



Análisis de agrupamiento espacial de la letalidad por COVID-19 en México

Núñez, Juan Manuel; Galeana-Pizaña, Mauricio; Jiménez-Ortega, Aldo Daniel; Quiroz-Cazares, Gabriela; Balderas-Cruz, Iyari; Seemann-Carús, Sofia; Ordorica-Mellado, Manuel; Lara-Pulido, José Alberto
Análisis de agrupamiento espacial de la letalidad por COVID-19 en México
CIENCIA *ergo-sum*, Número especial “SARS-CoV-2” 2021 | e141
Universidad Autónoma del Estado de México, México

Esta obra está bajo una Licencia Creative Commons Atribución-NoComercial-SinDerivar 4.0 Internacional.

Núñez, J. M., Galeana-Pizaña, M., Jiménez-Ortega, A. D., Quiroz-Cazares, G., Balderas-Cruz, I., Seemann-Carús, S., Ordorica-Mellado, M. y Lara-Pulido, J. A. (2021). Análisis de agrupamiento espacial de la letalidad por COVID-19 en México. CIENCIA *ergo-sum*, Número especial “SARS-CoV-2”. <https://doi.org/10.30878/ces.v28n4a2>

Análisis de agrupamiento espacial de la letalidad por COVID-19 en México

Spatial grouping analysis of COVID-19 lethality in Mexico

Juan Manuel Núñez

Universidad Iberoamericana, México

juan.nunez@ibero.mx

 <http://orcid.org/0000-0002-9835-0599>

Recepción: 10 de marzo de 2021

Aprobación: 03 de junio de 2021

Mauricio Galeana-Pizaña

Centro de Investigación en Ciencias de Información Geoespacial, México

mgaleana@centrogeo.edu.mx

 <http://orcid.org/0000-0003-3623-6851>

Aldo Daniel Jiménez-Ortega

Centro de Investigación en Ciencias de Información Geoespacial, México

ajimenez@centrogeo.edu.mx

 <http://orcid.org/0000-0003-1183-138X>

Gabriela Quiroz-Cazares

Centro de Investigación en Ciencias de Información Geoespacial, México

gquiroz@centrogeo.edu.mx

 <http://orcid.org/0000-0001-7221-7987>

Iyari Balderas-Cruz

Universidad Iberoamericana, México

iyari.balderas@correo.uia.mx

 <http://orcid.org/0000-0003-3013-9719>

Sofía Seemann-Carús

Universidad Iberoamericana, México

sofia.seemann@correo.uia.mx

 <http://orcid.org/0000-0001-9020-1444>

Manuel Ordorica-Mellado

El Colegio de México, A. C.

mordori@colmex.mx

 <http://orcid.org/0000-0002-7649-1729>

José Alberto Lara-Pulido

Universidad Iberoamericana, México

jose.lara@ibero.mx

 <http://orcid.org/0000-0003-1484-8451>

RESUMEN

Se lleva a cabo un análisis de agrupamiento espacial para identificar patrones espaciales y caracterizar los efectos de la letalidad por COVID-19 en México. Se explica en un nivel municipal las similitudes y diferencias en la tasa de letalidad acumulada del 22 de abril al 6 de agosto de 2020 en relación con el perfil demográfico de la población, factores socioeconómicos y ambientales, así como la accesibilidad a establecimientos de salud. Los resultados muestran que, si bien la mayor parte de la población agrupada mantiene una tasa de letalidad promedio en los centros urbanos, los grupos más aislados presentan tasas de letalidad inferiores a la media nacional. Las tasas de letalidad más altas se alcanzan en los municipios rurales e indígenas, donde las personas son más vulnerables.

PALABRAS CLAVE: análisis espacial, COVID-19, vulnerabilidad, capital natural, accesibilidad.

ABSTRACT

A spatial clustering analysis was implemented to identify spatial patterns and characterize the effects of COVID-19 lethality in Mexico. The goal of the research is to explain at the municipal level the similarities and differences in cumulative case lethality rate from April 22 to August 6, in relation to the demographic profile of the population, socioeconomic and environmental factors, as well as accessibility to health care. The obtained results show that, while most of the grouped population maintains an average lethality rate in urban centers, the most isolated groups maintain lower lethality rates than the national average. The highest lethality rates are reached in rural and indigenous municipalities, where people are deeply vulnerable to COVID-19.

KEYWORDS: spatial analysis, COVID-19, vulnerability, natural capital, accessibility.

INTRODUCCIÓN

La crisis sanitaria que se vive por el COVID-19, originado por el SARS-CoV-2, ha modificado drásticamente la vida de miles de millones de personas en todo el mundo y se ha convertido en un foco creciente de atención global, además de un factor crítico en la toma de decisiones de salud pública. Para combatirlo, los países de todo el mundo y las organizaciones internacionales han tomado diversas medidas, incluida la promoción de la transparencia y el acceso público a los datos sobre enfermedades a partir del uso de mapas que han jugado un papel importante (Gao *et al.*, 2020).

El uso de enfoques de análisis espacial y tecnologías geoespaciales como los Sistemas de Información Geográfica (SIG) y mapeo en línea (web) han sido elementos importantes para estudiar el COVID-19 (Franch-Pardo *et al.*, 2020). Un ejemplo de su impacto ha sido el desarrollo de diversas herramientas como la plataforma que puso en función la Universidad Johns Hopkins, la cual está dando soporte a un sitio web global con actualizaciones permanentes (Dong *et al.*, 2020). Boulos y Geraghty (2020) describen y dan seguimiento a una serie de aplicaciones geográficas a la enfermedad por COVID-19 a medida que evoluciona en todo el mundo. A partir de una revisión de 63 artículos científicos sobre la dimensión geográfica del COVID-19, Franch-Pardo *et al.* (2020) proponen cinco categorías distintas de mapeo de la enfermedad: *a)* análisis espaciotemporal, *b)* geografía social y de salud, *c)* variables ambientales, *d)* minería de datos y *e)* mapeo basado en la web. Otros trabajos exploran cómo la transmisión geográfica del número reproductivo del COVID-19 es mayor en comparación con el coronavirus del SARS (Liu *et al.*, 2020).

En México diversas plataformas gubernamentales han sido desarrolladas para brindar información sobre el COVID-19. La Secretaría de Salud en colaboración con el Consejo Nacional de Ciencia y Tecnología (Conacyt) todos los días pone a disposición de la ciudadanía información de la Dirección General de Epidemiología (DGE) acerca del COVID-19 para la rápida visualización de métricas epidemiológicas, seguimiento espacial de casos confirmados, sospechosos, negativos y defunciones (DGE, 2020). El INEGI, por su parte, ha desarrollado una serie de herramientas y levantamientos especiales que proporcionan insumos fundamentales para el análisis de la epidemia en nuestro país (INEGI, 2020a). El CONEVAL (2020) desarrolló el Visor geoespacial de la pobreza y la COVID-19. Igualmente, la Universidad Nacional Autónoma de México brinda información, mapas y recursos de interés sobre la evolución temporal y espacial en un nivel municipal del COVID-19 en México (Ghilardi *et al.*, 2020), así como el desarrollo de un índice de vulnerabilidad ante COVID-19 que incorpora consideraciones demográficas, socioeconómicas y de salud para la construcción de un mapa nacional con detalle por municipios (Suárez-Lastra *et al.*, 2020).

En este artículo se emplea el análisis de agrupamiento espacial mediante SIG para identificar el patrón espacial y caracterizar la letalidad acumulada a nivel municipal durante la Fase 3 por la epidemia de COVID-19 en México. Para ello, se empleó el enfoque de análisis espacial con el objetivo de implementar el método de agrupamiento que identifica patrones espaciales entre unidades con características similares. El análisis de agrupamiento realiza un procedimiento que intenta agrupar datos en grupos naturales basados en uno o más atributos especificados (Assunção *et al.*, 2006). Los resultados del agrupamiento están basados en las variables de análisis de entrada, el método de agrupamiento y el número de grupos especificados. Así, el estudio busca resaltar el papel del análisis espacial en la interpretación de los efectos y consecuencias del COVID-19 en México para avanzar en la comprensión de sus implicaciones demográficas, sociales, económicas, ambientales y de salud.

Lo anterior bajo la guía de las siguientes preguntas: ¿Los municipios con mayores tasas de letalidad son los más vulnerables? ¿Los municipios más aislados se mantienen con bajas tasas de letalidad? Para alcanzar el objetivo planteado se generó una base de datos municipal con la estimación de la tasa de letalidad acumulada desde el inicio de la Fase 3 (21 de abril de 2020) por la epidemia del COVID-19, el índice de vulnerabilidad ante el COVID-19 (Suárez-Lastra *et al.*, 2020), el índice de sustentabilidad de capital natural (Mora, 2019) y un indicador compuesto de accesibilidad a los establecimientos de salud.

La organización de este artículo parte de la introducción, la cual plantea los objetivos, así como una serie de preguntas que orientan el trabajo. En la siguiente sección se presentan las variables empleadas, el método de aprendizaje automático propuesto para el agrupamiento de los datos y los parámetros introducidos. Los resultados con los grupos obtenidos, así como su caracterización, se presentan en la siguiente sección. Finalmente, las conclusiones muestran una serie de reflexiones acerca del trabajo y las preguntas que guiaron esta investigación.

1. MATERIALES Y MÉTODOS

1. 1. Tasa de letalidad acumulada

Como se sabe, analizar efectos del virus en humanos compete a la ciencia biomédica, pero su propagación por el mundo involucra principios geográficos (García de León-Loza, 2020). La evolución de esta pandemia, en términos de tasas de incidencia, mortalidad y letalidad, es heterogénea, con diferencias entre países e incluso entre regiones del mismo país, lo que requiere comprender los elementos que impulsan su comportamiento (De Figueiredo *et al.*, 2020).

La mortalidad privilegia la cuantificación de defunciones y sus promedios respecto a la población total (García de León-Loza, 2020). La letalidad es un indicador que depende de la capacidad de identificar casos y muertes y, por lo tanto, sintetiza dos elementos cruciales de un proceso epidémico: la magnitud y el impacto (De Figueiredo *et al.*, 2020). Por ello, se considera que la letalidad es un estadístico muy importante para evaluar el resultado de la terapia (atención) que reciben los casos graves y críticos (Simón-Domínguez *et al.*, 2020).

De acuerdo con Porta (2014), la tasa de letalidad es la proporción de casos de una condición específica que son fatales dentro de un tiempo especificado, cuya fórmula se muestra a continuación:

$$\text{Tasa de letalidad} = \frac{\text{Número de muertes de una enfermedad} \\ \text{(en un periodo dado)}}{\text{Número de casos diagnosticados de esa enfermedad} \\ \text{(en el mismo periodo)}} * 100$$

En este sentido, la tasa de letalidad es un indicador muy útil, ya que permite conocer la respuesta de los servicios de salud a los enfermos diagnosticados por COVID-19 en términos del peor resultado posible: la muerte.

La letalidad del COVID-19 que reporta cotidianamente la Secretaría de Salud desde el 13 de abril de 2020 es válida, aunque es necesario aclarar que se trata de una cifra subestimada, porque se trata de casos de COVID-19 que fueron diagnosticados por los servicios de salud, lo cual no incluye los casos asintomáticos que nunca llegaron a los servicios y, por ende, no están registrados (DGE, 2020).

Con el objetivo de identificar tendencias en los datos para definir la temporalidad del estadístico de letalidad que se utilizará en el modelo, de acuerdo con los datos reportados en la base de datos de la Plataforma de información geográfica de la UNAM sobre COVID-19 en México (Ghilardi *et al.*, 2020), se obtuvo la tasa de

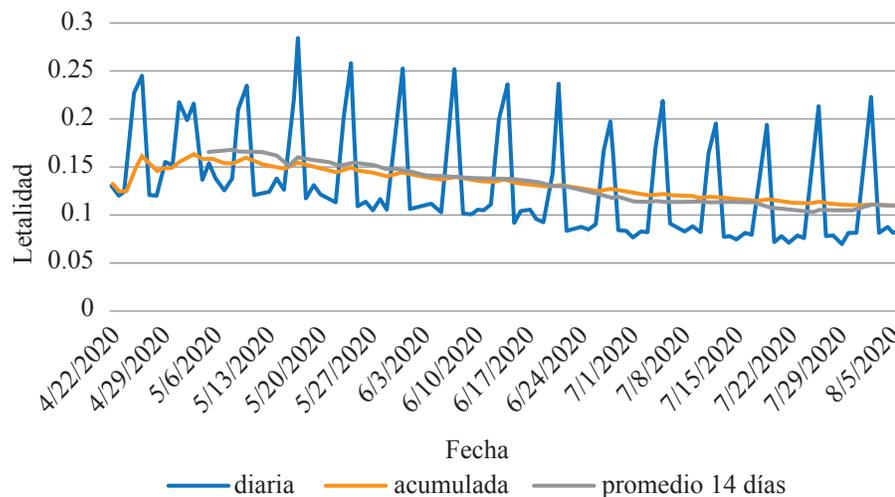
letalidad diaria de la cual se procedió al recálculo de la letalidad acumulada y el promedio de letalidad diaria cada 14 días, considerando como base el 22 de abril, a partir de las siguientes ecuaciones (Rothman *et al.*, 1998):

$$\text{Letalidad diaria} = \frac{\text{número defunciones en el día } n}{\text{número de personas infectadas en el día } n}$$

$$\text{Letalidad acumulada} = \frac{\text{número defunciones hasta el día } n}{\text{número de personas infectadas hasta el día } n}$$

$$\text{Promedio cada 14 días} = \frac{\text{número defunciones promedio desde el día } n_{-14} \text{ al día } n}{\text{promedio de personas infectadas desde el día } n_{-14} \text{ al día } n}$$

A partir del comportamiento observado de las tres variables, se consideró a la letalidad acumulada de este periodo como estadístico para el agrupamiento espacial, pues resume las tendencias observadas en las otras variables y agrega la información histórica de letalidad en un dato concreto (gráfica 1).



GRÁFICA 1

Letalidad diaria, acumulada y promedio de 14 días en México del 22 de abril al 6 de agosto

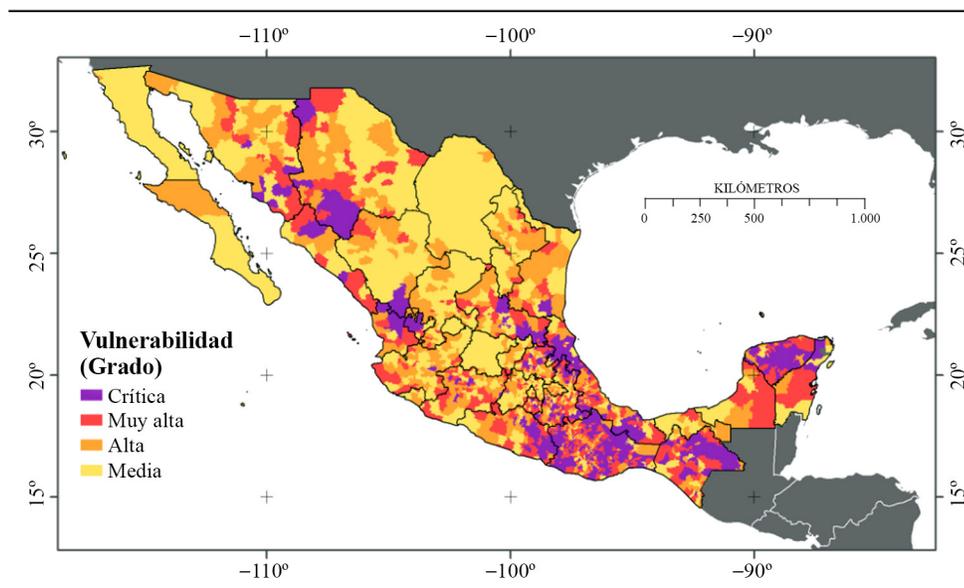
Fuente: elaboración propia a partir de datos de la Plataforma de información geográfica de la UNAM sobre COVID-19 en México (Ghilardi *et al.*, 2020).

1. 2. Índice de vulnerabilidad ante el COVID-19

En México, como parte del reconocimiento de los efectos diferenciados en cada lugar por COVID-19, así como por el entendimiento desde el punto de vista geográfico acerca de cómo éstos variarán de acuerdo con una serie de características sociales, económicas y de salud de la población, Suárez-Lastra *et al.*, (2020) elaboraron un índice de vulnerabilidad ante el COVID-19 en escala municipal que integra los principales aspectos demográficos, de salud y socioeconómicos de la población que influyen en su capacidad para anticipar, hacer frente, resistir y recuperarse de los efectos adversos del coronavirus, de acuerdo con la terminología de la Oficina de las Naciones Unidas para la Reducción del Riesgo de Desastre (UNISDR, 2017).

Cada una de las tres dimensiones que integran el índice de vulnerabilidad ante el COVID-19 está construida a partir de una serie de indicadores cuya relevancia se basa en la revisión de trabajos publicados en cada una de las áreas de conocimiento que, además, existen como información pública, aunado a que miden los aspectos considerados más críticos en cuanto a la vulnerabilidad a los efectos de la epidemia (Suárez-Lastra *et al.*, 2020). La dimensión demográfica la componen cinco variables asociadas con las características de la población que aumentan la vulnerabilidad al virus SARS-CoV-2. La dimensión de salud la conforman cinco variables asociadas al estado de salud de la población, a la infraestructura de salud y al personal médico disponible en los municipios respecto a su población. Finalmente, la dimensión socioeconómica está integrada por ocho variables relacionadas con el bienestar de la población de un municipio en términos de satisfactores básicos, derechos y capacidad económica. No obstante, para facilitar su interpretación y su representación cartográfica el índice se categorizó en cuatro grupos o clases (cuartiles) que representan cuatro grados de vulnerabilidad: medio, alto, muy alto y crítico (mapa 1).

De acuerdo con los resultados reportados por Suárez-Lastra *et al.* (2020), 63% de la población del país vive en municipios con un grado medio de vulnerabilidad, los cuales se encuentran asociados a grandes centros urbanos que cuentan con mayor infraestructura de salud y mayor capacidad económica. De la población nacional, 29% se ubica con valores altos y muy altos de vulnerabilidad; aquí se incluyen municipios con localidades urbanas de tamaño medio, infraestructura de salud escasa, producción económica precaria y acceso limitado a medios de comunicación. Finalmente, 8% de la población se ubica en valores críticos de vulnerabilidad, donde están presentes municipios rurales con el porcentaje más alto de personas mayores de 60 años y una importante población indígena, además de que dichos municipios cuentan con escasos servicios de salud y presentan tasas de marginación muy por arriba de la media nacional.



MAPA 1

Grado de vulnerabilidad al COVID-19 en México

Fuente: Suárez-Lastra *et al.* (2020).

1. 3. Índice de sustentabilidad de capital natural

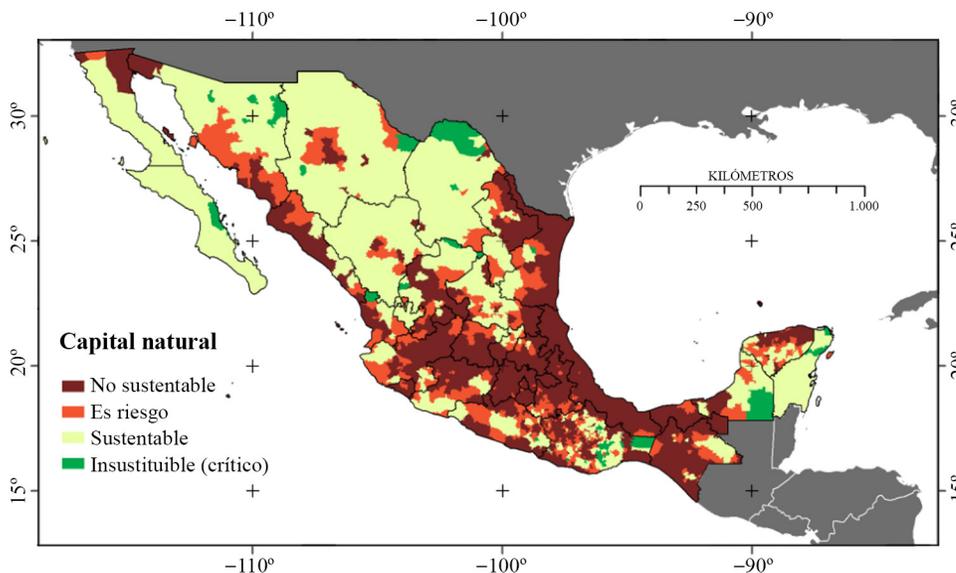
Con la aparición del virus SARS-CoV-2, en Wuhan, China, ha quedado claro que a medida que avanza la destrucción de la naturaleza aumenta el contacto entre los humanos y las especies potencialmente portadoras de infecciones. La deforestación, el tráfico ilegal de especies y la producción intensiva de alimentos son algunos

de los factores asociados al aumento de las enfermedades zoonóticas (Wallace, 2016). Los resultados de la actividad humana durante las últimas décadas se han traducido notablemente en la crisis de la biodiversidad y en la pérdida de la condición íntegra de los ecosistemas (Mora, 2017).

El capital natural se relaciona con la integridad ecológica cuando cualquier *stock* de recursos naturales o activos ambientales que proporcione un flujo de bienes o servicios útiles, ahora y en el futuro, se vuelve crítico para el uso sostenible (Brand, 2009). En este contexto, la integridad ecológica se puede utilizar como el factor más completo para evaluar la calidad restante del ecosistema después de los impactos humanos (Mora, 2019). El índice de sustentabilidad de capital natural es el producto del tamaño del ecosistema remanente, expresado por su cantidad, así como de su integridad ecológica, expresada como calidad (Czúcz *et al.*, 2012). Este indicador del estado y cambio en la biodiversidad de los ecosistemas naturales y ecosistemas agrícolas permite un análisis sobre la cantidad y calidad del capital natural que tiene México, el cual es importante no olvidar por la prisa para reconstruir economías devastadas.

A partir de diversas fuentes de información geográfica, se ha integrado un SIG de apoyo a la decisión espacial, con cobertura nacional, que utiliza modelos de representación tipo ráster con una resolución de 1 km²; éste permite estimar la cantidad de degradación del capital natural a nivel municipal. Lo anterior posibilita que se identifique el capital natural crítico insustituible para sustentar funciones ecológicas, que también es parte del capital natural sustentable, e indica las condiciones donde el ecosistema puede tener un legado evolutivo, es decir, mantener las propiedades de autorregulación, autoorganización y estabilización de procesos ecológicos como las interacciones depredador-presa. Además, el capital natural sustentable para las ecorregiones también se identifica cuando la condición del ecosistema no se ha degradado al grado de que las propiedades emergentes puedan mantener propiedades intactas, estables y autoorganizadas. El capital natural en riesgo se considera cuando el legado evolutivo ya no se sostiene y el capital natural no sustentable se identifica cuando la integridad ecológica se encuentra en un punto de máxima saturación de degradación ecológica.

En un nivel municipal, 5% del capital natural presenta condiciones críticas para su conservación, 31% es sostenible, 20% está en riesgo y 44% del capital natural se considera insostenible (mapa 2).



MAPA 2

Índice de sustentabilidad de capital natural por municipio para México

Fuente: Mora (2019).

1. 4. Accesibilidad a los establecimientos de salud

El índice de accesibilidad a los establecimientos de salud en México sintetiza el tiempo (minutos) de traslado esperado entre las localidades y los establecimientos de salud de primero, segundo y tercer orden en México. Lo anterior bajo el supuesto de que una persona puede acceder a cualquiera en el menor tiempo posible por medio de un transporte automotor o desplazándose a pie considerando aspectos como la velocidad máxima de circulación de la red vial o la velocidad a pie de acuerdo con las condiciones de la superficie como la pendiente y cobertura del suelo (Frakes *et al.*, 2015).

Para este índice se calcula el tiempo de viaje a partir de productos de información geoespacial disponibles tales como el Continuo de Elevaciones Mexicano 3.0 (INEGI, 2020b), la Red Nacional de Caminos (INEGI-SCT-IMT, 2017), el Conjunto de datos de uso de suelo y Vegetación, escala 1:250 000 Serie VI (INEGI, 2016) y el Conjunto de datos vectoriales de información topográfica digital, escala 1:250 000 (INEGI, 2012).

La herramienta empleada fue Travel Time Cost Surface (Frakes *et al.*, 2015), la cual calcula el tiempo de viaje de un conjunto de localizaciones puntuales (cualquier localidad) definidas a las ubicaciones de interés (*e. g.*, establecimientos de salud) bajo el supuesto de que una persona accede a éstas en el menor tiempo posible si emplea automóvil o si va a pie considerando aspectos como la velocidad máxima de circulación de la red vial y la velocidad a la que se puede caminar de acuerdo con las condiciones de la superficie como la presencia de cuerpos de agua, la pendiente y cobertura del suelo, ya que al presentarse zonas con densa vegetación se reduce la velocidad o al incrementar la pendiente la velocidad disminuye de manera no lineal, de acuerdo con la ecuación de Theobald (2010) y Tobler (1993) que se muestra a continuación:

$$\text{Tiempo en segundos} = 6 * \text{EXP}(-3.5 * \text{ABS}(\text{TAN} \left(\frac{\text{Pendiente en grados}}{57.29578} \right) + 0.05))$$

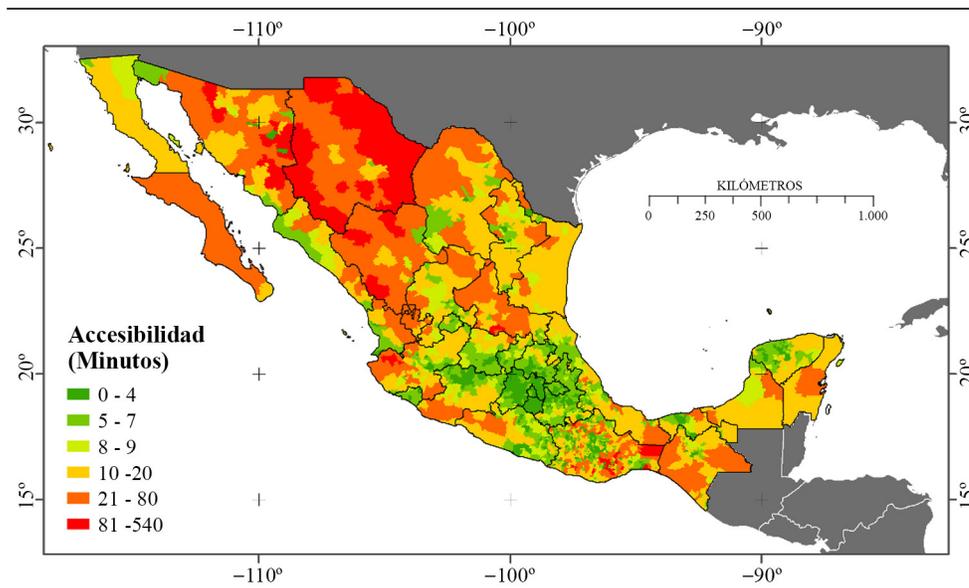
La parametrización de la velocidad en automóvil se definió en función de los límites especificados en la Red Nacional de Caminos (INEGI e IMT-SCT, 2017), mientras que para la velocidad a pie se consideró una velocidad promedio de 5.5 km/h. El cálculo de la fricción o nivel de impedancia de los usos de suelo se realizó con base en la serie de uso de suelo y vegetación VI (INEGI, 2016). Se reclasificaron las clases originales considerando las coberturas de suelo del National Land Cover Dataset (NLCD, 1992) de Estados Unidos. Como resultado se obtuvo una agrupación de ocho clases, donde a cada una de se le asignó el valor de su transitabilidad o porcentaje máximo de velocidad de viaje (PMTS). En el cuadro 1 se muestra el valor de transitabilidad en el que se calibraron los modelos.

CUADRO 1
Transitabilidad por tipo de uso de suelo y vegetación

Clases de uso de suelo y vegetación	Transitabilidad
Agricultura	0.80
Asentamientos humanos	0.95
Bosques	0.66
Cuerpos de agua	0.00
Matorrales	0.70
Otros	0.56
Pastizales	0.80
Selvas	0.68

Fuente: elaboración propia a partir la serie de uso de suelo y vegetación VI (INEGI, 2016) y del National Land Cover Dataset (NLCD, 1992).

Con base en la Clave Única de Establecimientos de Salud (CLUES) de la Dirección General Información en Salud, se procesó y estandarizó la base de datos a través de la depuración según el estatus de operación (en operación), nivel de atención (del primer al tercer nivel) y tipo del establecimiento (fijo), de acuerdo con la última actualización del 20 de junio de 2020. Cabe destacar que el nivel de atención se encuentra asociado con niveles de complejidad necesarios para resolver problemas de salud de diferente magnitud y severidad. Pese a que el COVID-19 generalmente no se atiende en el primer nivel, por su cantidad y distribución (~21 000 unidades), en localidades rurales distantes, se convierte en el primer punto de contacto y el medio encargado de canalizar a unidades médicas de nivel dos (~4 700 unidades) y tercero (~168 unidades). Para ilustrar, se presenta el índice de accesibilidad a establecimientos de salud en México como resultado de la integración para todos los niveles de atención (mapa 3).



MAPA 3

Índice de accesibilidad a establecimientos de salud en México

Fuente: elaboración propia a partir de Frakes *et al.* (2015).

1. 5. MÉTODO DE AGRUPAMIENTO ESPACIAL

El agrupamiento espacial es un método de aprendizaje automático empleado para resolver problemas de clasificación, predicción y agrupamiento de datos. Dado el número de grupos por crear, buscará una solución en la que todas las entidades dentro de cada uno sean lo más parecidas posible, ya que todos los grupos en sí son tan diferentes como es posible. La similitud se basa en el conjunto de atributos seleccionados para el análisis y también puede incorporar de manera opcional las propiedades espaciales o las propiedades espaciotemporales (Yan *et al.*, 2019).

La herramienta de análisis de agrupamiento proporciona valores R^2 que varían de 0 a 1 para cada atributo y, por lo tanto, se asocia el poder explicativo de la variable para discriminar entre las características (Moore y Dixon, 2015). Adicionalmente, el número de grupos creados juega otro papel importante en el análisis de agrupamiento espacial, ya sea porque se requiera crear *a priori* un número determinado de grupos o porque se precise evaluar el número óptimo de cada uno (Ruda, 2016). En cualquier caso, el empleo de la pseudoestadística F de Caliński-Harabasz (Caliński y Harabasz, 1974) permite expresar una relación que refleja la similitud dentro del grupo y las diferencias entre grupos. Los valores más grandes del pseudoestadístico F indican más similitud dentro del grupo y mayor diferencia entre ellos de acuerdo con la siguiente ecuación (Yan *et al.*, 2019):

$$\frac{(R^2/nc - 1)}{(1 - R^2)/(n - nc)}$$

Donde:

$$R^2 = \frac{SST - SSE}{SST}$$

Donde SST es empleado como una métrica de diferencias entre grupos y SSE como la similitud dentro del grupo, definidos respectivamente:

$$SST = \sum_{i=1}^{n_c} \sum_{j=1}^{n_i} \sum_{k=1}^{n_v} (V_{ij}^k - \bar{V}^k)^2$$

$$SSE = \sum_{i=1}^{n_c} \sum_{j=1}^{n_i} \sum_{k=1}^{n_v} (V_{ij}^k - \bar{V}_t^k)^2$$

Donde n es el número de elementos, n_i es el número de elementos en el grupo I , n_c es el número de clases (grupos), n_v es el número de variables utilizadas para agrupar elementos. V_{ij}^k es el valor de la k -ésima variable del j -ésimo elemento en el i -ésimo grupo, \bar{V}^k es el valor medio de la k -ésima variable, y \bar{V}_t^k es el valor medio de la k -ésima variable del grupo i .

Aunado al pseudoestadístico F de Caliński-Harabas, existen otras medidas para evaluar agrupaciones, dentro de las cuales destaca la relación de la suma de cuadrados entre las agrupaciones y la suma total de cuadrados, cuyo resultado se representa en una escala de medición de 0 a 1, y que sirve para interpretar la varianza total que es explicada por el modelo de agrupamiento (Everitt *et al.*, 2011).

Para comprender mejor el patrón de la letalidad por COVID-19 en México e identificar grupos de municipios espacialmente cercanos entre sí que comparten características similares, se empleó la herramienta de análisis de agrupamiento de ArcGIS Pro 2.6.1 (ESRI, 2020). Se utilizaron cuatro campos de atributos en la capa de 2 457 municipios para México, incluida la tasa de letalidad acumulada durante la Fase 3 por la epidemia de COVID-19, el índice de vulnerabilidad, el índice de sustentabilidad de capital natural y la accesibilidad a establecimientos de salud de primero, segundo y tercer nivel.

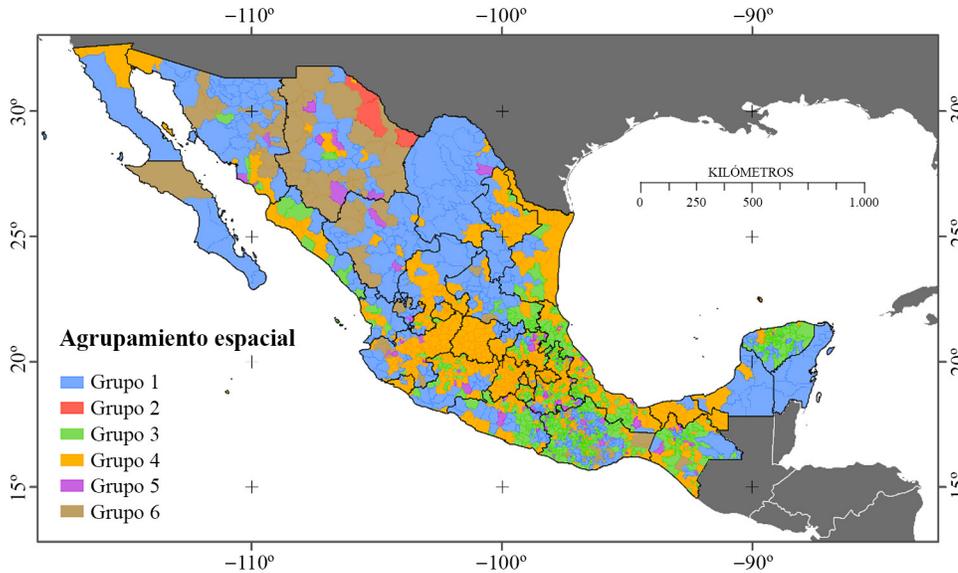
Para el método de agrupamiento, se empleó el algoritmo k -medoides donde, en lugar del valor medio de cada grupo, se toma como representativo un objeto en cada uno. Este objeto representativo, llamado *medoide*, está destinado a ser el que se encuentre ubicado más al centro dentro del grupo, por lo que después de determinar los k -medoides elegidos para los k grupos cada objeto restante se agrupa con el objeto representativo al que es más similar (Kaundinya *et al.*, 2013).

La eficacia del agrupamiento se midió mediante la pseudoestadística F para la que valores pseudoestadísticos F más grandes indican soluciones que funcionan mejor para maximizar tanto las similitudes dentro del grupo como las diferencias entre los grupos (Yan *et al.*, 2019).

2. RESULTADOS Y DISCUSIÓN

2. 1. El agrupamiento espacial de la letalidad por COVID-19 en México

De acuerdo con la metodología propuesta, se encontraron opciones similares en números de 4, 5 y 6 grupos. Sin embargo, al verificar los valores pseudoestadísticos F más grandes que indican mejores soluciones para maximizar tanto las similitudes dentro del grupo como las diferencias entre los grupos, 6 fue el número que se asoció con el mayor valor pseudoestadístico F para los 2 457 municipios de México (mapa 4).



MAPA 4
Agrupamiento espacial de la letalidad por COVID-19 en México
 Fuente: elaboración propia.

Este arreglo explica 64.7% de la varianza total de los datos, umbral que se interpreta similar al coeficiente de determinación ($r^2 = 0.6469$), que se encuentra dentro de la categoría de correlación positiva aceptable o media (Armitage y Berry, 1997). Rizkiana-Prima *et al.* (2020) aplican una técnica de agrupamiento, en el contexto del COVID-19, en la que reportan un valor de 84.4% de la varianza total de los datos, aunque cabe señalar que en ese estudio se realizaron tres agrupaciones a partir del empleo de cuatro variables altamente correlacionadas (número de casos, casos aliviados, muertes y la tasa de letalidad) para 35 provincias en Indonesia.

Comparando las variables analizadas, la variable de accesibilidad promedio a establecimientos de salud tiene el valor R^2 más alto (0.7713), lo que indica una influencia más efectiva en el proceso de agrupamiento, aunque los valores de R^2 de las otras variables como índice de sustentabilidad de capital natural (0.6785), letalidad (0.6079) e índice de vulnerabilidad (0.5409) muestran una influencia moderada (cuadro 2).

CUADRO 2
 Estadísticas de variables empleadas en el análisis de agrupamiento espacial

Variable	Unidad	Promedio	(SD)	Mínimo	Máximo	R^2
Accesibilidad a establecimientos de salud	Minutos	13.72	24.84	0.00	540.15	0.7713
Índice de sustentabilidad de capital natural	Sin unidad	0.21	0.20	0.00	0.91	0.6785
Letalidad acumulada	%	12.87	18.35	0.00	100.00	0.6079
Índice de vulnerabilidad ante el COVID-19	Sin unidad	3.04	0.43	1.85	4.22	0.5409

Fuente: elaboración propia.

2. 2. Caracterización de municipios en relación con la letalidad por COVID-19 en México

A partir del agrupamiento espacial de las cuatro variables propuestas en los seis grupos mejor evaluados, se describen cada uno de ellos considerando las principales características de la tasa de letalidad acumulada desde el inicio de la Fase 3, el índice de vulnerabilidad ante el COVID-19, el índice de sustentabilidad de capital natural y un indicador compuesto de accesibilidad a los establecimientos de salud. Los seis grupos de municipios en relación con las características que definen cada grupo y las diferencias entre ellos se describen en los siguientes apartados.

2. 2. 1. Grupo 1, municipios aislados

Se trata de 523 municipios que abarcan 45.3% del territorio nacional y agrupan 14% de la población. La tasa de letalidad promedio para el grupo es de 6.2%, caracterizada por una accesibilidad a establecimientos de salud de alrededor de 18 minutos, una vulnerabilidad de grado alto representado por municipios con localidades urbanas de tamaño medio, con una población con valores sociodemográficos y económicos ligeramente por debajo del promedio, pero que se caracterizan por un capital natural conservado.

2. 2. 2. Grupo 2, municipios extremadamente aislados

Este grupo está conformado por los municipios de Guadalupe, Coyame del Sotol y Manuel Benavides, que se ubican en el estado de Chihuahua. Se trata de tres municipios fronterizos que representan 1.1% de la superficie nacional, donde viven apenas 8 461 habitantes. Se trata de un grupo con letalidad de 0%, siete horas promedio de accesibilidad a establecimientos de salud, y un capital natural sustentable

2. 2. 3. Grupo 3, municipios precarios

En este grupo existen 749 municipios que representan 12% de la superficie y la población del país. Se caracteriza además por una tasa de letalidad acumulada de 9% y una accesibilidad promedio de nueve minutos a los establecimientos de salud; asimismo, se encuentran en un grado crítico de vulnerabilidad caracterizado por municipios primordialmente rurales con el porcentaje más alto de personas mayores de 60 años y una importante proporción de población indígena. En estos municipios se presenta el mayor nivel de hacinamiento, el menor acceso a medios de comunicación y una economía muy precaria, además de un capital natural caracterizado por una degradación ecológica que se encuentra en un punto de saturación máxima. Se trata de un conjunto de municipios donde las condiciones socioeconómicas y ecológicas son tan precarias que sobrellevar un periodo largo de contingencia significaría un mayor rezago social y ambiental

2. 2. 4. Grupo 4, municipios promedio

Se trata del grupo más numeroso conformado por 972 municipios que representan 25% de la superficie del país y 72% de la población nacional. Se caracteriza por un promedio de letalidad acumulada de alrededor de 12%, la máxima accesibilidad a establecimientos de salud y una vulnerabilidad de grado medio caracterizada por grandes centros urbanos que cuentan con la mayor infraestructura de salud y la mayor capacidad económica del país donde degradación ambiental se encuentra en un punto de saturación máxima. Se trata de municipios donde se ha presentado el mayor número de casos confirmados de COVID-19 en México.

2. 2. 5. Grupo 5, municipios críticos

Agrupar 119 municipios dispersos que representan 2.4% de la superficie del país y 0.9% de la población nacional, en donde se presenta un índice de vulnerabilidad en grado crítico y un índice de capital natural altamente en riesgo; la accesibilidad promedio a establecimientos de salud es similar al promedio nacional, pero donde la tasa de letalidad acumulada alcanza un promedio de 75.4%, lo que representa a los municipios en donde la pandemia ha tenido hasta ahora un mayor impacto por letalidad. Se trata de un conjunto de municipios dispersos principalmente en los estados de Puebla, Oaxaca, Veracruz y Chiapas, así como con menor presencia en Guerrero, Hidalgo, Jalisco, Chihuahua, Michoacán, Sonora, Durango y Coahuila.

2. 2. 6. Grupo 6, municipios muy aislados

Agrupar 91 municipios que representan 14.6% de la población nacional y 1.2% de la población del país. Se trata de un conjunto de municipios ubicados principalmente al noroeste del país con una letalidad de 7.5% y una accesibilidad promedio de 95 minutos a los establecimientos de salud. Al igual que los grupos 1 y 2, tiene características muy similares de vulnerabilidad y letalidad por COVID-19, un gradiente de poca accesibilidad e iguales valores de capital natural, por lo que estos tres grupos podrían formar parte de un mismo agrupamiento de municipios aislados caracterizados por un grado de marginación por arriba del promedio nacional con pequeñas áreas rurales, un porcentaje de población indígena por arriba del promedio nacional, infraestructura de salud muy escasa, cuyos hogares cuentan con pocos medios de comunicación y la producción económica es precaria (Suárez-Lastra *et al.*, 2020).

CONCLUSIONES

Este artículo expuso un análisis de agrupamiento espacial para comprender el patrón espacial de la letalidad acumulada en un nivel municipal del 22 de abril al 6 de agosto de 2020. Los resultados obtenidos permiten comprender las implicaciones demográficas, sociales, económicas, ambientales y de salud durante el inicio de la Fase 3 por la pandemia de COVID-19 en México.

Se identificaron seis agrupamientos que aglutinan espacialmente una serie de municipios que comparten características similares de la letalidad acumulada por COVID-19, la vulnerabilidad de la población ante los efectos adversos del coronavirus, la calidad de los ecosistemas naturales y ecosistemas agrícolas, así como de la accesibilidad a los servicios de salud.

Los resultados obtenidos permiten identificar múltiples regiones geográficas con características socioambientales similares que permiten diferenciar la letalidad por COVID-19 en México, donde es imperativo que la política pública atienda la emergencia y se encamine a evitar la propagación de contagios territorialmente diferenciada. Tres de los seis grupos identificados presentan condiciones de aislamiento y conservación de capital natural en entornos rurales caracterizados por grandes carencias socioeconómicas y una alta vulnerabilidad de su población frente al COVID-19. Sin embargo, cada uno de los agrupamientos espaciales identificados son una forma específica en la que el virus se desarrolla en el país, por lo que también son una oportunidad para la política pública de atender de manera diferenciada estos territorios y así evitar la propagación de contagio y el aumento de muertes.

El método empleado incorpora variables geográficas como demografía, acceso a la salud y capital natural para comprender la letalidad por COVID-19 en México. Se trata de un método de análisis espacial desarrollado en un ambiente de SIG para generar agrupamientos de datos en función de múltiples atributos, incluida su localización. Esta aproximación resalta la importancia de que, desde las ciencias de la salud, se considere la dimensión espacial para interpretar el fenómeno COVID-19, su dimensión espaciotemporal, su impacto geográfico en la toma de decisiones, la vida cotidiana y el entendimiento de la evolución de la enfermedad.

ANÁLISIS PROSPECTIVO

El reto que conlleva la investigación sobre las implicaciones del coronavirus SARS-CoV-2 demanda estudios interdisciplinarios entre ciencias de la salud y diversas disciplinas, incluidas las espaciales, de tal forma que se incluya la capacidad de realizar análisis espacial de diferente índole para comprender la dinámica espacio-temporal de COVID-19, así como para la modelación predictiva de cómo evoluciona la enfermedad. En este contexto, el análisis espacial es un marco analítico importante para identificar áreas de mayor riesgo que sirve para apoyar la implementación de medidas de control y estrategias de contención que apuesten a reducir de manera significativa las tasas de transmisión de la enfermedad por COVID-19 con reducciones en la intensidad y la velocidad

de contagio. Los resultados obtenidos ofrecen un entendimiento acerca del comportamiento geográfico de la letalidad por COVID-19, por lo que pueden emplearse para encaminar una serie de esfuerzos por contener geográficamente los contagios y así retrasar la propagación del virus en diferentes regiones del territorio nacional, lo cual proporcionaría así un tiempo valioso para activar una respuesta más exitosa.

AGRADECIMIENTOS

Este trabajo forma parte del proyecto “Construcción de modelos de agrupamiento espacial para el apoyo a la toma de decisiones sobre las políticas públicas de contención geográfica que puedan ser implementadas en México durante y después de la crisis por COVID-19” de la Convocatoria Ibero Frente al COVID-19 de la Universidad Iberoamericana Ciudad de México, para enfrentar los efectos de la pandemia en México. Se agradecen los fondos obtenidos para la investigación, así como los comentarios de los árbitros de la revista que mejoraron sustancialmente el contenido del artículo. Para finalizar, agradecemos a la Plataforma de información geográfica de la UNAM sobre COVID-19 en México por el preproceso personalizado de los datos abiertos para su descarga diaria.

REFERENCIAS

- Assunção, R. M., Neves, M. C., Câmara, G., & Da Costa Freitas, C. (2006). Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees. *International Journal of Geographical Information Science*, 20(7), 797-811.
- Armitage, P., & Berry, G. (1997). *Estadística para la investigación biomédica* (tercera edición). Madrid: Harcourt Brace.
- Boulos, M. N. K., & Geraghty, E. M. (2020). Geographical tracking and mapping of coronavirus disease COVID-19/severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) epidemic and associated events around the world: how 21st century GIS technologies are supporting the global fight against outbreaks and epidemics. *International Journal of Health Geographics*, 19(1), 8.
- Brand, F. (2009). Critical natural capital revisited: Ecological resilience and sustainable development. *Ecological economics*, 68(3), 605-612.
- Caliński, T., & Harabasz, J. (1974). A dendrite method for cluster analysis. *Communications in Statistics-theory and Methods*, 3(1), 1-27.
- CONEVAL (Consejo Nacional de Evaluación de la Política de Desarrollo Social) (2020). Visor Geoespacial de la Pobreza y la COVID-19 en los municipios de México. CONEVAL. Disponible en https://www.coneval.org.mx/Medicion/MP/Paginas/Visor_Geoespacial_Pobreza_COVID-19.aspx
- Czucz, B., Molnár, Z., Horváth, F., Nagy, G. G., Botta-Dukát, Z., & Török, K. (2012). Using the natural capital index framework as a scalable aggregation methodology for regional biodiversity indicators. *Journal for Nature Conservation*, 20(3), 144-152.
- De Figueiredo, A. M., Codina, A. D., De Figueiredo, D. C. M. M., Gil-García, E., & Kalache, A. (2020). Letalidad del COVID-19: ausencia de patrón epidemiológico. *Gaceta Sanitaria*.
- DGE (Dirección General de Epidemiología) (2020). *Datos abiertos*. Secretaría de Salud. Disponible en <https://www.gob.mx/salud/documentos/datos-abiertos-152127>
- Dong, E., Du, H., & Gardner, L. (2020). An interactive web-based dashboard to track COVID-19 in real time. *The Lancet infectious diseases*, 20(5), 533-534.

- ESRI (Environmental Systems Research Institute). (2020). *ArcGIS 2.6.1 for Desktop GIS application*. Redlands: ESRI Inc.
- Everitt, B., Landau, S., Leese M., & Stahl, D. (2011). *Cluster Analysis* (5th ed.). New York,: John Wiley.
- Frakes, B. T., Flowe, T., & Sherrill, K. R. (2015). National park service travel time cost surface model (TTCSM). *Natural Resource Report NPS/NRSS/NRR-2015/933*. Fort Collins: National Park Service.
- Franch-Pardo, I., Napoletano, B. M., Rosete-Verges, F., & Billa, L. (2020). Spatial analysis and GIS in the study of COVID-19. A review. *Science of The Total Environment*, 140033.
- Gao, P., Zhang, H., Wu, Z., & Wang, J. (2020). Visualising the expansion and spread of coronavirus disease 2019 by cartograms. *Environment and Planning A: Economy and Space*, 52(4), 698-701.
- García de León-Loza, A. (2020). Indicadores básicos y tendencias espacio-temporales en 20 países por mortalidad COVID-19. *Posición*, 3(1), 3-23
- Ghilardi, A., Ruiz-Mercado, I., Navarrete, A., Sturdivant, E., Larrazábal, A., Velasco, R., Gazcón Núñez, M. y Franch, I. (2020). *Plataforma de información geográfica de la UNAM sobre COVID-19 en México*. Centro de Investigaciones en Geografía Ambiental (CIGA), Laboratorio Nacional de Análisis y Síntesis Ecológica (LANASE), Escuela Nacional de Estudios Superiores, Unidad Mérida y Escuela Nacional de Estudios Superiores, Unidad Morelia. Disponible en <https://covid19.ciga.unam.mx>.
- INEGI (Instituto Nacional de Estadística y Geografía) (2012). *Conjunto de datos vectoriales de información topográfica digital*. Escala 1:250 000. Serie IV. México: Instituto Nacional de Estadística y Geografía.
- INEGI (Instituto Nacional de Estadística y Geografía) (2016). *Conjunto de datos vectoriales de uso de suelo y vegetación*. Escala 1:250 000. Serie VI (Capa Unión), escala: 1:250,000. México: Instituto Nacional de Estadística y Geografía.
- INEGI (Instituto Nacional de Estadística y Geografía), SCT (Secretaría de Comunicaciones y Transportes) e IMT (Instituto Mexicano del Transporte) (2017). Red Nacional de Caminos (RNC), versión 2018. Instituto Mexicano del Transporte.
- INEGI (Instituto Nacional de Estadística y Geografía) (2020a). *Visualizador analítico para el COVID-19*. Instituto Nacional de Estadística y Geografía. Disponible en <https://gaia.inegi.org.mx/covid19/>
- INEGI (Instituto Nacional de Estadística y Geografía) (2020b). Continuo de Elevaciones Mexicano 3.0. Disponible en <http://www.inegi.org.mx/geo/contenidos/datosrelieve/continental/descarga.aspx> (mayo de 2020).
- Kaundinya, D. P., Balachandra, P., Ravindranath, N. H., & Ashok, V. (2013). A GIS (Geographical Information System)-based spatial data mining approach for optimal location and capacity planning of distributed biomass power generation facilities: A case study of Tumkur district, India. *Energy*, 52, 77-88.
- Liu, Y., Gayle, A. A., Wilder-Smith, A., & Rocklöv, J. (2020). The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of Travel Medicine*, 27(2), <https://doi.org/10.1093/jtm/taaa021>
- Moore, T. W., & Dixon, R. W. (2015). A Spatiotemporal analysis and description of hurricane Ivan's (2004) tornado clusters. *Papers in Applied Geography*, 1(2), 192-196.
- Mora, F. (2017). A structural equation modeling approach for formalizing and evaluating ecological integrity in terrestrial ecosystems. *Ecological Informatics*, 41, 74-90.
- Mora, F. (2019). The use of ecological integrity indicators within the natural capital index framework: The ecological and economic value of the remnant natural capital of México. *Journal for Nature Conservation*, 47, 77-92.
- NLCD (National Land Cover Dataset). (1992). *National Land Cover Dataset*. The USGS Land Cover Institute (LCII). Disponible en <http://landcover.usgs.gov/natlcover.php>.

- Porta, M. (2014). A dictionary of epidemiology. *International Epidemiological Association* (6th ed. Oxford: Oxford University Press.
- Rizkiana-Prima, R., Yashintia Arien, E., & Sutikno Departemen Statistika, F. S. (2020). *Analisis Cluster Virus Corona (COVID-19) di Indonesia pada 2 Maret 2020-12 April 2020 dengan Metode K-Means Clustering*.
- Rothman K. J., Greenland S., & Lash T. (1998). *Types of Epidemiologic Studies. Modern Epidemiology* (2nd ed.). Philadelphia: Lippincott Raven.
- Ruda, A. (2016). Exploring tourism possibilities using GIS-based spatial association methods. *Geographia Technica*, 11(2).
- Simón-Domínguez, J. I., Simón-Domínguez, N., & Reyes-Núñez, M. A. (2020). Cómo estimar la letalidad del COVID-19. *Revista Mexicana de Patología Clínica y Medicina de Laboratorio*, 67(1), 4-8
- Suárez-Lastra, M., Valdés González, C., Galindo Pérez, M., Salvador Guzmán, L., Ruiz Rivera, N., Alcántara-Ayala, I., López Cervantes, M., Rosales Tapia, A., Lee Alardin, W., Benítez Pérez, H., Juárez Gutiérrez, M. del C., Bringas López, O., Oropeza Orozco, O., Peralta Higuera, A. y Garnica-Peña, R. (2020). Índice de vulnerabilidad ante el COVID-19 en México. *Investigaciones Geográficas*. <https://doi.org/10.14350/rig.60140>.
- Theobald, D. M. (2010). Estimating natural landscape changes from 1992 to 2030 in the conterminous US. *Landscape Ecology*, 25(7), 999-1011.
- Tobler, W. (1993). Non-isotropic geographic modeling. In W. Tobler (ed.), *Three presentations on geographic analysis and modeling*. Santa Barbara: National Center for Geographic Information and Analysis, University of California.
- UNISDR (United Nations International Strategy for Disaster Reduction/ Oficina de las Naciones Unidas para la Reducción del Riesgo de Desastre). (2017). *Report of the open-ended intergovernmental expert working group on indicators and terminology relating to disaster risk reduction*. Geneva: UNISDR
- Wallace, R. (2016). *Big farms make big flu: Dispatches on influenza, agribusiness, and the nature of science*. NYU Press.
- Yan, Y., Wang, D., Yue, S., & Qu, J. (2019). Trends in summer air temperature and vapor pressure and their impacts on thermal comfort in China. *Theoretical and Applied Climatology*, 138(3-4), 1445-1456.

CC BY-NC-ND