

Análisis geoespacial del COVID-19 en Ciudad de México y Estado de México

Catherine Montiel Porcayo, Carlos Alonso Medina Cortes,
Ana María Magdalena Saldaña Pérez

Instituto Politécnico Nacional,
Centro de Investigación en Computación,
México

{a190460, a190757}@sagitario.cic.ipn.mx, amsaldanap@ipn.mx

Resumen. El virus coronavirus (SARS-CoV-2), responsable de la enfermedad COVID-19 es hasta el día de hoy, la causante de una pandemia no vista desde la causada por la gripe española en 1918. Desde su aparición a finales de diciembre de 2019 en Wuhan China, se han reportado aproximadamente 22 675 000 casos en todo el mundo [1] y 543,806 casos en México [2]. Debido a la gravedad de esta enfermedad, se han llevado a cabo grandes esfuerzos desde diferentes frentes como el médico, el gubernamental, el social, y también el científico. Gracias al avance de la tecnología y a las herramientas computacionales con las que se cuenta, además de los análisis de casos de personas infectadas por el COVID-19, se pueden hacer aproximaciones del esparcimiento de contagios, conocer la ocupación hospitalaria, y generar estrategias para controlar el fenómeno. El área geoespacial no es la excepción y cobra amplia relevancia al proporcionar técnicas que permiten modelar el esparcimiento de los casos de personas infectadas en un área geográfica determinada, hacer predicciones del comportamiento de contagios, y conocer las relaciones geográficas que coadyuvan al esparcimiento del virus. En este trabajo se analiza desde una perspectiva geoespacial cómo es que ha evolucionado la epidemia en México. Aplicando técnicas y modelado geoespacial se realiza un análisis desde que surgió el primer caso en territorio mexicano hasta la tercera semana de agosto del 2020, las áreas geográficas son la ciudad de México y del estado de México, las dos urbes con más población por metro cuadrado en el territorio mexicano. En el artículo se presentan mapas generados por técnicas de geoprocésamiento, con datos obtenidos de fuentes oficiales del gobierno mexicano.

Palabras clave: COVID-19, coronavirus, pandemia, geoprocésamiento, análisis geoespacial.

Geospatial Analysis of COVID-19 in Mexico City and the State of Mexico

Abstract. Until today, the coronavirus virus (SARS-CoV-2), responsible for the COVID-19 disease, is the cause of a pandemic not seen since the one caused by

the Spanish flu in 1918. Since its appearance at the end of December 2019 in Wuhan China, approximately 22,675,000 cases have been reported worldwide [1] and 543,806 cases in Mexico [2]. Due to the seriousness of this disease, great efforts have been made from different fronts such as the medical, the government, the social, and also the scientific. Thanks to the advancement of technology and the computational tools that are available, in addition to the analysis of cases of people infected by COVID-19, it is possible to make approximations of the spread of infections, know the hospital occupation, and generate strategies to control the phenomenon. The geospatial area is no exception and is widely relevant by providing techniques that allow modeling the spread of cases of infected people in a given geographical area, making predictions of the behavior of contagions, and knowing the geographical relationships that contribute to the spread of the virus. This paper analyzes from a geospatial perspective how the epidemic has evolved in Mexico. Applying techniques and geospatial modeling, an analysis is carried out from the moment the first case arose in Mexican territory until the third week of August 2020, the geographical areas are Mexico City and the state of Mexico, the two cities with the most population per square meter. in the Mexican territory. The article presents maps generated by geoprocessing techniques, with data obtained from official sources of the Mexican government.

Keywords: COVID-19, coronavirus, pandemic, geoprocessing, geospatial analysis.

1. Introducción

En México el COVID-19 impactó drásticamente a la población, tomando la vida de miles de personas y cambiando la vida de toda la nación. A pesar de los esfuerzos del gobierno para reducir la tasa de contagios, en algunas localidades sobrepobladas fue imposible reducir el número de personas infectadas diariamente, muchas de ellas con comorbilidades relacionadas a la obesidad, problemas respiratorios, entre otros.

Otro aspecto importante que propició el avance del virus en la sociedad mexicana es la falta de compromiso para evitar el número de contagios, si bien las autoridades iniciaron varias campañas para promover las medidas de prevención como es el lavado de manos, el estornudo de etiqueta, evitar los saludos de mano, mantener la distancia entre personas y en lo posible trabajar desde casa no en todos los casos se siguieron adecuadamente [3], estas actividades requieren un cambio en el comportamiento de las personas y por lo tanto, un ritmo de aprendizaje para poder tomar las medidas de prevención con la seriedad necesaria.

En este trabajo se busca analizar el avance del virus a través de las dos grandes urbes de México, la Ciudad de México y del Estado de México, ambas presentan altos índices de contagios y defunciones, otro objetivo del trabajo es encontrar las zonas de dichas entidades federativas con los números de contagios más altos.

Ya que el avance del virus no es instantáneo es necesario realizar un estudio a lo largo del tiempo del comportamiento de su propagación, para lo cual se toma una ventana de tiempo que comprende desde la segunda semana de mayo hasta la tercera semana de agosto de 2020.

Dentro de los hallazgos que se pueden percibir en la presente investigación es el comportamiento espacial que presentan los datos, mostrado en visualizaciones espaciales; algunos de los análisis más importantes son la relación entre el área de las zonas de mayor concentración de casos con respecto al número de infectados, observándose que las dos grandes urbes fueron seriamente afectadas, al igual que sus periferias; es decir los focos de infección se concentraron en las ciudades y posteriormente el virus se esparcía hacia los municipios vecinos.

Este comportamiento se observa de manera puntual en la Ciudad de México y en sus municipios envolventes, llamados comúnmente zona metropolitana; de igual forma el fenómeno es fácilmente observable en la relación de contagios ocurridos en la ciudad de Toluca y sus municipios vecinos.

2. Estado del Arte

De acuerdo con la Organización Mundial de la Salud (OMS), la primera aparición del COVID-19 fue en Wuhan, China, a finales del 2019 y fue hasta marzo 11 del 2020 que OMS declaró pandemia a la enfermedad del COVID-19 [4]. Después de este evento varios países europeos reportaron un incremento en el número de casos a través del continente, resaltando a Italia, Francia y España. En México el primer caso de COVID-19 se reportó el 27 de febrero de 2020 por el titular de la Subsecretaría de Prevención y Promoción de la Salud, Hugo López-Gatell Ramírez [5]. El 19 de marzo, los 32 estados de la república mexicana reportaban casos positivos de COVID-19. Hasta el 23 agosto se reportaron 595,144 casos positivos de COVID-19, así como 62,241 defunciones a través de todo el territorio mexicano.

El análisis de una pandemia requiere que varias ramas de las ciencias trabajen de manera interdisciplinaria para estudiar y resolver la problemática, dentro de estas la geografía juega un papel clave, ya que se relacionan las actividades humanas con las biofísicas [6], esto nos permite estudiar cómo es que un virus se transporta a través de un espacio geográfico y el proceso que implica [7]. Si bien la geografía nos ayuda a conocer el desarrollo del virus, hay otras ramas como las ciencias de la salud que buscan encontrar métodos para sesgar el crecimiento del virus realizando análisis y pruebas clínicas con fármacos conocidos [8], así como soporte para realizar estudios en otras ramas como es la inteligencia artificial, rama en la cual se han realizado estudios para conocer el diagnóstico de pacientes para así de manera no invasiva saber si una persona es diagnosticada o no con COVID-19 [9, 10].

Los Sistemas de Información Geográfica (GIS) han resultado ser una herramienta fundamental en el análisis de la distribución de enfermedades a través del mundo, para así determinar su alcance, así como para la toma de medidas de prevención sanitarias para reducir el nivel de mortalidad. Desde el surgimiento del virus se han desarrollado varias aplicaciones utilizando GIS como son los paneles geográficos del Centro de Sistemas e Ingeniería de la universidad Johns Hopkins (JHU CSSE) [2], o el de la Organización Mundial de la Salud [11], para así tener una representación de los países con un mayor número de casos positivos por COVID-19, el número de defunciones o el número de personas recuperadas de la enfermedad.

La densidad de población de cada ciudad es particular, en un estudio por Gibson y Rush. los autores realizando un análisis geoespacial que consistía en la comparación

entre dos poblados usando un buffer de dos metros para simular el distanciamiento social, y con el uso del algoritmo de aprendizaje automático k vecinos cercanos (knn, k nearest neighbours) pudieron determinar qué tan cercanas están las personas una de las otras cuando ocurre un contagio, apoyando ideas de prevención de contagio como el uso de la sana distancia. En la ciudad de Masiphumelele, en Cabo Sudáfrica fue posible llevar a cabo las prácticas de distanciamiento social debido a que hay una menor densidad de población, sin embargo, en la ciudad de Klipfontein Glebe fue difícil mantener 2 metros entre persona y persona [12].

3. Metodología

En este apartado se describe la metodología por medio de la cual fue posible procesar, modelar, y analizar el comportamiento que el virus COVID-19 ha tenido sobre varias ciudades de México; en el presente artículo el área de estudio se limita específicamente a la Ciudad de México y al Estado de México. La metodología aplicada consta de cuatro etapas, a través de las cuales se lleva a cabo la recuperación y procesamiento de los datos, para su posterior análisis geoespacial.

La metodología propuesta cuenta con 2 características principales:

- Hace uso de datos provenientes de un repositorio de datos abiertos proporcionados por el gobierno mexicano, para obtener información sobre el esparcimiento del virus COVID-19 en la ciudad de México y en el Estado de México.
- Modela geoespacialmente los datos consultados y los resultados obtenidos del análisis del esparcimiento del virus en el área de interés.

Etapas 1.-Extracción de los datos: en esta etapa se obtuvieron los datos provenientes del programa de datos abiertos del gobierno de México, este corpus era actualizado todos los días por el mismo gobierno reportando todas las pruebas negativas y positivas que se realizan a pacientes que presentaban síntomas del virus COVID-19.

Etapas 2.- Tratamiento de datos: los datos son procesados para limpiarlos, filtrarlos por áreas deseadas (Estado de México y Ciudad de México), y obtener los atributos de interés, que para la presente investigación son los pacientes que han salido positivos a la prueba en cada división política de las áreas de interés.

Etapas 3.- Modelado espacial, en el que se analizan los datos obtenidos y se mapean en el área de estudio. Se aplican técnicas geoespaciales para permitir un mayor análisis de los datos procesados, y obtener el modelado geoespacial adecuado.

Etapas 4.- Análisis de las variables generadas a partir de los datos previamente procesados de manera geoespacial, para determinar patrones y características en el comportamiento que tiene el virus sobre el área de estudio.

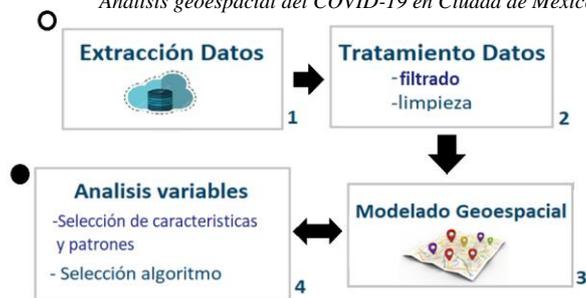


Fig. 1. Diagrama de actividades que muestra la secuencia de las etapas.

Tabla 1. Distribución de las variables a utilizar.

ETIDAD_UM	MUNICIPIO_RES	FECHA_DEF	RESULTADO
1	1	dd/mm/aaaa (fallecido)	1 (Positivo)
2	2	9999-99-99 (vivo)	2 (No Positivo)
3	3		3 (Pendiente)
...	...		
31	570		
32	999 (No especificado)		

La metodología planteada permite llevar a cabo el análisis espacio temporal del comportamiento del COVID-19 al procesar y modelar los datos en periodos de tiempo específicos planteados con base en la variación significativa del número de infectados. En la Figura 1, se muestra un diagrama de actividades de la secuencia de estas etapas, cabe destacar que la etapa 3 y 4 se retroalimentan una de la otra ya que algunos de los patrones y tendencias que muestra el virus en su esparcimiento se pueden visualizar de una mejor manera con un modelo geoespacial reiterativo.

La implementación de la metodología se llevó a cabo de la siguiente manera, primero se realizó una extracción continua de los datos de COVID-19 semanalmente, ya que en el trabajo realizado por el laboratorio en el transcurso de tres meses se mantuvo la actualización semanal de los mismos.

Como segundo paso e indagando en la etapa dos, se realizó un filtrado de los datos ya que se trabajaron tres áreas geográficas específicas, la República Mexicana dividida por estados, la Ciudad de México dividida por alcaldías, y el Estado de México dividido por municipios; posterior a esto, se realizó una limpieza de los datos ya que solo se trabajó con el número positivos de infectados por cada una de estas zonas y también con el número de fallecidos.

Para llegar a este filtrado se trabajó con las siguientes características dentro del corpus de datos abiertos de COVID-19 proveniente del gobierno: ENTIDAD_UM, que representa el código numérico de cada estado de la república, cabe destacar que cada tupla dentro del corpus representa una prueba de covid aplicada a un paciente con síntomas sospechosos de COVID-19, durante el desarrollo del proyecto se observó que el número de tuplas iba en aumento con el paso de los días empezando

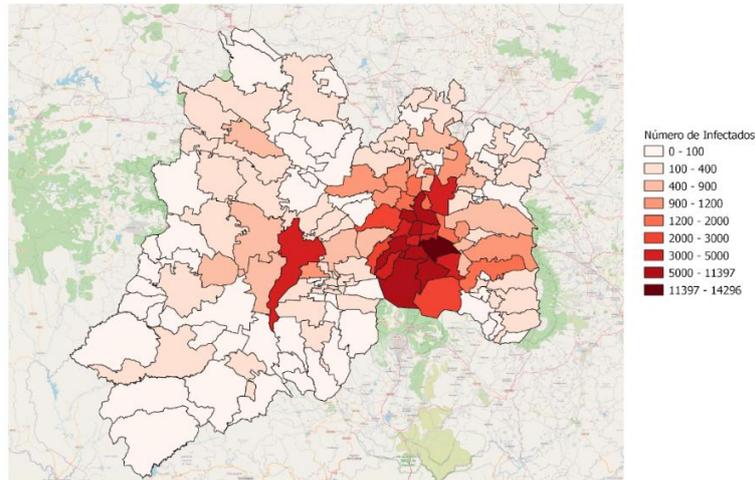


Fig. 2. Mapa de la ciudad de México y el Estado de México que muestra el número de infectados con fecha hasta el 20 agosto 2020.

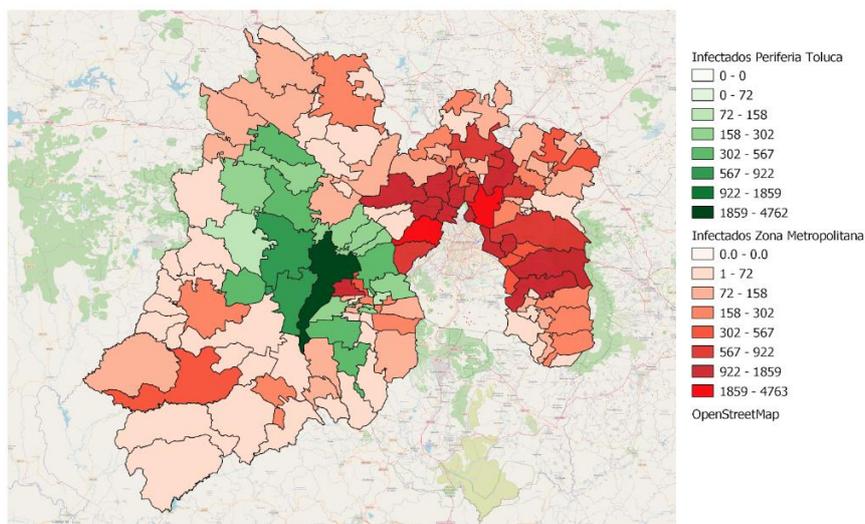


Fig. 3. Mapa del Estado de México que muestra el número de infectados por municipio hasta el 20 agosto de 2020.

con un corpus inicial de 110,995 tuplas con fecha del 6 Mayo del 2020, hasta un total de tuplas de 1,048,576 hasta el 20 de Agosto 2020.

Otra de las características a utilizar fue MUNICIPIO_RES que representa el código de municipio por cada estado de la república o el código de alcaldía para el caso de la ciudad de México, FECHA_DEF fue otra característica utilizada ya que esta nos indicaba si el paciente había muerto o no, y por último RESULTADO en donde sólo contabilizábamos los casos positivos, la siguiente tabla nos muestra las características

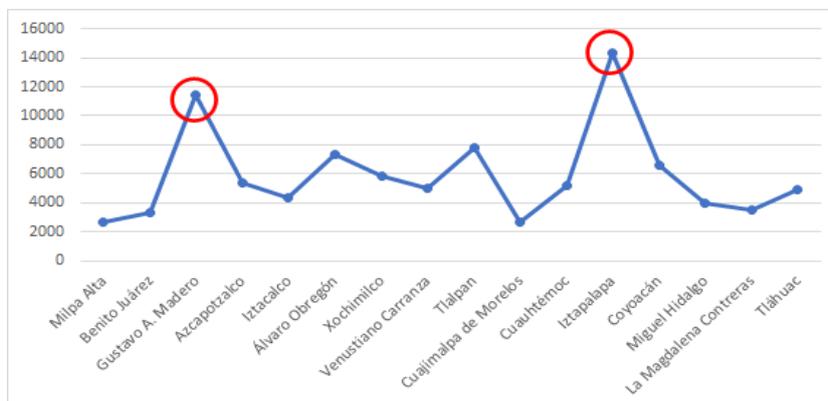


Fig. 4. Gráfica que muestra el número de pacientes positivos en la Ciudad de México en cada una de sus alcaldías.

de las variables que se utilizaron y qué valores podrían tomar cada una, los valores fueron procesados para ser representados como numéricos.

Posterior a la selección de características se realizó el conteo de datos, utilizando herramientas QGIS, consultas con PGADMIN (gestor de base de datos espacial para manejo de datos geoespaciales), y código en python.

En la tercera etapa se procedió a realizar el modelado Geoespacial en QGIS con periodicidad semanal, del número de infectados y el número de muertos en las tres zonas señaladas anteriormente.

Para el desarrollo del presente artículo, se decidió poner mayor interés en el Estado de México y en la Ciudad de México. La Figura 2 muestra un mapa de la ciudad de México y el Estado de México tomando como referencia el número de infectados con fecha última de 20 Agosto 2020, en el mapa se puede notar un principal foco de interés sobre la ciudad de México en donde algunas de sus alcaldías llegan hasta un pico arriba de 14,000 infectados, en contraste con los demás municipios del Estado de México donde los puntos más altos llegan aproximadamente a 4,700 infectados, uno de estos ejemplos es la ciudad de Toluca, esto se puede visualizar de manera rápida en el mapa de la Figura 2, basándonos en estas evidencias se decidió indagar en la hipótesis inicial, la expansión del virus tomando focos de infección para su posterior esparcimiento por sus municipios vecinos.

Al notar el mapa anterior y darnos cuenta que la diferencia de infectados en la Ciudad de México y el Estado de México era considerable decidimos crear un mapa aparte con coropletas guiándonos solo con los datos en el estado de México, con lo que llegamos al mapa de la Figura 3, de esta manera pudimos apreciar de manera espacial los principales focos de infección llegando a dos grupos y para fines explicativos decidimos mostrar de color verde para la zona de Toluca y roja para la zona metropolitana, en el mapa es claro apreciar estas zonas de contagio por su color más profundo.

Tomando en cuenta esta distinción de zonas infectadas y partiendo por los puntos de infección que en el mapa se muestran, las Ciudades de Toluca, Ecatepec y Naucalpan además de dos delegaciones en específico en la ciudad de México,

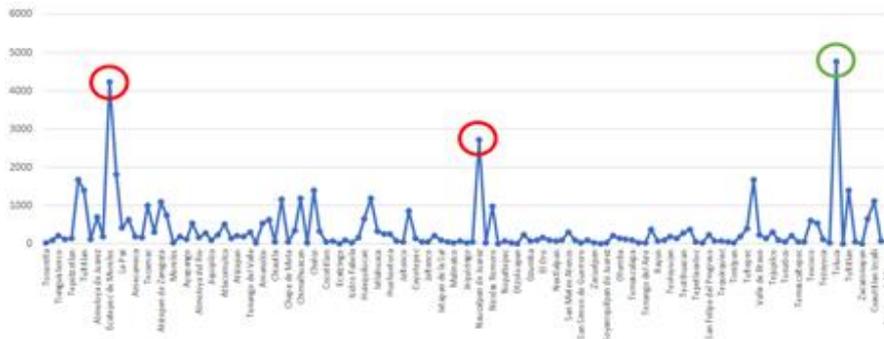


Fig. 5. Gráfica que muestra el número de pacientes positivos en el estado de México en cada uno de sus municipios.

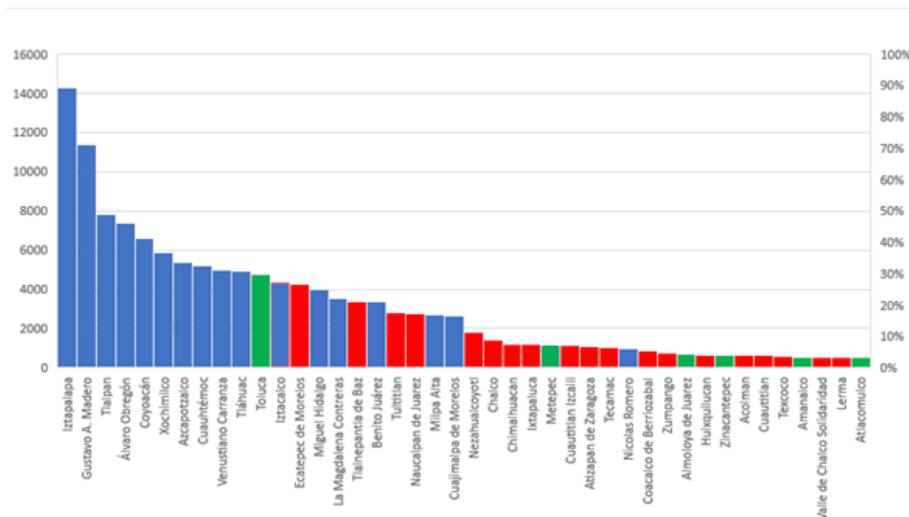


Fig. 6. Histograma que muestra el número de infectados ordenados de mayor a menor por municipio en el Estado de México y Ciudad de México.

Gustavo A Madero e Iztapalapa, decidimos corroborar los datos con gráficas e histogramas que nos mostraran los resultados esperados.

4. Discusiones

Analizando los datos de casos positivos en la Ciudad de México y el Estado de México se encontró que hay cierto patrón en el número de contagiados, agrupando el mayor número de contagiados en la alcaldía de Iztapalapa y Gustavo A. Madero en la Ciudad de México como se muestra en la figura 4, llegando a los 14,000 y 12,000 casos positivos respectivamente. Con lo que respecta al Estado de México, los

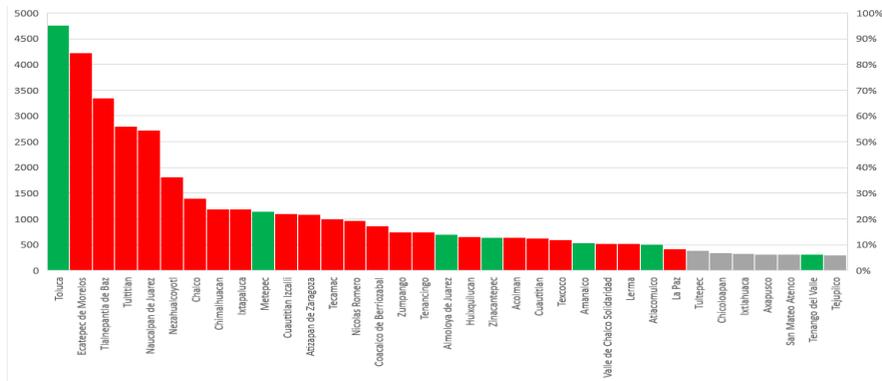


Fig. 7. Histograma que muestra el número de infectados ordenados de mayor a menor por municipio en el Estado de México, considerando solo los 35 municipios más afectados.

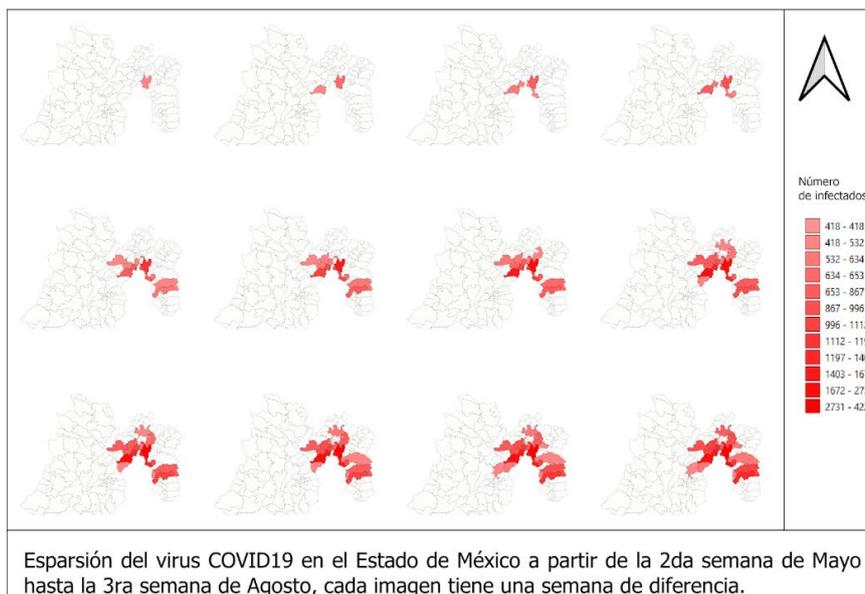


Fig. 8. Expansión del virus Covid 19 en el Estado de México con foco en la ciudad de México a partir de la 2da semana de mayo hasta 3ra semana de agosto, cada imagen tiene una semana de diferencia.

municipios de Toluca, Ecatepec y Naucalpan fueron los municipios con mayor número de casos positivos, como se muestra en la figura 5, alcanzando los 5,000,4,500 y 2,800 casos respectivamente. Las zonas analizadas representan los focos de infección y es a partir de estos que el número de casos positivos en las alcaldías y municipios colindantes fue aumentando de manera exponencial.

Se observa que los picos de infección se localizan en Iztapalapa y en la alcaldía Gustavo A. Madero, la gráfica muestra cómo las alcaldías próximas a las que

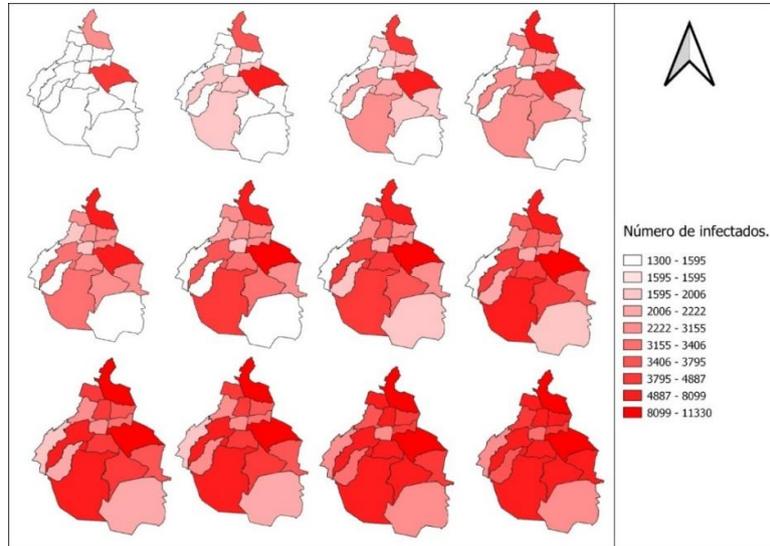


Fig. 9. Expansión del virus Covid 19 en la Ciudad de México a partir de la segunda semana de mayo hasta 3ra semana de agosto, cada imagen tiene una semana de diferencia.

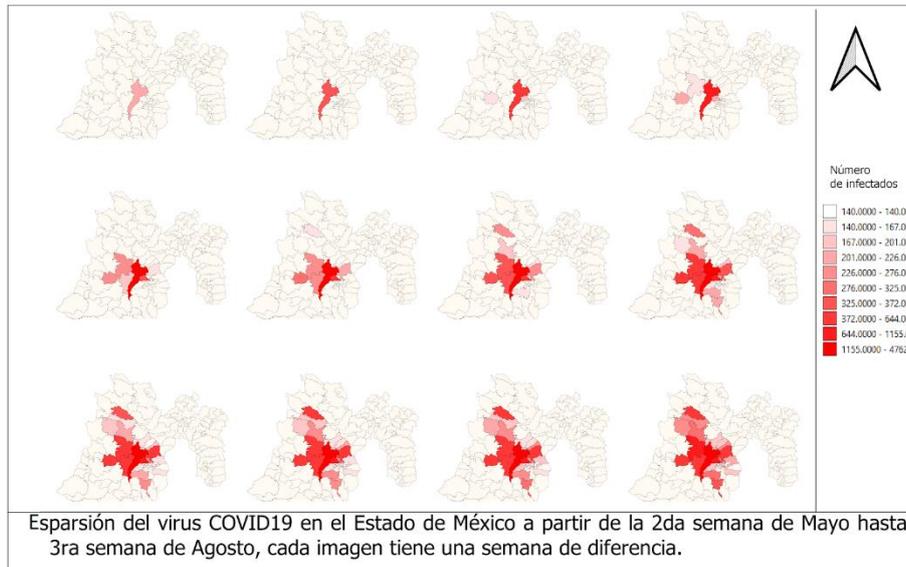


Fig. 10. Expansión del virus COVID-19 en el Estado de México con foco en la ciudad de Toluca a partir de la segunda semana de mayo hasta la tercera semana de agosto, cada imagen tiene una semana de diferencia.

presentan los mayores índices, también poseen un marcador elevado, a diferencia de las que se ubican más retiradas de los focos geográficos de infección.

Tomando en cuenta las gráficas anteriores se esperaba que los siguientes puntos en magnitud fuesen los municipios circundantes a los principales focos de infección, la figura 6 muestra un histograma ordenado de mayor a menor por número de infectados en las alcaldías de la ciudad de México (color azul) y los municipios del estado de México (color verde para la zona de Toluca, y color rojo para la zona metropolitana).

Basándonos en los resultados de la gráfica es posible percatarse de que las alcaldías de la Ciudad de México ocupan los primeros 10 lugares por arriba de Toluca y Ecatepec, analizando y esperando que la hipótesis planteada se cumpla en esta zona se esperaba que los dos primeros puestos fueran ocupados por Iztapalapa y la Gustavo A. Madero y que las alcaldías circundantes fuesen los siguientes puestos, viendo los resultados de la gráfica nos damos cuenta que se cumple la hipótesis los municipios en las posiciones 3 hasta 10 son las alcaldías que circundan a los principales focos, en la Figura 8 se puede ver este tipo de expansión en el mapa de manera semanal durante 3 meses.

Por otra parte, debido a que como en el mapa de la Figura 2 las alcaldías de la ciudad de México eclipsan los resultados de las dos zonas de Toluca y Zona metropolitana decidimos presentar otro histograma, solo tomando en cuenta los datos del Estado de México esperando que la hipótesis se cumpla también en estas, la Figura 7 muestra un histograma ordenado de mayor a menor por número de infectados en el estado de México rojo para la zona Metropolitana, Verde para la zona de circundante a Toluca y gris para los municipios restantes que no se encuentran involucrados en las zonas de esparcimiento.

Al analizar el comportamiento de la gráfica nos damos cuenta que las tres primeras posiciones corresponden a los principales 3 puntos que se mostraban en la gráfica 2, las siguientes barras muestran el comportamiento del virus por los municipios vecinos ya que podemos ver que las siguientes barras desde la posición 4 hasta la posición 29 corresponden a todos los municipios vecinos, también podemos notar que el segundo foco de infección más fuerte después de la Ciudad de México se encontró dentro de la zona metropolitana ya que la gráfica está pintada en su mayoría de color rojo.

Como tercer foco localizamos Toluca, es importante hacer notar que la diferencia de casos positivos a COVID-19 entre los municipios vecinos de Toluca y esta ciudad es grande, ello puede ser atribuido a la cantidad de hospitales que se encuentran en la ciudad a comparación de sus municipios vecinos; el conteo se hizo a partir de la característica **ENTIDAD_UM** que se refiere a la entidad donde se encuentra el hospital o la unidad de medicina que asistió al paciente.

A lo largo de los tres meses que duró este estudio se puede concluir que los focos de infección jugaron un importante lugar en la tasa de contagios en México, esto se puede visualizar en los mapas de coropletas de las Figuras 8, 9 y 10, cada uno representando las tres zonas antes mencionadas, donde se puede ver cómo es que a partir de las zonas con mayor número de contagios el virus se empieza a esparcir entre las demás vecindades, cambiando el color de blanco a los diferentes tonos de rojo, mostrando así la severidad de la situación y la expansión del virus en el territorio nacional.

Estos mapas muestran claramente los resultados ya antes discutidos en las gráficas analizadas, de tal manera que el lector pueda ver a través del tiempo la expansión del virus.

5. Conclusiones

El virus proveniente de Wuhan, China ha avanzado miles de kilómetros a través de la tierra debido a distintos factores. En México, el Covid se ha logrado contener en ciertos epicentros de las urbes. Gracias a este trabajo podemos localizar geográficamente los territorios más afectados, así como las zonas donde se reportan un pequeño número de infectados. El número de infectados varía de acuerdo con qué tan cerca están los poblados de los focos de la enfermedad.

La investigación proveniente de datos gubernamentales resulta muy útil para modelar el ritmo de infección del virus en cualquiera de las regiones del territorio nacional a través del tiempo. Con esta investigación se pudo modelar la situación actual del impacto de la epidemia del COVID-19 en el centro del territorio nacional, en trabajos futuros podrían analizarse los datos en todo el territorio mexicano para conocer qué estados o regiones fueron las más afectadas por el virus.

Los mapas generados semanalmente pueden ser consultados en el tablero del Laboratorio de Procesamiento Inteligente de Información Geoespacial. <https://blup.com.mx/piig/>. Como continuación de este trabajo de investigación se realizará un sitio web con un tablero en tiempo real del avance del COVID-19 a lo largo del territorio mexicano.

Referencias

1. Gobierno de México: Comunicado técnico 20 de Agosto (2020)
2. University Johns Hopkins University: Mapa global Johns Hopkins (2020)
3. Acuña-Zegarra, M.A., Santana-Cibrian, M., Velasco-Hernandez, J.X.: Modeling behavioral change and COVID-19 containment in Mexico: A trade-off between lockdown and compliance. *Mathematical Biosciences*, 325, pp. 108370 (2020)
4. Organización Mundial de la Salud: COVID-19: Cronología de la actuación de la OMS (2020)
5. Wikipedia: Pandemia de COVID-19 en México (2020)
6. Turner, B.L.: Contested identities: Human-environment geography and disciplinary implications in a restructuring academy. In: *Annals of the Association of American Geographers*, 92(1), pp. 52–74 (2002)
7. Franch-Pardo, I., Napoletano, B.M., Rosete-Verges, F., Billa, L.: Spatial analysis and GIS in the study of COVID-19. *Science of The Total Environment*, 140033 (2020)
8. Tobaiqy, M., Qashqary, M., Al-Dahery, S., Mujallad, A., Hershman, A.A., Kamal, M.A., Helmi, N.: Therapeutic management of COVID-19 patients: A systematic review. *Infection Prevention in Practice*, 100061 (2020)
9. Laguarda, J., Hueto, F., Subirana, B.: COVID-19 artificial intelligence diagnosis using only cough recordings. In: *IEEE Open Journal of Engineering in Medicine and Biology*, pp. 275–281 (2020)
10. Mei, X., Lee, H.C., Diao, K.Y., Huang, M., Lin, B., Liu, C., Bernheim, A.: Artificial intelligence-enabled rapid diagnosis of patients with COVID-19. *Nature Medicine*, 26, pp. 1–5 (2020)
11. Organización Mundial de la Salud: Who coronavirus disease dashboard (2020)
12. Gibson, L., Rush, D.: Novel coronavirus in Cape Town informal settlements: feasibility of using informal dwelling outlines to identify high risk areas for COVID-19 transmission from a social distancing perspective. *JMIR Public Health and Surveillance*, 6(2), pp. 9 (2020)