

Implementación del modelo matemático espacio-temporal para el COVID-19 en México

María Beatríz Bernábe Loranca¹,
Armando Benjamín Cruz Hinojosa²

¹ Benemérita Universidad Autónoma de Puebla,
Facultad de Ciencias de la Computación,
México

² Universidad Nacional Autónoma de México,
Facultad de Ciencias,
México

beatriz.bernabe@gmail.com,
aleph.g@ciencias.unam.mx

Resumen. En este trabajo se presenta la implementación computacional de un modelo matemático-epidemiológico que estima distintas variables sobre el problema del virus SARS-COV2 que produce la enfermedad COVID-19. El modelo considera tanto los datos estadísticos del INEGI como los datos epidemiológicos de la Dirección General de Epidemiología. La implementación del modelo consiste en un programa que genera estimaciones del número de infecciones, el número de pacientes hospitalizados en una Unidad de Cuidados Intensivos UCI, el número de muertes y la tasa de contagio. La importancia del desarrollo del modelo matemático consiste en una herramienta que estima los contagios de COVID-19 en distintas regiones de México.

Palabras clave: COVID, modelo matemático, predicción, python.

Mathematical Spatiotemporal Model Implementation for COVID-19 in Mexico

Abstract. This paper presents the computational implementation of a mathematical-epidemiological model that estimates different variables on the SARS-COV2 virus problem that produces COVID-19 disease. The model considers both statistical data from INEGI and epidemiological data from the General Directorate of Epidemiology. The implementation of the model consists of a program that generates estimates of the number of infections, the number of patients hospitalized in an Intensive Care Unit (ICU), the number of deaths and the infection rate. The implication of the development of the mathematical model consists of a tool for the visualization of COVID-19 infections in different regions of Mexico.

Keywords: COVID, mathematical model, prediction, python.

1. Introducción

Un año ha pasado desde que comenzó la emergencia sanitaria mundial ocasionada por el virus SARS- CoV-2 pero los contagios y muertes no han cedido. El reto social ha sido dramático, incidiendo en la calidad de vida y cambiando la cotidianidad de las personas alrededor del mundo. A pesar de los esfuerzos, se continúa amenazando la integridad de los hospitales y sistemas de salud de las naciones.

Para enfrentar la crisis, se requieren tomar decisiones y generar medidas de políticas públicas precisas fundamentadas en información actualizada y tecnológica que permita sistematizar la información que se genera de distintas fuentes de datos confiables. En este trabajo se presenta la implementación computacional de un modelo matemático para la propagación de la epidemia espacio-temporal del COVID-19, el cual se adaptó a la realidad mexicana usando información estadística del INEGI y los datos abiertos de la DGE (Dirección General de Epidemiología) sobre los casos de COVID-19 en México.

El software desarrollado produce estimaciones de número de contagios, número de hospitalizaciones, número de muertes y tasa de contagios. El modelo formulado en este artículo se ha inspirado de un modelo español y se adaptó a la situación Mexicana. El modelo que se presenta se implementó en el lenguaje de programación Python [1]. Los coronavirus son una familia de virus que causan infección en seres humanos y en animales como camellos, gatos y murciélagos.

Los coronavirus que afectan al ser humano (HCoV) pueden provocar desde el resfriado común hasta otros más graves como los producidos por los virus del Síndrome Respiratorio Agudo Grave SARS y del Síndrome Respiratorio de Oriente Próximo (MERS-CoV). La primera aparición del SARS en 2003, ocasionó más de 8.000 casos en 27 países con mortalidad del 10 %. Por otro lado, desde 2012 se han reportado 2499 casos de MERS-CoV en 27 países, principalmente en Arabia Saudita hasta con una letalidad de 34 % [2].

El virus que enfrentamos desde el año 2020, es un coronavirus en un sentido amplio. Sin embargo, dentro de la taxonomía de los coronavirus, el virus actual pertenece a la categoría Beta, su nombre oficial es SARS-CoV-2, el cual produce la enfermedad COVID-19 en humanos. El SARS-CoV surgió en 2002 en China aproximadamente después del HCoV-HKU1, sin embargo, el SARS-CoV generaba un síndrome respiratorio agudo en grupos de riesgo. Por su carácter nocivo, produjo un brote que infectó a más de 8.400 personas en 26 países entre Asia, Europa y América.

Se registraron 800 muertes traducidas en una letalidad del 9,6 %. La pandemia fue controlada año y medio después [3]. A partir de este problema de salud mundial, muchos son los desafíos en dar solución al aplanamiento de la curva de contagios, la cual obedece a que las medidas de restricción puedan ser modeladas en la matemática y la computación. Es en este punto donde se inserta el desarrollo del proyecto: Se describe un modelo matemático propuesto para mapear la evolución de la pandemia de 2020 a nivel nacional ante múltiples escenarios. El presente artículo se encuentra organizado en las siguientes secciones:

1. Introducción.
2. Modelo de esparcimiento epidémico.
3. Búsqueda y extracción de datos nacionales relacionados al COVID-19.

4. Implementación computacional de un modelo matemático flexible para predecir el comportamiento de la epidemia a nivel nacional, estatal y municipal en diferentes escenarios.
5. Hacia la interpretación de datos en un SIG.
6. Discusión de resultados y conclusiones.

2. Modelo de esparcimiento epidémico

Desde que la pandemia sorprendió a principios del 2020, distintos investigadores tanto de manera individual como en grupos y grandes empresas se dieron a la tarea de desarrollar propuestas diversas para dar respuesta al problema de salud del COVID-19. Respecto a estudios de cohorte retrospectivo y multicéntrico para adultos enfermos hospitalizados de COVID-19 y mayores a 18 años de edad en Wuhan-China, se analizaron los enfermos que habían sido dados de alta o habían fallecido antes del 31 de enero de 2020.

Se extrajeron datos demográficos, clínicos, de tratamiento y de laboratorio, incluidas muestras seriadas para la detección de ARN viral y se compararon entre supervivientes y no supervivientes para identificar los factores de riesgo asociados con la muerte intrahospitalaria. Para ello los autores utilizaron métodos de regresión logística univariable y multivariable [4].

Otros hallazgos metodológicos se ocupan de 3 prioridades iniciales para estimaciones de contagios y decesos categorizadas como: a) locales (la población, movilidad, morbilidad de la zona, cultura y economía), b) globales (impacto en otras zonas) y c) totales (casos confirmados, casos recuperados fallecimientos etc.). Para ello se utiliza el modelo SIR (Susceptible, Recuperada, Infectada) y técnicas multicriterio [5].

Recientemente una propuesta de simulador considera 13 variables extraídas de la secretaria de salud en México. El software simulador genera distintos escenarios para alertar a la población si las medidas de salud se relajan [6]. Por otra parte, los datos de la (Organización Mundial de la Salud) fueron procesados para la predicción del comportamiento de los contagios por COVID-19. Los datos adquiridos se sometieron a una estimación que utiliza una función logística para generar pronósticos de contagios con buena aproximación [5, 7].

En este apartado se describe el modelo de esparcimiento epidémico de COVID-19 descrito en el artículo de Arenas-Cota y Gómez-Gardeñes [1]. En tal manuscrito, los autores se basan en un estudio previo de evolución de pandemias en metapoblaciones estructuradas que incluyen agentes sujetos a patrones de movilidad recurrentes. Es decir, se describe matemáticamente que la dispersión de la enfermedad en una región es estimada en pequeños grupos geográficos, donde el resultado se apoya de la observación de interacción entre personas que se derivan de la movilidad entre los grupos.

La idea subyacente da lugar a introducir el concepto de meta-población, las cuales son creadas a partir de una población que se distribuye en una colección de grupos poblacionales bien definidos, de distinto número de individuos. Los individuos dentro de un grupo están mezclados y el patógeno puede ser transmitido entre cualesquiera agentes del grupo con la misma probabilidad.

El otro aspecto importante de la meta-población es la movilidad de los agentes, cada individuo tiene la posibilidad de migrar de su grupo a otro, esparciendo el patógeno a nivel del sistema. La movilidad de estos agentes se representa como una red en la que cada nodo es un grupo y flechas entre ellos representan la posibilidad de moverse de un grupo a otro. Esta información de importancia sustancial, se codifica en una matriz denominada matriz de movilidad, R .

Las distintas investigaciones epidemiológicas relevantes al COVID-19, sugieren que se agreguen propiedades a los grupos poblacionales para simular adecuadamente la forma en la que el patógeno se propaga. Una de estas consiste en separar dentro de cada grupo a los individuos en distintos estratos de edad como jóvenes, adultos y ancianos. Esta consideración es relevante dado que se ha observado que personas jóvenes reaccionan de formas distintas al virus respecto a las personas mayores.

Esta situación mal atendida podría implicar consecuencias catastróficas a la salud. La otra propiedad en los grupos es que cada individuo puede pertenecer a uno de ocho grupos posibles: susceptible (S), expuesto (E), asintomático (A), infectado (I), hospitalizado en una unidad de cuidados intensivos (H), muerto (D), y recuperado (R). La evolución de las fracciones de agentes en el estado $m \in \{S, E, A, I, H, D, R\}$ y estrato de edad $g \in \{1, 2, 3\}$ en el grupo $i \in \{1, \dots, N\}$, denotadas por $\rho_i^{m,g}(t)$ esta dada por:

$$\rho_i^{S,g}(t+1) = (1 - \Pi_i^g(t))\rho_i^{S,g}(t), \quad (1)$$

$$\rho_i^{E,g}(t+1) = \Pi_i^g(t)\rho_i^{S,g}(t) + (1 - \eta^g)\rho_i^{E,g}(t), \quad (2)$$

$$\rho_i^{A,g}(t+1) = \eta^g\rho_i^{E,g}(t) + (1 - \alpha^g)\rho_i^{A,g}(t), \quad (3)$$

$$\rho_i^{I,g}(t+1) = \alpha^g\rho_i^{A,g}(t) + (1 - \mu^g)\rho_i^{I,g}(t), \quad (4)$$

$$\rho_i^{H,g}(t+1) = \mu^g\gamma^g\rho_i^{I,g}(t) + \omega^g(1 - \psi^g)\rho_i^{H,g}(t) + (1 - \omega^g)(1 - \chi^g)\rho_i^{H,g}(t), \quad (5)$$

$$\rho_i^{D,g}(t+1) = \omega^g\psi^g\rho_i^{H,g}(t) + \rho_i^{D,g}(t), \quad (6)$$

$$\rho_i^{R,g}(t+1) = \mu^g(1 - \gamma^g)\rho_i^{I,g}(t) + (1 - \omega^g)\chi^g\rho_i^{H,g}(t) + \rho_i^{R,g}(t). \quad (7)$$

Estas ecuaciones representan la dinámica en tiempo discreto de los agentes (cada paso temporal representa un día), y están basadas en trabajos previos en dinámicas de pandemias con enfoque en cadenas de markov microscópicas. Los parámetros involucrados en estas ecuaciones están descritos en la Tabla 1, a excepción de $\Pi_i^g(t)$, que representa la probabilidad de que un agente susceptible sea contagiado por un agente expuesto o infectado, convirtiéndose en un agente expuesto. La expresión analítica de dicho valor es:

$$\Pi_i^g(t) = (1 - p^g)P_i^g(t) + p^g \sum_{j=1}^N R_{i,j}P_j^g(t), \quad (8)$$

donde $P_i^g(t)$ es la probabilidad que agentes de estrato de edad g se infecten del patógeno en el grupo poblacional i y p^g es la probabilidad de movilidad de los agentes en el estrato de edad g . Para determinar $P_i^g(t)$ se asume que los contactos en el grupo i incrementan monotónicamente según la densidad poblacional, como dicta la función $f(x) = 1 + (1 + e^{-\xi x})$, además se considera la matriz de contactos C cuyas entradas

$C_{g,h}$ definen la fracción de contactos que un individuo el el estrato de edad g tiene con individuos en el estrato de edad h . De esta forma la fórmula para calcular $P_i^g(t)$ se define como:

$$P_i^g(t) = 1 - \prod_{h=1}^3 \prod_{j=1}^N (1 - \beta_A)^{z^g k^g f\left(\frac{n_i^{eff}}{s_i}\right) C_{g,h}\left(\frac{n_{j \rightarrow i}^{A,h}(t)}{(n_h^{eff})}\right)} (1 - \beta_I)^{z^g k^g f\left(\frac{n_i^{eff}}{s_i}\right) C_{g,h}\left(\frac{n_{j \rightarrow i}^{I,h}(t)}{(n_h^{eff})}\right)}. \quad (9)$$

Los exponentes de las bases $(1 - \beta_A)$, $(1 - \beta_I)$ en la Ecuación 9 representan el número de contactos estimado de un agente en el grupo i en el estrato g con pacientes asintomáticos e infectados respectivamente. De esta forma el doble producto es la probabilidad que un agente no se infecte después de dichos contactos. Es importante explicar que el término $z^g k^g f(n_i^{eff}/s_i)$ es el número promedio de contactos de un agente, ya sea con agentes infectados o no infectados, y z^g es un factor normalizante del efecto de la función $f(x)$, este se calcula como:

$$z^g = \frac{3}{\sum_{i=1}^N f\left(\frac{n_i^{eff}}{s_i}\right) (n_i^g)^{eff}}. \quad (10)$$

Y las poblaciones efectivas representan la cantidad esperada de agentes activos en el grupo poblacional i , tomando en consideración las migraciones:

$$n_i^{eff} = \sum_{g=1}^3 (n_i^g)^{eff}, \quad (11)$$

$$(n_i^g)^{eff} = p_g \sum_{j=1}^N R_{j,i} n_j^g + (1 - p_g) n_i^g. \quad (12)$$

Finalmente, los últimos miembros de los exponente en la Ecuación 9 son $n_{j \rightarrow i}^{A,h}(t)$ y $n_{j \rightarrow i}^{I,h}(t)$, los cuales representan el número de agentes del estrato de edad h en estado asintomático o infectado y que han migrado de la región j a la región i . Y pueden ser calculados de la siguiente forma:

$$n_{j \rightarrow i}^{m,h}(t) = n_j^h \rho_j^{m,h}(t) [(1 - p^h) \delta(i, j) + p^h R_{j,i}]. \quad (13)$$

Con $\delta(i, j)$ la función delta de Kronecker. El artículo original señala en su formulación políticas de aislamiento mediante un porcentaje k_0 de población en aislamiento y se han omitido los detalles, sin embargo se recomiendan las lecturas [1, 4, 5]. El modelo formulado aquí puede ser aplicado a la dinámica espacio temporal de otros fenómenos, como la transmisión de información y congestión del tráfico.

3. Búsqueda y extracción de datos

En la Sección 2 se menciona que una de las bondades del modelo reside en la posibilidad de modificar las variables para ajustarlo a la realidad de una población particular a tratar, definiendo una macro-población adecuada.

Tabla 1. Parámetros del modelo.

Símbolo	Significado	Valor estimado en España
β_A	Infectividad de pacientes asintomáticos	0,06
β_I	Infectividad de pacientes infectados	0,06
$\langle k^g \rangle$	Número promedio de contactos	(11,8, 13,3, 6,6)
η^g	Tasa de latencia	$\frac{1}{2,34}$
α^g	Tasa de infección asintomática	$\left(\frac{1}{5,06}, \frac{1}{2,86}, \frac{1}{2,86} \right)$
μ^g	Tasa de salida	$\left(\frac{1}{1,0}, \frac{1}{3,2}, \frac{1}{3,2} \right)$
γ^g	Porcentaje de casos necesitados de UCI	(0,002, 0,05, 0,36)
ω^g	Tasa de fatalidad en UCI	0,42
ψ^g	Tasa de mortalidad	$\frac{1}{7,0}$
χ^g	Tasa de descarga de UCIs	$\frac{1}{10,0}$
C	Matriz de contactos	$\begin{pmatrix} 0,5980 & 0,3849 & 0,01711 \\ 0,2440 & 0,7210 & 0,0350 \\ 0,1919 & 0,5705 & 0,2376 \end{pmatrix}$
ξ	Factor de densidad	0,01
$\langle p^g \rangle$	Factor de movilidad	(0,0, 1,0, 0,0)
σ	Habitantes promedio por hogar	2,5
k_0	Porcentaje de confinamiento	Ajustable

La macro-población para México consiste en tomar como grupo poblacional a cada uno de los municipios de la República. Esta conceptualización se justifica porque las interacciones económicas y sociales en cada municipio obligan a los individuos a estar en constante contacto, cumpliendo la hipótesis de ser un grupo bien mezclado donde el patógeno puede ser transmitido en la misma probabilidad. En esta sección se describen los métodos de obtención de datos necesarios para esta macro-población.

3.1. Datos pandémicos de México

Los datos corresponden a los casos de COVID-19 en México, los cuales se han descargado de la página de datos abiertos del gobierno mexicano [7,8,9]. En este sitio también se consigue información proveniente de la DGE sobre las pruebas de COVID-19 realizadas a nivel nacional. La información sobre COVID-19 respecto a los pacientes, contiene la siguiente información:

- Entidad y municipio de residencia del paciente.
- Edad del paciente.
- Resultado de la prueba (negativo/positivo).

- Tipo de paciente (ambulatorio/hospitalizado).
- Si el paciente requirió de cuidados intensivos (sí/no).
- Si el paciente fue intubado (sí/no).
- Fecha de inicio de síntomas en el paciente.
- Fecha de ingreso a hospitalización (Si es que ingresó).
- Fecha de defunción (Si es que el paciente falleció).

De acuerdo con las características de estos datos, una muestra representativa del problema es tomada para el ejercicio del modelo y los datos se consideran como epidemiológicos para el modelo matemático formulado en este trabajo.

3.2. Información demográfica

La información demográfica que nutre al modelo: a) La población total presente en la región. b) La densidad poblacional de la región. c) El número promedio de habitantes por hogar. Estos parámetros se valoran a nivel nacional, estatal y municipal tal que las cifras son obtenidas del INEGI (Instituto Nacional de Estadística y Geografía). En este trabajo se utilizaron datos recientes que el INEGI reserva para el público y corresponden al censo nacional realizado en 2015 [10, 11].

3.3. Información geográfica

Los datos corresponden a la información topográfica de los municipios, en particular los polígonos que describen la aproximación territorial de estos se obtuvieron de la plataforma CARTO [12]. El formato de esta información es geoJSON y con esta pueden graficarse regiones en un mapa. Cada municipio y estado de la república tiene su propio geoJSON, y a través del módulo de manejo de datos JSON de Python pueden ser manipulados. La exposición de los mapas con los resultados que arroja el modelo se encuentra en desarrollo.

4. Implementación computacional del modelo

En esta sección se describen aspectos importantes de la implementación que se codificó en Python y adicionalmente se describen los cálculos de las variables claves que han sido modificadas a partir de la información demográfica de México. El resultado alcanzado es un software capaz de simular escenarios futuros de la pandemia en México a escala municipal. Se explican las variables manteniendo la notación del modelo de original, descritas en la Tabla 1.

4.1. Definición de variables del modelo matemático

Tres estratos de edad son necesarios: jóvenes (0 a 25 años), adultos (26 a 65 años) y personas mayores (más de 65 años). Los estados en los que cada agente en un grupo puede encontrarse se denominan como: susceptible (S), expuesto (E), asintomático

(A), infectado (I), hospitalizado en unidad de cuidados intensivos (H), recuperado (R) y muerto (D). La matriz de contactos C , así como el número promedio de contactos de un agente por estrato de edad k , se reutilizaron del modelo español [1]. Se justifica tal situación dadas las similitudes culturales entre ambas naciones. Sin embargo, el cálculo de estos valores pueden adaptarse de mejor forma e incluso añadir el estado *vacunados*, lo cual se resolverá en un trabajo futuro.

De acuerdo con los datos obtenidos en la Sección 3.2, se estimó el número promedio de habitantes en cada hogar $\sigma = 3,7$. La lógica para el cálculo de la matriz de movilidad de regiones R se resume en la siguiente premisa: un agente tendrá mayor probabilidad de migrar a un municipio cercano al que pertenece que de migrar a un municipio alejado.

Este planteamiento se ha inspirado en leyes físicas de fenómenos naturales, en particular en la ecuación de Coulomb, por lo que se asume que tal relación es inversamente proporcional al cuadrado de la distancia entre los municipios. Usando los datos recopilados en la Sección 3.3, se toma el polígono de aproximación territorial de cada municipio y se calcula el centroide geométrico del municipio para almacenarlo en una estructura llamada tabla de municipios. La matriz de movilidad se calcula como:

$$R_{i,j} = \frac{f(d(i,j))}{\sum_{k \neq i} f(d(i,k))}, i \neq j; \quad R_{i,i} = 0, \quad (14)$$

donde $f(x) = 1/x^2$, y $d(i,j)$ es la distancia euclidiana entre los centroides del i -ésimo municipio y el j -ésimo municipio de la tabla de municipios. Cabe señalar que $f(x)$ puede ser sustituida por otra función monótona decreciente adecuada y $d(i,j)$ por otra métrica que mejor describa la distancia entre municipios, por ejemplo la generada por la red de carreteras de México.

Finalmente, para el cálculo de la distribución inicial de la población $\rho(t_0)$, por municipio, por estrato de edad y por cada uno de los siete estados, se procesa la información epidemiológica descrita en la Sección 3.1. A cada registro de los datos se le asigna el estrato de edad al que pertenece y se clasifica en uno de los estados posibles de la siguiente forma:

- Si el resultado de la prueba fue positivo, el paciente vive y han pasado más de 14 días desde la fecha de síntomas, se clasifica como paciente recuperado.
- Si el resultado de la prueba fue positivo, el paciente vive y han pasado menos de 14 días desde la fecha de síntomas, se clasifica como paciente portador.
- Si el resultado de la prueba fue positivo, el paciente vive y el paciente está hospitalizado en una Unidad de Cuidados Intensivos (UCI) o está intubado, se clasifica como paciente hospitalizado en UCI.
- Si el resultado de la prueba fue positivo, y el paciente se encontraba intubado o en una UCI, y tiene asociada una fecha de defunción, se clasifica como paciente difunto.

Posteriormente, con la información de la población de cada municipio por estratos de edad, se calcula el porcentaje de población en cada uno de los estados. El resto de la población supone se encuentra en estado susceptible.

El resto de las variables como el ritmo de mortandad, infectividad de pacientes asintomáticos e infectados, entre otros se reutilizan de la Tabla 1, por ser propiedades del patógeno y no de la región.

4.2. Implementación en Python

El modelo matemático ha sido implementado en un módulo de Python llamado Mappand, apoyándose de la librería de cómputo Numpy. La librería se compone de la clase MMC (Micro-Markov Chain) que implementa las fórmulas expuestas en la Sección 2, las cuales son necesarias para evolucionar el esparcimiento del patógeno en el tiempo y la clase Modelo que se encarga de encapsular los métodos necesarios para leer los datos de entrada y almacenar los resultados en una base de datos. La lectura puede ser desde una base de datos o de archivos CSV. Las entradas de la clase MMC son:

- El número de grupos poblacionales de la macropoblación, N .
- La matriz de contactos C , arreglo de numpy con forma $(3, 3)$.
- La matriz de movilidad R , arreglo de numpy con forma (N, N) .
- La matriz población por municipio y estrato de edad n , arreglo de numpy con forma $(N, 3)$.
- La densidad de población por municipio s , arreglo de numpy con forma $(N, 1)$.
- La distribución de población inicial $\rho(t_0)$, un diccionario de Python, cuyas llaves son los 8 distintos estados $\{S, E, A, I, H, D, R\}$ y cuyos valores son arreglos de numpy con forma $(N, 3)$.
- El diccionario de parámetros del modelo, que guarda la información de las variables descritas en la Tabla 1.

Una vez instanciado un objeto con estos datos, el método `evolve` se encarga de evolucionar en el tiempo el esparcimiento del patógeno en la macro-población. Las entradas al método son: el número de días a evolucionar en el tiempo Δt y el porcentaje de aislamiento social estimado k_0 . Esta evolución se realiza de la siguiente forma:

1. Se actualizan los parámetros $\{n_i^{eff}, z^g, n_{j \rightarrow i}^{A,g}(t), n_{j \rightarrow i}^{I,g}(t), P_i^g(t), \Pi_i^g(t)\}$ al tiempo actual t , como indican las Ecuaciones 8 a 13.
2. Se calcula la evolución de la población $\rho(t + 1)$ según las fórmulas descritas de la Ecuación 1 a la Ecuación 7.
3. Se guarda la nueva población $\rho(t + 1)$ en un arreglo y se repite el procedimiento tomando $\rho(t + 1)$ como nueva población a evolucionar.

El algoritmo se repite Δt veces, y el valor de retorno es el arreglo con las poblaciones estimadas. Las fórmulas que implementa la clase MMC hacen uso de técnicas de vectorización de la librería numpy. MMC también es generalizable cambiando las entradas del programa por variables particulares de una meta población arbitraria.

Tabla 2. Simbología.

Notación	Significado
CE	Total de casos estimados
CA	Total de casos actuales
ΔC	Nuevos casos estimados
HE	Total de hospitalizaciones estimadas
HA	Total de hospitalizaciones actuales
ΔH	Nuevas hospitalizaciones estimadas
ME	Total de muertes estimadas
MA	Total de muertes actuales
ΔM	Nuevas muertes estimadas

5. Hacia la interpretación de datos en un SIG

La parte final del proyecto consiste en la interpretación de los resultados del modelo matemático. Una vez se ha calculado la distribución de la población hipotética por cada uno de los municipios, se calculan los valores:

$$\text{Casos}_i(t_0 + \Delta t) = \sum_{g=1}^3 \left(\rho_i^{R,g}(t_0 + \Delta t) + \rho_i^{H,g}(t_0 + \Delta t) + \rho_i^{D,g}(t_0 + \Delta t) \right) n_i^g. \quad (15)$$

Esta suma encapsula los casos estimados por la vigilancia centinela pasados Δt días, suponiendo un porcentaje de aislamiento nacional k_0 :

$$\text{Hosp}_i(t_0 + \Delta t) = \sum_{g=1}^3 \rho_i^{H,g}(t_0 + \Delta t) n_i^g. \quad (16)$$

Esta ecuación reúne el número de pacientes hospitalizados en una UCI pasados Δt días, suponiendo un porcentaje de aislamiento nacional k_0 :

$$\text{Muertes}_i(t_0 + \Delta t) = \sum_{g=1}^3 \rho_i^{D,g}(t_0 + \Delta t) n_i^g. \quad (17)$$

Aquí se concentra el número de muertes acumuladas, pasados Δt días, asumiendo un porcentaje de aislamiento nacional k_0 . En trabajo futuro mediante el módulo de manejo de datos JSON de Python, se añadirán estos resultados al archivo geoJSON de cada entidad y municipio como nuevas propiedades llamadas casos, hospitalizados y defunciones respectivamente.

Que posteriormente con la librería JavaScript de código abierto, Leaflet, se encargará de mostrar los datos geoJSON de los municipios y entidades aplicando una escala de colores pertinente en base a la propiedad de casos.

Tabla 3. Evolución: $\Delta t = 7, k_0 = 0,5$. El cuarto renglón indica que Baja California tiene 5023 casos confirmados y el modelo predice 5269 casos dentro de una semana si el 50 % de la población está en confinamiento.

Entidad	CE	CA	ΔC	HE	HA	ΔH	ME	MA	ΔM
República Mexicana	95104	87347	7757	814	312	502	10012	9767	245
Aguascalientes	912	815	97	7	2	5	36	34	2
Baja California	5269	5023	246	39	7	32	873	861	12
Baja California Sur	625	605	20	2	0	2	36	36	0
Campeche	595	563	32	9	12	-3	72	68	4
Coahuila	1174	1084	90	8	1	7	79	77	2
Colima	157	141	16	2	2	0	20	19	1
Chiapas	1982	1705	277	33	21	12	148	137	11
Chihuahua	1631	1569	62	13	16	-3	320	314	6
Ciudad de México	26181	24263	1918	223	42	181	2273	2203	70
Durango	416	365	51	6	3	3	38	37	1
Guanajuato	1712	1463	249	19	10	9	115	108	7
Guerrero	1966	1827	139	19	9	10	276	270	6
Hidalgo	1852	1670	182	16	5	11	286	284	2
Jalisco	1969	1660	309	31	17	14	148	136	12
Edo. De México	15698	14264	1434	107	31	76	1615	1588	27
Michoacán	2043	1869	174	14	1	13	163	160	3
Morelos	1523	1376	147	14	2	12	265	262	3
Nayarit	607	552	55	7	3	4	53	50	3
Nuevo León	1546	1372	174	14	6	8	93	88	5
Oaxaca	1421	1291	130	12	5	7	151	148	3
Puebla	3282	2914	368	39	23	16	357	345	12
Querétaro	973	875	98	10	5	5	94	92	2
Quintana Roo	1944	1843	101	20	19	1	354	346	8
San Luis Potosí	996	882	114	10	3	7	52	50	2
Sinaloa	3433	3261	172	26	15	11	493	484	9
Sonora	1992	1864	128	13	2	11	124	121	3
Tabasco	4336	4009	327	34	8	26	519	510	9
Tamaulipas	1688	1554	134	14	6	8	102	99	3
Tlaxcala	1084	989	95	10	12	-2	143	139	4
Veracruz	3872	3624	248	29	17	12	533	524	9
Yucatán	1920	1771	149	14	6	8	148	144	4
Zacatecas	305	284	21	0	1	-1	33	33	0

6. Discusión de resultados y conclusiones

A continuación, se muestran las predicciones de número de casos, número de hospitalizaciones en UCI (unidad de cuidados intensivos) y número de muertes en cada entidad mexicana obtenidas del modelo implementado, usando la base de datos abiertos de COVID-19 de la DGE correspondiente al día 30/05/2020 y distintos valores para las variables Δt y k_0 . Estos resultados se encuentran en las Tablas 3, 4 y 5. En la Tabla 2 se describen los componentes.

Tabla 4. Evolución: $\Delta t = 7$, $k_0 = 0,75$.

Entidad	CE	CA	ΔC	HE	HA	ΔH	ME	MA	ΔM
República Mexicana	89817	87347	2470	726	312	414	10010	9767	243
Aguascalientes	846	815	31	6	2	4	36	34	2
Baja California	5113	5023	90	35	7	28	873	861	12
Baja California Sur	610	605	5	1	0	1	36	36	0
Campeche	573	563	10	8	12	-4	72	68	4
Coahuila	1117	1084	33	8	1	7	79	77	2
Colima	145	141	4	2	2	0	20	19	1
Chiapas	1793	1705	88	31	21	10	148	137	11
Chihuahua	1587	1569	18	12	16	-4	320	314	6
Ciudad de México	24961	24263	698	199	42	157	2272	2203	69
Durango	384	365	19	6	3	3	38	37	1
Guanajuato	1536	1463	73	17	10	7	115	108	7
Guerrero	1874	1827	47	18	9	9	276	270	6
Hidalgo	1723	1670	53	15	5	10	286	284	2
Jalisco	1754	1660	94	28	17	11	148	136	12
Edo. De México	14667	14264	403	87	31	56	1615	1588	27
Michoacán	1926	1869	57	13	1	12	163	160	3
Morelos	1412	1376	36	13	2	11	265	262	3
Nayarit	570	552	18	7	3	4	53	50	3
Nuevo León	1419	1372	47	11	6	5	93	88	5
Oaxaca	1329	1291	38	9	5	4	151	148	3
Puebla	3027	2914	113	35	23	12	357	345	12
Querétaro	903	875	28	10	5	5	94	92	2
Quintana Roo	1882	1843	39	19	19	0	354	346	8
San Luis Potosí	920	882	38	9	3	6	52	50	2
Sinaloa	3323	3261	62	24	15	9	493	484	9
Sonora	1902	1864	38	12	2	10	124	121	3
Tabasco	4124	4009	115	28	8	20	518	510	8
Tamaulipas	1599	1554	45	13	6	7	102	99	3
Tlaxcala	1010	989	21	10	12	-2	143	139	4
Veracruz	3689	3624	65	28	17	11	533	524	9
Yucatán	1813	1771	42	12	6	6	148	144	4
Zacatecas	286	284	2	0	1	-1	33	33	0

6.1. Interpretación de resultados

Según muestran las predicciones en las Tablas 3, 4 y 5, en el mejor de los casos el valor mínimo de nuevos casos pasados siete días es de 2470 casos a nivel nacional, que corresponde al escenario en el que el 75 % de la población permanece en cuarentena durante toda la semana. El hecho que el valor de Casos no disminuya en comparación a días anteriores, evidencia que la curva de casos no se está aplanando.

Además, se puede decir que partir de lo experimentado en este trabajo, la curva solo se puede aplanar si se siguen manteniendo las medidas preventivas de salud en cuanto a higiene, y disminuir el contacto con otras personas para evitar en la medida de lo posible transmitir el virus. De otro modo, con una observación empírica, la curva no podrá ser aplanada a corto plazo.

Tabla 5. Evolución: $\Delta t = 30, k_0 = 0,5$. El doceavo renglón indica que Guanajuato tiene 10 pacientes en UCI y se predicen 48 nuevos pacientes en un mes si el 50% de la población está en confinamiento.

Entidad	CE	CA	ΔC	HE	HA	ΔH	ME	MA	ΔM
República Mexicana	216707	87347	129360	1926	312	1614	11889	9767	2122
Aguascalientes	2281	815	1466	17	2	15	57	34	23
Baja California	7964	5023	2941	42	7	35	933	861	72
Baja California Sur	814	605	209	2	0	2	41	36	5
Campeche	923	563	360	4	12	-8	81	68	13
Coahuila	2420	1084	1336	17	1	16	100	77	23
Colima	438	141	297	5	2	3	23	19	4
Chiapas	6536	1705	4831	44	21	23	206	137	69
Chihuahua	2528	1569	959	10	16	-6	335	314	21
Ciudad de México	52577	24263	28314	569	42	527	2788	2203	585
Durango	1177	365	812	10	3	7	51	37	14
Guanajuato	5779	1463	4316	58	10	48	167	108	59
Guerrero	4173	1827	2346	27	9	18	312	270	42
Hidalgo	5070	1670	3400	39	5	34	330	284	46
Jalisco	7320	1660	5660	78	17	61	219	136	83
Edo. De México	41539	14264	27275	448	31	417	1955	1588	367
Michoacán	5110	1869	3241	31	1	30	194	160	34
Morelos	4106	1376	2730	46	2	44	305	262	43
Nayarit	1344	552	792	8	3	5	63	50	13
Nuevo León	4604	1372	3232	52	6	46	135	88	47
Oaxaca	4182	1291	2891	28	5	23	180	148	32
Puebla	9991	2914	7077	85	23	62	443	345	98
Querétaro	2527	875	1652	20	5	15	118	92	26
Quintana Roo	3070	1843	1227	13	19	-6	376	346	30
San Luis Potosí	2720	882	1838	20	3	17	74	50	24
Sinaloa	5378	3261	2117	29	15	14	538	484	54
Sonora	3276	1864	1412	16	2	14	147	121	26
Tabasco	8251	4009	4242	61	8	53	591	510	81
Tamaulipas	3483	1554	1929	28	6	22	137	99	38
Tlaxcala	3055	989	2066	20	12	8	169	139	30
Veracruz	9020	3624	5396	63	17	46	596	524	72
Yucatán	4188	1771	2417	32	6	26	186	144	42
Zacatecas	863	284	579	4	1	3	39	33	6

6.2. Trabajo futuro

Actualmente, se están generando archivos de documentación del módulo de Python con los resultados hasta hoy alcanzados. Se busca distinguir este proyecto de otros similares con la creación de un portal web que presente la posibilidad de realizar predicciones personalizadas de contexto particular, adaptando el software de los autores de este artículo un servicio en la nube.

Referencias

1. Arenas, A., Cota, W., Gómez-Gardeñes, J., Gómez, S., Granell, C., Matamalas, J. T., Soriano, D., Steinegger, B.: A mathematical model for the spatiotemporal epidemic spreading of COVID19. Cold Spring Harbor Laboratory (2020) doi: 10.1101/2020.03.21.20040022
2. Arranz, J., María, J.: COVID-19, SARS-CoV-2. GdT-semFYC en Enfermedades Infecciosas (2020) www.semfyec.es/wp-content/uploads/2020/03/COVID-19-semFYC.pdf
3. Bernábe-Loranca, M. B., Sarmiento-Barrios, E., Cerón-Garnica, C., Rubio-Quintero, R., Martínez-Guzmán, G.: Simulador estadístico de contagios para COVID-19 usando 13 variables del sistema de salud. *Revista Pistas Educativas*, vol. 42, no. 138 (2021)
4. Bernábe-Loranca, M. B., González-Velázquez, R., Granillo-Martínez, E., Ruiz-Vanoye, J. A., Canan, A. C.: Towards an approach of the contagion curve for COVID-19 in Mexico. In: *Advances in Intelligent Systems and Computing*. Springer International Publishing, pp. 553–566 (2021) doi: 10.1007/978-3-030-71187-0_51
5. INEGI: Datos. www.inegi.org.mx/datos/ (2020)
6. INEGI: Encuesta Intercensal 2015. www.inegi.org.mx/programas/intercensal/2015/default.html#Tabulados (2016)
7. Li, F.: Structure, function, and evolution of coronavirus spike proteins. *Annual Review of Virology*, vol. 3, no. 1, pp. 237–261 (2016) doi: 10.1146/annurev-virology-110615-042301
8. Secretaría de Salud: COVID-19 Tablero México (2020) coronavirus.gob.mx/datos/#DOView
9. Secretaría de Salud: Personal de Salud (2020) coronavirus.gob.mx/personal-de-salud/
10. Secretaría de Salud: Datos Abiertos - Dirección General de Epidemiología (2020) www.gob.mx/salud/documentos/datos-abiertos-152127
11. Serdán: Mapa con los polígonos de los municipios de México (2015) albertoserdan.carto.com/tables/mapa_municipios_de_mexico/public/map
12. Zhou, F., Yu, T., Du, R., Fan, G., Liu, Y., Liu, Z., Xiang, J., Wang, Y., Song, B., Gu, X., Guan, L., Wei, Y., Li, H., Wu, X., Xu, J., Tu, S., Zhang, Y., Chen, H., Cao, B.: Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: A retrospective cohort study. *The Lancet*, vol. 395, no. 10229, pp. 1054–1062 (2020) doi: 10.1016/s0140-6736(20)30566-3