso espacio-temporal o la obtención de configuraciones futuras para la predicción de eventos en el espacio o en el tiempo.

6.1. Métodos del análisis espacio temporal

Los métodos comúnmente usados en el análisis de datos espacio temporales corresponden a técnicas de: visualización, detección de atípicos, predicción, particionamiento y agregación, identificación de patrones de agrupamiento y detección de cambio.

La visualización exploratoria de datos espacio temporales permite examinar la evolución de los datos a través del tiempo, mediante el uso de animaciones, diagramas espacio-tiempo, serie de mapas en una secuencia de tiempo que pueden representar cambios en áreas definidas o eventos que se expanden en el territorio.

Adicionalmente, las técnicas de estadística espacial para la detección de atípicos son aplicables en el análisis de la evolución espacial. Entendiendo el atípico espacio temporal como aquel elemento cuyo atributo temático difiere significativamente de otros objetos en su vecindario espacio temporal (Shekhar et al., 2015).

Para la predicción de variables, se utilizan modelos espacio temporales los cuales tienen como objetivo predecir el valor de una variable dependiente en función de un conjunto de variables explicativas. Cuando la variable dependiente es discreta, el proceso es denominado clasificación espacio temporal, cuando es continua se denomina regresión espacio temporal.

Para realizar una clasificación espacio temporal se pueden utilizar métodos de interpretación de imágenes de sensores remotos, donde las variables explicativas corresponden al valor espectral de cada una de las bandas que componen la imagen y la variable dependiente corresponde a una cobertura del suelo (urbano, bosque, agua, etc.). Dado el avance tecnológico de los sensores remotos en las últimas décadas, es factible recolectar imágenes sobre una zona específica con una frecuencia temporal de días, meses o años; lo que permite realizar estudios de cambios en la cobertura del suelo, crecimiento de las áreas urbanas, deforestación, disminución de fuentes hídricas, entre otros.

Las técnicas de evaluación multicriterio también son utilizadas para la obtención de configuraciones futuras de variables discretas a partir de la modificación de valores de las variables como simulación de cambios a ser realizados (Buzai, 2010). Varios trabajos utilizaron modelos espaciales multicriterio para identificar la vulnerabilidad y susceptibilidad por amenazas naturales, es el caso del estudio de Chau et al., (2004) que realizó un análisis espacial de deslizamientos para la Isla de Hong Kong y del trabajo de Ruiz-Pérez & Gelabert Grimalt (2012) quienes realizaron un análisis de la vulnerabilidad social frente a desastres naturales en la isla de Mallorca.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

En el contexto de la simulación espacial dinámica, destacan los modelos de autómatas celulares, en los cuales un sistema de celdas interactúan de manera simple para revelar un comportamiento global complejo (Wolfram, 1983). En los sistemas de autómatas celulares, el estado de una celda en un arreglo depende del estado previo de las celdas vecinas y de una conjunto de reglas preestablecidas (Câmara & Monteiro, 2001).

Un modelo de autómatas celulares consta de: a) un espacio euclidiano dividido en un arreglo de celdas idénticas; b) una estructura de vecindario aplicada a cada celda; c) variables estado, es decir, un conjunto de atributos para un autómata, que lo describe en un punto en particular en el tiempo; d) reglas de transición, las cuales determinan como el estado de un autómata cambia en el tiempo; e) tiempo considerado como pasos discretos: las reglas de transición aplican en cada paso o incremento de un intervalo de tiempo discreto (Câmara & Monteiro, 2001; Goodchild, 2015).

Los modelos de autómatas celulares son aplicados en la simulación de escenarios para la planificación territorial, cambios en la cobertura del suelo, crecimiento de la población, modelado del tráfico, entre otros.

Finalmente, la integración y partición espacio temporal es un método de análisis espacio temporal con amplias aplicaciones, que tiene como objetivo agrupar elementos espacio temporales similares o particionarlos en el espacio y tiempo. El uso de este método en datos de crimen permite identificar tendencias, focalizar zonas y horarios más críticos, facilitando la gestión de los recursos policiales. Si se analizan datos de accidentes de tráfico es factible agregarlos en las principales rutas para identificar las áreas con mayor accidentalidad. Los algoritmos para la integración y partición espacio temporal corresponden a técnicas de estimación de densidad de puntos y correlación temporal para series espacio temporales (Shekhar et al., 2015).

7. Conclusiones

En los últimos años se ha asistido a una proliferación de geodatos, causada principalmente por desarrollo de las actuales tecnologías de información, la explotación de fuentes de Big Data y el crecimiento de la cartografía colaborativa. Con el conocimiento adecuado para procesar dicha información, se puede generar un amplio valor agregado para usuarios institucionales tanto del sector público como el privado, así como para ciudadanos.

El análisis espacial es el eje articulador de las intervenciones de la geografía aplicada en la prestación de servicios a la población. Así lo demuestra el hecho de que sus cinco conceptos, tal como fueron enunciados por Buzai, estén respaldados por metodologías que permiten llevarlos a la práctica. Estos conceptos son: localización, distribución; asociación, interacción, y evolución espacial.

En términos de localización se resaltó la proliferación de datos geográficos y sus implicaciones para el empoderamiento de la población, así como para una mejor comunicación de información entre organizaciones y el público general. Por su parte el concepto de distribución consideró especialmente las llamadas técnicas centrográficas y el Análisis Exploratorio de Datos Espaciales –AEDE-, las cuales permiten obtener una primera aproximación de las tendencias generales y los valores atípicos en los datos espaciales.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

En el marco del concepto de asociación espacial se expusieron tres técnicas: 1. Geodemografía; 2. Regresión espacialmente ponderada; 3. Estudios de localización y asignación. La ventaja que tienen es que permiten identificar la estructura de un espacio dado en función de las relaciones entre sus variables.

La interacción espacial, aunque es exigente en datos, tiene la ventaja de que permite entender los espacios funcionales a partir de la comprensión de los espacios relacionales. Esto es claro, por ejemplo, en el tema del comercio y en la provisión óptima de servicios a la población. En este documento se trabajaron técnicas de interacción espacial y modelos de localización y asignación.

Finalmente, se abordó brevemente el concepto de evolución espacial, que investiga el camino de transición espacial de un estado a otro, considerando la incorporación de la dimensión temporal. Entre otras se mencionaron técnicas de visualización espaciotemporal, sensores remotos, clasificación multicriterio y de autómatas celulares. Estas técnicas tienen el potencial de permitir la comprensión de los cambios en los espacios geográficos y la formulación de escenarios futuros, lo cual es muy útil para la toma de decisiones.

Referencias bibliográficas

- Anselin, L. (1999). Interactive techniques and exploratory spatial data analysis. *Geographical Information Systems: Principles, Techniques, Management and Applications*, eds., P. Longley, M. Goodchild, D. Maguire, and D. Rhind. Cambridge: Geoinformation Int. Recuperado a partir de http://www.geos.ed.ac.uk/~gisteac/gis_book_abridged/files/ch17.pdf
- Anselin, L. (2005). *Exploring Spatial Data with GeoDa: A Workbook*. Center for Spatially Integrated Social Science. Recuperado a partir de <http://geodacenter.asu.edu/learning/tutorials>
- Bailey, T., & Gatrell, T. (1996). *Interactive Spatial Data Analysis* (1a ed.). Prentice Hall.
- Barron, C., Neis, P., & Zipf, A. (2014). A Comprehensive Framework for Intrinsic OpenStreetMap Quality Analysis. *Transactions in GIS*, 18(6), 877–895. <https://doi.org/10.1111/tgis.12073>
- Brown, P. J., Hirschfield, A. F., & Batey, P. W. (2000). Adding value to census data: public sector applications of the super profiles geodemographic typology. *Journal of Cities and Regions*, 10(19), 7.
- Brunsdon, C., Fotheringham, A. S., & Charlton, M. (2008). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. *Encyclopedia of Geographic Information Science*, 558.
- Buzai, G. D. (2003). *Mapas sociales urbanos*. Lugar Buenos Aires. Recuperado a partir de <http://www.biblioteca.unlpam.edu.ar/pubpdf/huellas/n08a08medus.pdf>
- Buzai, G. D. (2010). Análisis Espacial con Sistemas de Información Geográfica, sus cinco conceptos fundamentales. En *Geografía y Sistemas de Información Geográfica. Aspectos conceptuales y aplicaciones*. (Universidad Nacional de Lujan-GESIG, pp. 163–195). Lujan. Recuperado a partir de <http://www.gesig-proeg.com.ar/documentos/libros/libro-13/CAPITULO-07.pdf>

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

Buzai, G. D., Baxendale, C. A., Cruz, M. R., & González, J. (2003). Análisis Linkage de los patrones de localización socio-habitacional urbana: el caso de Luján. *Anuario de la División Geografía 2002-2003*, 151–178.

Câmara, G., & Monteiro, A. M. V. (2001). Geocomputation techniques for spatial analysis: are they relevant to health data? *Cadernos de Saúde Pública*, 17(5). <https://doi.org/10.1590/S0102-311X2001000500002>

Centro de Excelencia para Información Estadística de Gobierno, Seguridad Pública, Victimización y Justicia. (2015). Una aproximación geográfica al estudio de los delitos violentos en América Latina, el caso de Zapopan México.

Chau, K. T., Sze, Y. L., Fung, M. K., Wong, W. Y., Fong, E. L., & Chan, L. C. P. (2004). Landslide hazard analysis for Hong Kong using landslide inventory and GIS. *Computers & Geosciences*, 30(4), 429–443. <https://doi.org/10.1016/j.cageo.2003.08.013>

Cliquet, G. (1995). Implementing a subjective MCI model: An application to the furniture market. *European Journal of Operational Research*, 84(2), 279–291. [https://doi.org/10.1016/0377-2217\(93\)E0313-M](https://doi.org/10.1016/0377-2217(93)E0313-M)

Cliquet, G. (2006). Retail location models. En *Geomarketing: Methods and Strategies in Spatial Marketing (Geographical Information Systems Series)* (1a ed., pp. 137–156). Wiley-ISTE.

Consejo Nacional de Política Económica y Social. (2009, mayo). Conpes 3585. Recuperado a partir de http://www.ipgh.org/Difusion/2009/Files/CONPES-3585_PNIG-e-ICDE-02.PDF

Dewulf, B., Neutens, T., Van Dyck, D., De Bourdeaudhuij, I., & Van de Weghe, N. (2012). Correspondence between objective and perceived walking times to urban destinations: Influence of physical activity, neighbourhood walkability, and socio-demographics. *International journal of health geographics*, 11(1), 43.

Díaz, V., & Acosta, J. (2011). Segregación residencial y división social del espacio, elementos para el análisis de la estructura urbana de Puerto Vallarta, México. *REVISTA NODO*, 6(11). Recuperado a partir de <http://revistas.uan.edu.co/index.php/nodo/article/view/152>

Dodge, M., McDerby, M., & Turner, M. (2008). The Power of Geographical Visualizations. En *Geographic Visualization* (pp. 1–9). Recuperado a partir de http://samples.sainsburysebooks.co.uk/9780470987636_sample_411898.pdf

Douard, J. P. (2006). Geomarketing and consumer Behaviour. En *Geomarketing: Methods and Strategies in Spatial Marketing (Geographical Information Systems Series)* (1a ed., pp. 91–114). Wiley-ISTE.

Estrada, L., & Moreno, S. L. (2013). Análisis espacial de la pobreza multidimensional en Colombia a partir del censo de población del 2005. CANDANE. Recuperado a partir de <http://www.dane.gov.co/candane/files/Analisis%20Espacial%20de%20Pobreza.pdf>

Fotheringham, A. S., Brunson, C., & Charlton, M. (2002). *Geographically weighted regression: the analysis of spatially varying relationships*. Chichester: Wiley.

González-Benito, Ó., Gatrex, M., & Muñoz-Gallego, P. A. (2000). Assessment of potential retail segmentation variables An approach based on a subjective MCI resource allocation model. *Journal*

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

of Retailing and Consumer Services, 7(3), 171–179. [https://doi.org/10.1016/S0969-6989\(99\)00026-0](https://doi.org/10.1016/S0969-6989(99)00026-0)

Goodchild, L. (2015). *Geospatial Analysis 5th Edition*, 2015. Recuperado el 26 de abril de 2017, a partir de <http://www.spatialanalysisonline.com/HTML/>

GoogleCode. (2009). *Google Maps JavaScript API V2 Services*. Recuperado a partir de <http://code.google.com/intl/es-ES/apis/maps/documentation/javascript/v2/services.html>

Grigsby-Toussaint, D. S., Chi, S.-H., & Fiese, B. H. (2011). Where they live, how they play: Neighborhood greenness and outdoor physical activity among preschoolers. *International Journal of Health Geographics*, 10(1), 66. <https://doi.org/10.1186/1476-072X-10-66>

Gutiérrez, J., Cardozo, O. D., & García, J. C. (2013). Estimación directa de la demanda de transporte a nivel de estación mediante el uso de la regresión geográficamente ponderada. En *Congreso Chileno de Ingeniería de Transporte*. Recuperado a partir de <http://www.revistas.uchile.cl/index.php/CIT/article/viewArticle/28422>

Haklay, M., & Weber, P. (2008). Openstreetmap: User-generated street maps. *IEEE Pervasive Computing*, 7(4), 12–18.

Harries, K., & LeBeau, J. (2007). Issues in the Geographic Profiling of Crime: Review and Commentary. *Police Practice and Research*, 8(4), 321–333. <https://doi.org/10.1080/15614260701615029>

Harris, R. J., Sleight, P., & Webber, R. J. (2005). *Geodemographics, GIS and Neighbourhood Targeting*. London: Wiley. Recuperado a partir de http://metalib.ucl.ac.uk:9003/sfx_local?sid=google&auinit=R&aulast=Harris&title=Geodemographics%20and%20Neighbourhood%20Targeting&genre=book&isbn=0470864133&date=2005

Hernando, M. Á. M. (1990). APLICACIÓN DEL MODELO DE HUFF EN EL ESTUDIO DEL COMERCIO MINORISTA EN LA REGIÓN DE MURCIA. *Papeles de Geografía*, 0(16), 217–235.

Hindle, A., Spollen, M., & Dixon, P. (2004). Developing indicators of the effect of geodemographic factors on cost and performance of public services. *London: DTLR*. Recuperado a partir de <http://www.local.communities.gov.uk/finance/0607/swg0503/swg-05-16full.pdf>

Holderness, T., & Turpin, E. (2015). White paper—PetaJakarta. org: Assessing the role of social media for civic co-management during monsoon flooding in Jakarta, Indonesia. *University of Wollongong, Wollongong*.

Huff, D. L. (1963). A Probabilistic Analysis of Shopping Center Trade Areas. *Land Economics*, 39(1), 81. <https://doi.org/10.2307/3144521>

IEAG, U. (2014). *A World that Counts—Mobilising the Data Revolution for Sustainable Development*. Report prepared for the UN Secretary General by the Independent Expert Advisory Group on a Data Revolution for Sustainable Development.

Instituto Nacional de Estadística y Geografía -INEGI. (s/f). Directorio Estadístico Nacional de Unidades Económicas (DENUE). Recuperado a partir de <http://www.inegi.org.mx/est/contenidos/proyectos/denue/presentacion.aspx>

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

Jácome, J. (2013). Patrones espaciales en las elecciones locales del 2007 en Bogotá, Colombia, *22(1)*, 141–168.

Jácome Molina, J. M. (2015). Aproximación al análisis espacial urbano para la caracterización del sector educativo en Bogotá y su posible uso como apoyo a las políticas públicas. *Revista de Investigación Básica*, 4(1), 43–63.

Kellerman, A. (1981). *Centographic measures in geography* (First Edition edition). Norwich: Geo Abstracts.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. *META Group Research Note*, 6, 70.

Laurent, T., Ruiz-Gazen, A., & Thomas-Agnan, C. (2009). GeoXp: an R package for exploratory spatial data analysis. *TSE Working Paper*, 9. Recuperado a partir de <http://publications.ut-capitole.fr/3239/>

LeBeau, J. L. (1987). The methods and measures of centrography and the spatial dynamics of rape. *Journal of Quantitative Criminology*, 3(2), 125–141.

Li, S., Dragicevic, S., Castro, F. A., Sester, M., Winter, S., Coltekin, A., ... others. (2016). Geospatial big data handling theory and methods: A review and research challenges. *ISPRS Journal of Photogrammetry and Remote Sensing*, 115, 119–133.

Longley, P. A., & Tobón, C. (2004). Spatial Dependence and Heterogeneity in Patterns of Hardship: An Intra-Urban Analysis. *Annals of the Association of American Geographers*, 94(3), 503–519. <https://doi.org/10.2307/3693927>

Lopes, S. D. F. (2011). Geo-segmentación y geo-posicionamiento en el análisis de la preferências de los turistas. *Estudios y Perspectivas en Turismo*, 20(4), 842–854.

López, H. D. (2005). Georeferenciación De Puentes Peatonales En Ciudad De México Y Su Relación Con Peatones Atropellados. *Universidad Nacional Autónoma de México, México, Distrito Federal*, 1–17.

Louail, T., Lenormand, M., Ros, O. G. C., Picornell, M., Herranz, R., Frias-Martinez, E., ... Barthelemy, M. (2014). From mobile phone data to the spatial structure of cities. *Scientific reports*, 4. Recuperado a partir de <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4055889/>

Machado, R. P. P. (2008). PROCESOS DE GEOCODIFICACIÓN URBANA: LOS CASOS DE SÃO PAULO Y BARCELONA. *Revista Catalana de Geografia.[En línea]. Barcelona: Institut Cartogràfic de Catalunya, febrer.* Recuperado a partir de <http://www.academia.edu/download/30961447/103919.pdf>

Mateos, P. (2013). Geovisualización de la población: Nuevas tendencias en la web social. Recuperado a partir de <http://rua.ua.es/dspace/handle/10045/34746>

Mayer-Schönberger, V., & Cukier, K. (2014). *Big Data: A Revolution That Will Transform How We Live, Work, and Think* (Reprint edition). Boston: Eamon Dolan/Mariner Books.

Miller, F. L. (2007). *GIS Tutorial for Marketing*. ESRI Press.

Miller, H. J., & Goodchild, M. F. (2015). Data-driven geography. *GeoJournal*, 80(4), 449–461.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

Miller, H. J., & Han, J. (2009). *Geographic Data Mining and Knowledge Discovery, Second Edition*. CRC Press.

Mitchell, A. (2005). *The Esri Guide to GIS Analysis, Volume 2: Spatial Measurements and Statistics* (1 edition). Redlands, Calif: Esri Press.

Moreno, A., & Bosque, J. (2010). Los modelos de localización óptima como herramientas para la planificación territorial y urbana de instalaciones y equipamientos. *Ciudad y Territorio: Estudios Territoriales*, 461–480.

Moreno Jiménez, A. (1995). Planificación y. gestión de servicios a la población desde la perspectiva territorial. *Boletín de la Asociación de Geógrafos españoles*, (20), 115–134.

Moreno Jimenez, A., Rodríguez Rodríguez, V., Jiménez Blasco, B., & Martínez de Sola, B. (1986). Informática y Geografía. En *Geografía teórica y cuantitativa: concepto y métodos* (pp. 127–146). Universidad de Oviedo.

Mostern, R., Southall, H., & Berman, M. L. (2016). *Placing Names: Enriching and Integrating Gazetteers*. Indiana University Press.

National Academy of Sciences. (2006). *Learning to Think Spatially: GIS as a Support System in the K-12 Curriculum*. Recuperado a partir de <http://www.nap.edu/catalog/11019.html>.

Nyerges, T. L. (1991). Analytical map use. *Cartography and Geographic Information Systems*, 18(1), 11–22.

Ojo, A., Vickers, D., & Ballas, D. (2013). Creating a small scale area classification for understanding the economic, social and housing characteristics of small geographical areas in the Philippines. *Regional Science Policy & Practice*, 5(1), 1–24. <https://doi.org/10.1111/j.1757-7802.2012.01076.x>

Petersen, J., Gibin, M., Longley, P., Mateos, P., Atkinson, P., & Ashby, D. (2011). Geodemographics as a tool for targeting neighbourhoods in public health campaigns. *Journal of Geographical Systems*, 13(2), 173–192. <https://doi.org/10.1007/s10109-010-0113-9>

Presidencia de la República. (2000). Directiva Presidencial 02. Recuperado a partir de <http://www.alcaldiabogota.gov.co/sisjur/normas/Norma1.jsp?i=6399>

Rivero, M. S. (2008). Análisis espacial de datos y turismo: nuevas técnicas para el análisis turístico. Una aplicación al caso extremeño. *Revista de Estudios Empresariales. Segunda Época*, (2). Recuperado a partir de <http://revistaselectronicas.ujaen.es/index.php/REE/article/view/365>

Ruiz-Pérez, M., & Gelabert Grimalt, M. (2012). Análisis de la vulnerabilidad social frente a desastres naturales: el caso de la isla de Mallorca. *GeoSig*, 4, 1–26.

Ruiz-Santacruz, J., & Durán, C. A. (2013). *Análisis espacial multitemporal de la tasa de mortalidad infantil municipal en Colombia*. DEPARTAMENTO ADMINISTRATIVO NACIONAL DE ESTADÍSTICA-DANE. Recuperado a partir de <http://econpapers.repec.org/paper/col000482/012670.htm>

Schabenberger, O., & Gotway, C. (2004). *Statistical methods for spatial data analysis* (CRC press).

Serrano, R. M., & Valcarce, E. V. (2000). *Técnicas econométricas para el tratamiento de datos espaciales: la econometría espacial*. Edicions Universitat Barcelona.

Moreno Mayorga S.L., Jácome Molina, J. M (2017): "Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura", *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

Shekhar, S, Jiang, Z, Ali, R, Eftelioglu, E, Tang, X, & Gunturi, V. (2015). Spatiotemporal Data Mining: A Computational Perspective. *ISPRS International Journal of Geo-Information*, 4, 2306–2338. <https://doi.org/10.3390/ijgi4042306>

Singleton, A. D., & Spielman, S. E. (2013). The past, present and future of geodemographic research in the United States and United Kingdom. *Professional Geographer, Forthcoming*.

Sleight, P. (2004). An Introductory Review of Geodemographic Information Systems. *Journal of Targeting, Measurement and Analysis for Marketing*, 12(4), 379–388.

Smith, S., & Burce, C. (2008). *CrimeStat III, User Workbook*. National Institute of Justice. Recuperado a partir de http://www.icpsr.umich.edu/CrimeStat/workbook/CrimeStat_Workbook.pdf

Song, Y., & Knaap, G.-J. (2007). Quantitative Classification of Neighbourhoods: The Neighbourhoods of New Single-family Homes in the Portland Metropolitan Area. *Journal of Urban Design*, 12(1), 1–24. <https://doi.org/10.1080/13574800601072640>

Starmanss, R. (2016). The Advent of Data Science: some considerations about the Unreasonable Effectiveness of Data. En *Handbook of Big Data* (pp. 3–20). CRC Press.

Turner, A. (2006). *Introduction to neogeography*. O'Reilly Media, Inc. Recuperado a partir de [https://books.google.es/books?hl=es&lr=&id=oHgDv4feV-8C&oi=fnd&pg=PA24&dq=Turner,+A.+\(2006\).+Introduction+to+neogeography.+%22+O%27Reilly+Media,+Inc.%22.&ots=wZtaUHYYcZ&sig=53b7k_18-UipcWTK8ucz1C1v0JQ](https://books.google.es/books?hl=es&lr=&id=oHgDv4feV-8C&oi=fnd&pg=PA24&dq=Turner,+A.+(2006).+Introduction+to+neogeography.+%22+O%27Reilly+Media,+Inc.%22.&ots=wZtaUHYYcZ&sig=53b7k_18-UipcWTK8ucz1C1v0JQ)

United Nations. (2015, julio 2). Integration of geospatial, statistical and other information. Recuperado a partir de <http://ggim.un.org/docs/meetings/GGIM5/E-C20-2015-8%20Integrating%20Geo-Statistical%20Info%20Report.pdf>

Vickers, D., & Rees, P. (2007). Creating the UK National Statistics 2001 output area classification. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, 170(2), 379–403.

Waller, L. A., & Gotway, C. . (2004). *Applied Spatial Statistics for Public Health Data* (Vol. 368). John Wiley & Sons. Recuperado a partir de https://books.google.com/books/about/Applied_Spatial_Statistics_for_Public_He.html?hl=es&id=OuQwgShUdGAC

Webber, R. (2004). Neighbourhood inequalities in the patterns of hospital admissions and their application to the targeting of health promotion campaigns [Working / discussion paper]. Recuperado el 18 de septiembre de 2017, a partir de <http://discovery.ucl.ac.uk/1291/>

Webber, R., & Ashby, D. (2006). *High crime: High disorder neighbourhoods -spatial analysis and geodemographics*. Centre for Advanced Spatial Analysis. Recuperado a partir de http://www.casa.ucl.ac.uk/cjec/phd/downloads/AuditCommission_Report.pdf

Wieland, T. (2017). Market Area Analysis for Retail and Service Locations with MCI. *R JOURNAL*, 9(1), 298–323.

Wolfram, S. (1983). Statistical mechanics of cellular automata. *Reviews of Modern Physics*, 55(3), 601–644. <https://doi.org/10.1103/RevModPhys.55.601>

Yrigoyen, C. (1996). Interacción espacial y regional: análisis de los flujos comerciales de Castilla y León. En *5º Congreso de Economía Regional de Castilla y León: Comunicaciones*, Vol. 2, 1996

Moreno Mayorga S.L., Jácome Molina, J. M (2017): “Usos de información georreferenciada para prestación de servicios a la población: una revisión de literatura”, *GeoFocus (Artículos)*, nº 20, p. 201-229. ISSN: 1578-5157 <http://dx.doi.org/10.21138/GF.572>

(*Comunicaciones 2*), págs. 681-694 (pp. 681–694). Consejería de Economía y Hacienda. Recuperado a partir de <https://dialnet.unirioja.es/servlet/articulo?codigo=4686256>

Yrigoyen, C. (1997). Modelos de determinacion de areas de mercado de comercio al por menor.

Yrigoyen, C. (2004). *Modelos De Heterogeneidad Espacial*. EconWPA. Recuperado a partir de <http://core.ac.uk/download/pdf/9310145.pdf>

Yrigoyen, C., & Calderón, G. F. A. (2009). *Análisis de Datos Espacio-Temporales Para la Economía Y El Geomarketing*. Netbiblio, S. L.

