

Artículo de investigación

Modelado espacial y exploratorio de la enfermedad renal crónica en la zona metropolitana de San Luis Potosí, México

Karla Issel Lara Rojas^{1*}, Jaqueline Calderón Hernández², Idida Eunice Tello Cruz³

¹ Programa Multidisciplinario de Posgrado en Ciencias Ambientales, Universidad Autónoma de San Luis Potosí, México.

² Coordinación para la Innovación y Aplicación de la Ciencia y la Tecnología, Universidad Autónoma de San Luis Potosí, México.

³ Facultad de Medicina, Universidad Autónoma de San Luis Potosí, México.

* E-mail: issellr@live.com

Recibido: 22/06/2021; Aceptado: 01/08/2021; Publicado: 30/12/2021

Resumen

El modelaje espacial permite analizar en patrones espaciales el comportamiento y el grado de autocorrelación de factores de riesgo medioambientales y sociodemográficos involucrados en el desarrollo de una enfermedad. El objetivo fue modelar la dinámica espacial con factores de riesgo medioambiental y sociodemográficos en los conglomerados identificados en la zona metropolitana de San Luis Potosí, México. De los casos residentes dentro de los conglomerados, se calculó la autocorrelación espacial mediante el Índice de Morán con el software GeoDa. Se construyó un modelo de correlación espacial con el software GWR4.09 con los factores de riesgo medioambiental y sociodemográficos. Los resultados del índice de Morán mostraron autocorrelación positiva en la distribución de los casos residentes dentro de los tres conglomerados (0.552, 0.581 y 0.763, respectivamente ($p=0.001$)). El modelo de ajuste a nivel local entre el riesgo relativo de desarrollar enfermedad renal crónica con la concentración (mg/dL) de flúor y el índice de marginación arrojó un coeficiente de regresión de R y R² ajustado de 0.82 y 0.79 respectivamente. La metodología propuesta permitió relacionar factores de riesgo presentes en las zonas identificadas con el desarrollo de la enfermedad para tener un mejor entendimiento de su comportamiento a nivel de área geostatística básica.

Palabras clave: modelado espacial, análisis exploratorio, insuficiencia renal crónica.

Spatial and exploratory modeling of chronic kidney disease in the metropolitan area of San Luis Potosí, Mexico

Abstract

Spatial modeling makes it possible to analyze the behavior and degree of autocorrelation of environmental and sociodemographic risk factors involved in the development of a disease in spatial patterns. The objective was to model the spatial dynamics with environmental and sociodemographic risk factors in the conglomerates identified in the metropolitan area of San Luis Potosi, Mexico. From the resident cases within the clusters, spatial autocorrelation was calculated using the Moran Index with the GeoDa software. A spatial correlation model was built with the GWR4.09 software with the environmental and sociodemographic risk factors. The results of the Moran index showed positive autocorrelation in the distribution of resident cases within the three clusters (0.552, 0.581 and 0.763,

respectively ($p = 0.001$)). The local adjustment model between the relative risk of developing chronic kidney disease with the concentration (mg / dL) of fluoride and the marginalization index yielded a regression coefficient of R and adjusted R² of 0.82 and 0.79 respectively. The proposed methodology made it possible to relate risk factors present in the areas identified with the development of the disease to have a better understanding of its behavior at the level of the basic geostatistical area.

Keywords: spatial modeling, exploratory analysis, chronic kidney disease.

1. Introducción

Con el diseño de técnicas espaciales y una mayor accesibilidad a softwares especializados se han desarrollado herramientas eficaces para identificar áreas geográficas que presenten altas tasas de prevalencia, incidencia o mortalidad de aquellas enfermedades de interés; además identifican y predicen la presencia de conglomerados que poseen un riesgo elevado de desarrollar una condición de salud y un exceso en el número de casos (Baneerje et al, 2016) (Kang et al, 2016). El modelaje espacial permite analizar el comportamiento y el grado de autocorrelación de los casos residentes dentro de estos patrones espaciales, así como en las unidades vecinas (Celemín, 2009) y comprobar hipótesis que relacionen la proximidad y similitud de factores de riesgo medioambientales o sociodemográficos involucrados dentro de las unidades espaciales (Kang et al, 2016) a través de la estructuración de un modelo que relacione la presencia de la enfermedad donde el conjunto de estos factores presentan un mayor o menor poder explicativo (Gutiérrez-Puebla et al, 2012). El análisis de los datos obtenidos se integra en un sistema de información geográfico (SIG) como herramienta para visualizar, consultar y presentar el comportamiento de la enfermedad y su correlación con variables explicativas en el espacio geográfico para mejorar los estudios de salud ambiental, evaluar los riesgos y mejorar las necesidades de atención médica (Shaw et al, 2017).

En un artículo anterior, abordamos la enfermedad renal crónica (ERC), un problema de salud pública en México, en el que se estima que uno de cada diez mexicanos se encuentra en algún estadio de la enfermedad (CUCS, 2020). Nuestra zona de estudio abarcó la zona metropolitana San Luis Potosí-Soledad de Graciano Sánchez (ZM SLP-SGS), capital del estado de SLP, México; a través de un análisis puramente espacial, se detectaron tres conglomerados estadísticamente significativos ($p \leq 0.05$) con alto riesgo relativo (RR) de desarrollar la enfermedad, dentro de los cuales se identificaron factores de riesgo medioambiental y sociodemográficos que pudiesen estar involucrados en el exceso de casos presentes en estas zonas. Para determinar este patrón de distribución, se requiere un índice de comparación estadístico; como el índice de Morán (IM) para evaluar la autocorrelación espacial (Celemín, 2009) y para determinar la proximidad y similitud de los posibles factores de riesgo identificados, la técnica de Regresión Geográficamente Ponderada (GWR, por sus siglas en inglés), nos permite realizar la estructuración de un modelo que relacione la presencia de una variable dependiente (ERC) y un conjunto de variaciones heterogéneas de las que se conoce su ubicación dentro de los conglomerados para estimar un coeficiente de regresión local (R²) en cada una de las unidades geográficas donde las variables independientes tienen un mayor, o menor, poder explicativo (Gutiérrez-Puebla et al, 2012) en el comportamiento de la ERC en la zona de estudio para manejar adecuadamente la salud poblacional expuesta y continuar vinculando la información en un sistema de vigilancia epidemiológico para el diseño de intervenciones en salud pública. El objetivo de la investigación es modelar la dinámica espacial en los conglomerados identificados con los factores de riesgo medioambiental y sociodemográficos involucrados en las zonas de estudio con exceso de casos de ERC en la ZM SLP-SGS.

2. Materiales y métodos

2.1. Base de datos de casos residentes en los conglomerados: Partiendo del análisis puramente espacial realizado previamente con el software SatScanTM, de los tres conglomerados identificados ubicados al centro-sur (RR 5.61), suroeste (RR 2.23) y noroeste (RR 3.1) de la ZM SLP-SGS, se prosiguió a extraer los casos residentes en estas zonas y se concentraron en una base de datos, que incluían características sociodemográficas, clínicas y su ubicación geográfica en el sistema de coordenadas World Geodetic System 1984 (WGS84). Cada base se convirtió a archivo Shapefile usando el software ArcGis versión 10.2 (Environmental Systems Research Institute, 2013).

2.2. Autocorrelación espacial: Para conocer el grado de concentración (autocorrelación positiva) o dispersión (autocorrelación negativa) de los casos de ERC dentro de los conglomerados estadísticamente significativos, se utilizó el software GeoDaTM versión 9/16/2018 subversión 1.12.1.161 (Geodacenter, 2018). Como insumo requerido, se ingresaron los archivos Shapefile correspondientes de cada conglomerado, de los cuales se prosiguió a calcular un gestor de pesos, con contigüidad de primer orden tipo reina mediante la selección de una variable de matriz para calcular el estadístico de IM local. La hipótesis nula que se planteó en éste análisis fue que la distribución geográfica de los casos dentro de los conglomerados es completamente aleatoria. El modelo se corrió bajo una simulación de 999 permutaciones de Monte Carlo.

2.3. Correlación espacial de variables: Partiendo de los factores de riesgo ambientales y sociodemográficos identificados en la primer parte de la investigación; se obtuvo la cartografía de uso de suelo y vegetación escala 1:250000, serie VI (continuo nacional) del portal de la Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO, 2020) para identificar algún tipo de suelo relacionado con la agricultura en las Áreas Geoestadísticas Básicas (AGEB) ubicados en la zona noroeste y la cartografía de marginación por AGEB urbana 2010 del portal del Consejo Nacional de Población (CONAPO, 2020), para obtener el índice de marginación de cada uno de los AGEB que conforman los tres conglomerados estadísticamente significativos.

Así mismo, de la revisión de documentos de investigación científica (Hergt, 2009) (Stevens-Vázquez, 2012) y de la Plataforma de Evaluación de Aguas Subterráneas (GAP, por sus siglas en inglés) (Groundwater Assessment Platform, 2018), se obtuvieron las concentraciones (mg/dL) de flúor (F) y arsénico (As) de los pozos que abastecen agua potable en estos AGEB. Con esta información, se conformó una base de datos que se introdujo al software GWR4.09 (School of Geographical Sciences & Urban Planning, 2020) para la construcción del modelo de correlación espacial. Las variables de ubicación fueron las claves geoestadísticas de los AGEB pertenecientes a los conglomerados estadísticamente significativos y sus respectivas coordenadas en el sistema WGS84, la variable dependiente fue el RR de desarrollar ERC (calculado anteriormente con el software SatScanTM) en cada AGEB, las variables independientes fueron el índice de marginación, las respectivas concentraciones (mg/dL) de F y As; así como la superficie agrícola. La configuración del modelo fue de tipo gaussiano y en la asignación de pesos, se utilizó la función de Kernel adaptativo y la búsqueda de banda dorada (opciones más utilizadas) para determinar automáticamente el mejor tamaño de ancho de banda y para la medida de bondad de ajuste del modelo se utilizó el criterio de información de Akaike (AICc).

2.4. Representación espacial: El grado de correlación de las variables en los AGEB que conforman los conglomerados, se representó en un SIG mediante el software ArcGis versión 10.2 (Environmental Systems Research Institute, 2013) en el sistema de coordenadas WGS84 escala 1:1 50 000.

3. Resultados

El estadístico de IM mostró una autocorrelación positiva en la distribución de los casos residentes dentro de cada uno de los tres conglomerados. De acuerdo al valor de significancia p obtenido para cada uno de ellos ($p = 0.001$) y valores de puntuaciones z (equivalentes a las desviaciones estándar dentro de una curva de distribución normal) presentan valores positivos (los casos tienden a agruparse), mostrando así que los patrones espaciales formados por los casos no se distribuyen aleatoriamente, por lo tanto, la hipótesis nula es rechazada (Tabla 1), indicándonos que los casos residentes dentro de los conglomerados presentan un componente geográfico.

Tabla 1.- Análisis de autocorrelación espacial con Índices de Morán de los casos residentes dentro de los conglomerados estadísticamente significativos.

	Conglomerado 1	Conglomerado 2	Conglomerado 3
Índice de Morán	0.552	0.581	0.763
Puntuaciones z	5.1208	5.0873	9.7965
Valor de p	0.001	0.001	0.001
RR	5.61	3.10	2.23
Ubicación	Noroeste	Centro-Sur	Suroeste

Elaboración propia

En el análisis de las variables independientes, sobre la capa de uso de suelo y vegetación, esta no presentó relación con algún tipo de agricultura que se ejerciera en los AGEB que se ubican al noroeste en la periferia de la ZM SLP-SGS. Esta capa se tomó en cuenta, ya que, en la primera parte de éste artículo, en la exploración de factores de riesgo medioambiental, se consideró que la cercanía de estos AGEB a zonas agrícolas pudiera contribuir al desarrollo de la enfermedad.

Se realizaron diversas corridas para determinar el mejor modelo de ajuste de la variable dependiente (RR de desarrollar ERC) con las variables independientes como predictores del comportamiento geográfico de la enfermedad dentro de los conglomerados; las variables de mejor ajuste fueron la concentración (mg/dL) de F en agua de pozos y el índice de marginación.

El ajuste del modelo de regresión global (RG) arroja un coeficiente de regresión de R y ajustado R^2 de 0.328654 y 0.254060 respectivamente; es decir, que sólo éstas dos variables independientes, explican alrededor del 25% del exceso de casos de enfermedad en los conglomerados identificados. Sin embargo, al análisis de GWR, muestra un R y ajustado R^2 de 0.819132 y 0.791309 respectivamente, lo que significa que, a nivel geográfico específico de los AGEB dentro de los conglomerados, existe una mejor correlación entre estas dos variables independientes con el RR de desarrollar ERC. El criterio de bondad de ajuste AICc se redujo entre los dos modelos, esto significa que el modelo tiene un mejor ajuste a nivel de los AGEB que conforman los conglomerados estadísticamente significativos que a nivel de toda la ZM (Tabla 2).

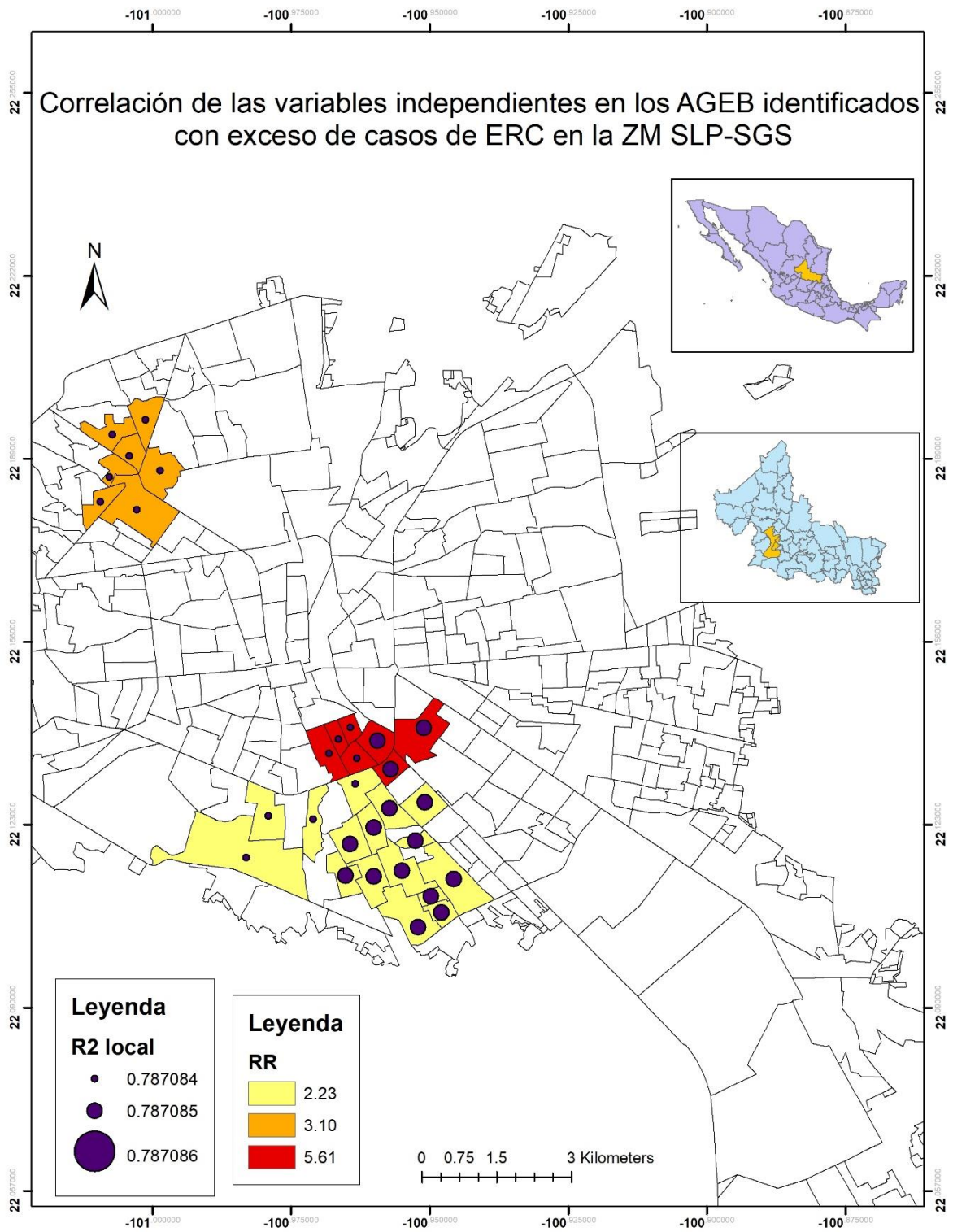
Tabla 2.- Resultados entre los parámetros del modelo de regresión global y modelo de regresión geográficamente ponderada.

Variables	Modelo de RG	Modelo de GWR			
		Min	Max	Media	DE
Intercepto	4.567521	0.10584	7.03800	6.81425	1.22478
Concentración de F	-0.803486	-1.5065	0.46556	-1.4428	0.34843
Índice de marginación	-0.976223	-2.2030	-0.2984	-0.3599	0.33650
R	0.328654	0.819132			
R ²	0.254060	0.791309			
AICc	108.496995	70.643578			
No. de parámetros	3	3			
Suma de residuales cuadrados	44.182152	11.903189			
Sigma global estimada	1.256158	0.619656			
Ancho de banda	-----	33.402679			
Número efectivo de parámetros	-----	3.980494			
Grados de libertad	-----	27.019506			

Elaboración propia

En la figura 1 se representa espacialmente los conglomerados estadísticamente significativos con su respectivo RR y su relación con el R² obtenido del análisis de GWR.

Figura 1



Elaboración propia

3. Discusión

Debido a que las estadísticas y las historias clínicas del sector salud no involucran factores de riesgo medioambiental ni de carácter social, la metodología aquí propuesta nos permitió relacionar estas variables a través del uso de técnicas de epidemiología espacial, nos permitió tener un mejor entendimiento del comportamiento de la enfermedad mediante la exploración del espacio geográfico; ya que al identificar zonas con exceso de casos; estas deben ser tomadas como prioritarias para la atención por parte del sector salud; no sólo en el enfoque clínico de las causas médicas “tradicionales” que desencadenan la enfermedad como lo son la diabetes mellitus y la hipertensión arterial (Ávila-Saldívar et al, 2013), sino como observamos, en un primer paso, al identificar conglomerados en el espacio geográfico, esto nos permite crear hipótesis sobre la presencia de otros factores involucrados al conocer mecanismos de acción de agentes medioambientales reportados ampliamente en la literatura científica y que deben ser considerados para el conocimiento de la enfermedad y sobre todo incluir las características sociodemográficas (nivel de educación, calidad de vivienda, servicios públicos, ingresos económicos, etc.); ya que también la calidad de la salud humana depende de la calidad del medio ambiente, ya que es notable mencionar que en aquellos lugares donde existe mayor desigualdad social, existe un mayor nivel de vulnerabilidad al padecimiento de enfermedades ya que se encuentran más expuestos a los contaminantes ambientales y con ello aumenta el riesgo de sufrir efectos negativos en la salud poblacional.

En conjunto, estas determinantes proximales nos llevan a corroborar estas hipótesis al determinar un patrón de autocorrelación espacial (calculado mediante el IM), las cuales se deben a la similitud en sus características sociodemográficas y en la exposición a factores de riesgo medioambiental con ubicación geográfica específica. El análisis de GWR, permite detectar aquellos factores que presentan mayor peso dentro del modelo espacial y que explican el exceso de casos de la enfermedad; ya que como se describió en la primera parte del artículo, debido a la naturaleza geológica del subsuelo, el F presente en el agua subterránea que abastece en la ZM SLP-SGS (Hergt, 2009) (Stevens-Vázquez, 2012), principalmente en los AGEB pertenecientes a los conglomerados estadísticamente significativos, este excede el límite establecido en la Norma Oficial Mexicana NOM-127 (≤ 1.5 mg/dL) (SEGOB, 2019), es de notar, que principalmente, en el conglomerado ubicado al suroeste es el que presenta las concentraciones más altas de F en pozos de agua de consumo humano (3.1 - 3.3 mg/dL) (Hergt, 2009) (Stevens-Vázquez, 2012), el análisis muestra que en esta zona se presenta el mayor índice de autocorrelación espacial ($IM = 0.763$) y con el mayor coeficiente de regresión ajustado ($R^2 = 0.787086$) a nivel local, estas zonas presentan un índice de marginación medio y alto. La misma situación se observa en los conglomerados ubicados al centro-sur (índice de marginación medio) y al noroeste (índice de marginación alto).

La visualización de las enfermedades a un nivel de mínima división territorial ha sido ampliamente estudiada por la Unidad de Estadísticas de Salud de Área Pequeña (SAHSU, por sus siglas en inglés) del Colegio Imperial de Londres, la cual evalúa el riesgo para la salud de la población debido a factores ambientales mediante el uso de datos de salud y población recopilados de forma rutinaria a una escala pequeña. Con este objetivo, SAHSU posee conjuntos completos de datos demográficos y de salud y una variedad de conjunto de datos ambientales de alta resolución espacial para obtener una alerta temprana de cualquier efecto adverso para la salud poblacional (Imperial Collage London, 2020). Por ello, la importancia de emplear estas metodologías al explorar el espacio geográfico, es comprender mejor el comportamiento de la ERC, y de cualquier otra enfermedad que represente un desafío a la salud pública de cualquier sistema de salud, con el objetivo de crear una red de seguimiento, la cual nos permita conjuntar la información de las diversas instituciones de salud, mediante técnicas geoespaciales; como las que aquí se emplean, nos permitan realizar una vigilancia epidemiológica, geográfica y pasiva de la enfermedad (identificación de conglomerados), la información generada permite la planeación de acciones basadas en la evidencia; mediante la mejora

continua, se evalúan y se perfeccionan las estrategias, para el diseño de programas de prevención e intervención, al conocer las causas distales (factores ocupacionales, ambientales y sociales) y proximales que determinen el comportamiento de la enfermedad, llevan a la mejora en la atención de los servicios de salud, lo que disminuiría la saturación de los servicios especializados, el gasto institucional en terapias sustitutivas y mejorar la salud poblacional.

Referencias bibliográficas

Ávila-Saldívar M, Conchillos-Olivares G, Rojas-Báez IC, Ordoñez-Cruz AE, Ramírez-Flores HJ. Enfermedad renal crónica: causa y prevalencia en la población del Hospital General La Perla. *Med Int Mex* [Internet]. Septiembre-octubre 2013 [consultado 13 de junio de 2020]; 29 (5): [aproximadamente 1 p.]. Disponible en: <https://www.medigraphic.com/pdfs/medintmex/mim-2013/mim135e.pdf>

Baneerje S. Spatial Data Analysis. *Annu. Rev. Public Health* 2016, 37. DOI: 10.1146/annurev-publhealth-032315-021711.

Celemín JP. Autocorrelación espacial e indicadores de asociación espacial. Importancia, estructura y aplicación. *Revista Universitaria de Geografía* [Internet]. 2009 [consultado el 16 de junio del 2020]; 18 [aproximadamente 1 p.]. Disponible en: <https://www.redalyc.org/comocitar.oa?id=383239099001>

Comisión Nacional para el Conocimiento y Uso de la Biodiversidad [Internet], México: Uso de uso de suelo y vegetación, escala 1:250000, serie VI (continuo nacional) [consultado el 15 de mayo del 2020] [aproximadamente 2 pantallas]. Disponible en: http://www.conabio.gob.mx/informacion/metadatos/gis/usv250s6gw.xml? httpcache=yes& xsl=/db/metadatos/xsl/fgdc_html.xsl& indent=no

Consejo Nacional de Población [Internet]. México: Índices de marginación [consultado el 15 de mayo del 2020]. Datos abiertos del índice de marginación [aproximadamente 2 pantallas]. Disponible en: https://www.conapo.gob.mx/es/CONAPO/Datos_Abiertos_del_Indice_de_Marginacion

cucs.udg.mx [Internet]. Centro Universitario de Ciencias de la Salud; c1997-2019 [consultado el 20 de junio de 2020]. Disponible en: cucs.udg.mx/noticias/

Environmental Systems Research Institute. ArcGis version 10.2 for desktop [software]. c1999-2013 [consultado el 19 de junio de 2020]. Sitio web: <https://www.arcgis.com>. Se requiere licencia.

geodacenter.github.io [Internet]. GeoDa an Introduction to Spatial Data Analysis. Version 9/16/2018 subversión 1.12.1.16 [software]. 2020 [consultado el 19 de junio de 2020]. Disponible en: <https://geodacenter.github.io/download.html>

Groundwater Assessment Platform [Internet]. Swiss: Swiss Agency for Development and Cooperation SDC. c2018. Groundwater Assessment Platform Maps [consultado el 20 de mayo del 2020]. Disponible en: <https://www.gapmaps.org/Home/Public>

Gutiérrez-Puebla J, García-Palomares JC, Daniel-Cardozo O. Regresión Geográficamente Ponderada (GWR) y estimación de la demanda de las estaciones del Metro de Madrid [Internet]. Madrid: XV Congreso Nacional de Tecnologías de la Información Geográfica; 2012 [consultado el 16 de junio del 2020]. Disponible en: http://tig.age-geografia.es/2012_Madrid/ponencia2/GutierrezPuebla_final_par.pdf

Hergt T. Diseño optimizado de monitoreo de la calidad del agua de los sistemas de flujo subterráneo en el acuífero 2411 "San Luis Potosí": Hacia un manejo sustentable. Universidad Autónoma de San Luis Potosí; 2009.

Imperial Collage London [Internet]. Londres, Reino Unido; c2020 [consultado el 20 de junio de 2020]. Small Area Health Statistics Unit; [alrededor de 2 pantallas]. Disponible en: <https://www.imperial.ac.uk/school-public-health/epidemiology-and-biostatistics/small-area-health-statistics-unit/>

Kang SY, Cramb SM, White NM, Ball SJ, Mengersen KL. Making the most of spatial information in health; a tutorial in Bayesian disease mapping for areal data. *Geospat. Health* 2016, 11 (428). DOI: 10.4081/gh.2016.428.

School of Geographical Sciences & Urban Planning [Internet]. GW4R (Geographically Weighted Regression) versión 4.09 [software]. 2020 [consultado el 22 de junio de 2020]. Disponible en: <https://sgsup.asu.edu/sparc/gwr4>

Secretaría de Gobernación SEGOB [Internet]. Ciudad de México: Diario Oficial de la Federación c2019 [consultado el 11 de junio de 2020]. Proyecto de Norma Oficial Mexicana PROY-NOM-127-SSA-2017, Agua para uso y consumo humano. Límites permisibles de la calidad del agua [alrededor de 2 pantallas]. Disponible en: https://dof.gob.mx/nota_detalle.php?codigo=5581179&fecha=06/12/2019

Shaw NT, McGuire SK. Understanding the use of the geographical information systems (GISs) in health informatics research: a review. *J Innov Health Inform.* 2017, 24 (2). DOI: <http://dx.doi.org/10.14236/jhi.v24i2.940>

Stevens-Vázquez GS. Análisis espacial de la vulnerabilidad hídrica en la zona metropolitana de la ciudad de San Luis Potosí-Soledad de Graciano Sánchez, México. Universidad Autónoma de San Luis Potosí, 2012.



Esta obra se encuentra bajo Licencia Creative Commons Reconocimiento-NoComercial 4.0. Internacional. Reconocimiento - Permite copiar, distribuir, exhibir y representar la obra y hacer obras derivadas siempre y cuando reconozca y cite al autor original. No Comercial - Esta obra no puede ser utilizada con fines comerciales, a menos que se obtenga el permiso.