

Modelación bayesiana de patrones espacio-temporales de la incidencia acumulada de COVID-19 en municipios de México

Bayesian modeling of spatiotemporal patterns of the cumulative incidence of COVID-19 in municipalities of Mexico

Gerardo Núñez Medina

Orcid: 0000-0001-8038-091X

gnunez@colef.mx

COLEF, Departamento de Estudios de Población, México

Resumen

El trabajo busca modelar la distribución de la tasa de incidencia acumulada de COVID-19 en los municipios de México a través del ajuste de tres modelos lineales generalizados (en competencia) con efectos espaciales y temporales y función de enlace *Poisson*. Se utilizaron datos de casos confirmados de COVID-19, reportados por la Secretaría de Salud de febrero a julio de 2020. Con el objetivo de reducir los costos computacionales asociados a la estimación de múltiples parámetros, con grandes cantidades de datos, se optó por utilizar la aproximación integrada anidada de Laplace en lenguaje R (R-INLA). Los modelos fueron evaluados a través del criterio de información Akaike (AIC), donde el mejor resultó ser el Modelo No Paramétrico de Interacción Espacio-Temporal. Los resultados permiten confirmar la presencia de importantes niveles de heterogeneidad en la distribución espacio-temporal de las tasas de incidencia de COVID-19 entre municipios de México.

Abstract

The work seeks to model the distribution of cumulative incidence rate of COVID-19 in municipalities of Mexico through the adjustment of three generalized linear models (in competition) with spatial and temporal effects and Poisson link function. Data of confirmed cases of COVID-19, reported

Palabras Clave

Modelos lineales
generalizados mixtos
COVID-19
R-INLA

by Health Secretary of Mexico, from February to July 2020, were used. In order to reduce the computational costs associated with the estimation of multiple parameters with large amounts of data, we chose the Integrated Nested Laplace Approximation implemented in R language (R-INLA). The models were evaluated through the Akaike (AIC) criterion, and the best was the Non-parametric Model of Space-Time Interaction. The results confirm the presence of significant levels of heterogeneity in the spatial-temporal distribution of COVID-19 incidence's rates among municipalities of Mexico.

Keywords

Bayesian generalized linear models
COVID-19
R-INLA

Recibido: 31/08/2020
Aceptado: 23/10/2020

Introducción

El inicio de la pandemia de SARS-CoV-2 (*Severe Acute Respiratory Syndrome Coronavirus 2*), también llamada COVID-19 (Coronavirus Disease 2019), desató la amenaza más importante del mundo en materia de salud pública en más de un siglo. Esto se debe a que la tasa básica de reproducción R_0 del COVID-19 es de 2.68 personas por individuo contagiado —1.28 en el caso de la gripe común—, cifra que representa el número medio de nuevos contagios generados por un solo infectado durante su fase infecciosa (Martínez y Vargas, 2020), además de una elevada tasa de letalidad que es en promedio superior al 2%.

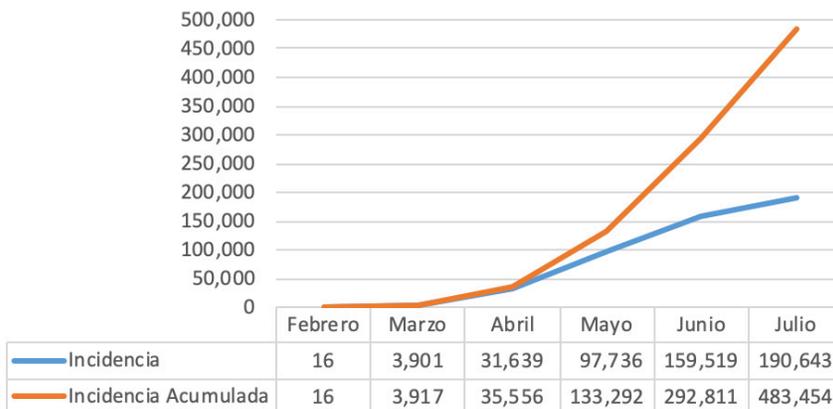
La rápida expansión del COVID-19, a nivel global, se debe en principio a su elevada tasa de reproducción, pero también es producto de las dificultades prácticas de romper las cadenas de contagios, lo cual depende de una adecuada aplicación de prácticas de prevención orientadas a la detección, atención y contención oportuna de todos los casos positivos (Ramos, 2020), para lo cual, es importante dar seguimiento a los niveles de dispersión espacial de la enfermedad. El COVID-19 inició en la ciudad de Wuhan, China, en noviembre de 2019; llegó a Europa y al Norte de América entre diciembre de 2019 y enero de 2020. Los primeros casos en México se presentaron el mes de febrero de 2020 (Organización Panamericana de la Salud, 2020), a la par que el gobierno de México inició la aplicación de medidas orientadas a la contención y mitigación del proceso de transmisión.

La tasa básica de reproducción del COVID-19 implica que la epidemia tendrá un crecimiento exponencial y solo podrá ser atenuada con medidas de contención como el distanciamiento social y el uso de mascarillas, además de otras prácticas de prevención, las cuales desafortunadamente

tienen costos económicos y sociales muy elevados (Comisión Económica para América Latina y el Caribe, 2020). De esta forma, la evolución del COVID-19 en México registró un total de 16 casos positivos confirmados en febrero de 2020, mientras que para el último día del mes de julio se registraron 483,454 casos confirmados acumulados (Gráfica 1). Durante los últimos seis meses, la pandemia ha evolucionado de forma creciente y continuada, no solo a lo largo del tiempo, sino a través del espacio, donde el virus se propagó de 13 a 1,847 municipios entre los meses de febrero a julio de 2020. Bajo este escenario, es de esperarse que la pandemia continúe evolucionando hasta llegar a un punto máximo de contagios a finales del mes de agosto y logre infiltrarse a la totalidad de municipios del país.

El trabajo hace una revisión de la evolución espacio-temporal de la pandemia de COVID-19 en los municipios de México, para lo que se utilizan tres modelos bayesianos espacio-temporales: (1) Modelo Clásico Paramétrico, (2) Modelo Dinámico No Paramétrico y (3) Modelo No Paramétrico de Interacción Espacio-Temporal, con la finalidad de evaluar la velocidad de propagación del virus, en términos de la tasa de incidencia acumulada estandarizada (Tiae) de contagios a lo largo de los municipios del país y de seis meses de evolución de la pandemia.

Gráfica 1. Evolución mensual de casos de COVID-19 en México



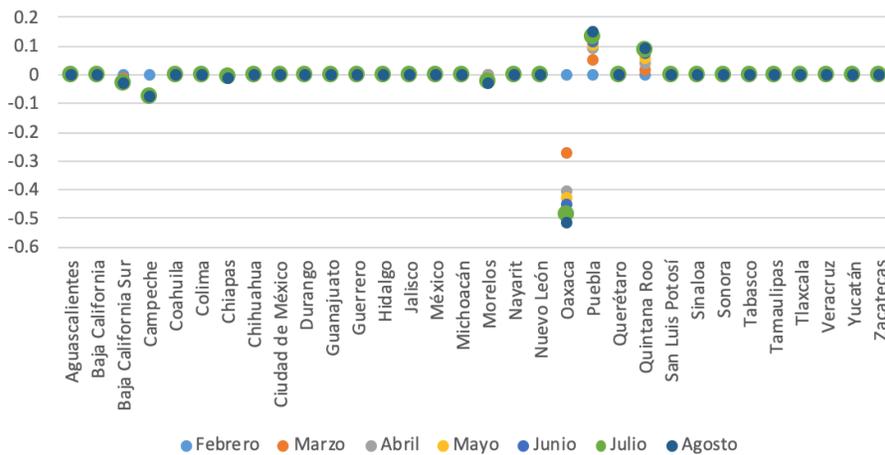
Fuente: elaboración propia con datos de la Secretaría de Salud.

La dispersión espacial de casos confirmados de COVID-19 en México inició en los municipios con mayor densidad poblacional, como los ubicados en grandes zonas urbanas: Ciudad de México, Monterrey y Tijuana. En enero se reportaron dos casos en dos municipios, para febrero se registraron 16 casos

en 13 municipios; mientras que para finales de julio 1,847 municipios presentaron poco más de 480,000 casos confirmados de COVID-19 (Gráfica 1). En contraste, 609 municipios se encontraron libres de la pandemia, debido básicamente al importante grado de aislamiento geográfico y social en el que viven normalmente y, posiblemente, por la adecuada aplicación de los programas de prevención.

Un análisis estadístico preliminar de la evolución por entidad federativa del COVID-19 en México se presenta en la Gráfica 2. En la misma se observa la evolución del logaritmo natural de la tasa de incidencia acumulada de COVID-19 para los meses de febrero a julio de 2020. En general, la mayor parte de las entidades presentan un riesgo bajo y constante de contagio en el periodo analizado, sin embargo, resaltan dos entidades vecinas, Oaxaca y Puebla, con patrones de comportamiento extremo.

Gráfica 2. Evolución del logaritmo natural de la tasa de incidencia acumulada de COVID-19 por entidad federativa, febrero-julio de 2020



Fuente: elaboración propia con datos de la Secretaría de Salud.

Según se muestra en la Gráfica 2, las entidades de Puebla y Quintana Roo presentan las tasas de contagio más elevadas; mientras que los estados de Oaxaca y Campeche alcanzan las menores tasas de contagio, muy por debajo de la media nacional, en especial el estado de Oaxaca. Esta tendencia parece sostenerse a lo largo de los meses analizados, no obstante, debe considerarse que un análisis a nivel de entidad federativa oculta la variabilidad y dispersión del riesgo por contagio en niveles geográficos de agregación menores, como es el caso de los municipios.

Materiales y Método

Fuentes de datos

Desde el 13 de abril del año 2020, la Secretaría de Salud de México puso a disposición de la comunidad científica las bases de datos que permiten conocer la evolución diaria del número de casos y defunciones por COVID-19 a nivel municipal. Si bien, los datos presentan problemas y limitaciones relacionadas con la estimación del número de contagios, estos problemas derivan de la fragmentación del sistema de salud mexicano, donde cada estado recaba su información con una metodología distinta, y es la Secretaría de Salud quien agrupa los datos y aplica un modelo denominado "Centinela", para estimar la totalidad de casos leves y ambulatorios de COVID-19 dispersos en el país. El modelo Centinela utiliza una muestra para estimar el número de casos, es decir, en México no se cuentan la totalidad de casos positivos de COVID-19 identificados a partir de una prueba de laboratorio. Sin embargo, no se tienen elementos para afirmar que la Secretaría de Salud oculte o tergiverse información sobre el impacto del COVID-19 en México (Castañeda y Garrido, 2020).

Las bases de datos utilizadas para modelar la evolución espacio-temporal del COVID-19 en municipios de México fueron publicadas por la Dirección de Información del Sistema de Vigilancia Epidemiológica de la Secretaría de Salud de México (Secretaría de Salud, 2020). Los datos sobre contagios de COVID-19 fueron recuperados el día 3 de agosto del 2020, mientras que las cifras relativas a la población municipal fueron tomadas de las Proyecciones de la Población de los Municipios de México 2015-2030 base 1 y 2, realizadas por el Consejo Nacional de Población (Partida y García, 2018).

De la base de datos colectada el 3 de agosto, se procesaron un total de 1.2 millones de casos registrados como sospechosos, de ellos, se analizaron 483,454 casos confirmados, es decir, los casos donde el resultado de la prueba de COVID-19 fue positivo. La localización espacial se asignó de acuerdo a la entidad y municipio de residencia del paciente, y la temporalidad conforme la fecha de inicio de síntomas reportada por el enfermo, no se asignó la fecha de arribo a la unidad médica, debido a la discrepancia entre fechas. Así, la tasa de incidencia acumulada de COVID-19 se estimó como el cociente del total de casos acumulados confirmados hasta el mes t , entre la población expuesta.

Modelos espacio-temporales

Los datos espacio-temporales definen un proceso estocástico a partir de una sucesión indexada de la forma:

$$Y(s,t) = \{y(s,t) | (s,t) \in \mathbb{R}^2 \times \mathbb{R}\}$$

Donde $y(s,t)$ representa el número observado de casos confirmados de COVID-19 en el municipio s entre 1 y 2,457 municipios, a un tiempo t entre 1 y 6 meses (Blangiardo y Cameletti, 2015, p. 235). De esta forma, la incidencia acumulada de casos de COVID-19, en el municipio s , puede modelarse a partir de una distribución *Poisson* dada por:

$$Y(s,t) \sim \text{Poisson}(\lambda_{st})$$

Bajo el supuesto que datos observados se comportan siguiendo una distribución *Poisson*, con media λ_{st} definida como el producto de la tasa de incidencia acumulada estandarizada de COVID-19 ($Tiae$) y el número esperado de casos de COVID-19, $\lambda_{st} = Tiae_{st} \times E_{st}$ (Moraga, 2019). Donde la tasa de incidencia acumulada estandarizada, o índice estandarizado de incidencia, se define como el cociente del total de casos acumulados confirmados de COVID-19 y el número esperado de casos:

$$Tiae(s,t) = \frac{Y(s,t)}{E(s,t)}$$

La tasa cuantifica el riesgo de enfermarse en el municipio s , al mes t , mientras que los casos esperados $E(s,t)$ representan el número de casos que se esperaría observar, si la población del municipio s presentará el mismo comportamiento que la población estándar (o estatal). El número esperado de casos puede estimarse utilizando el método de estandarización indirecta como:

$$E(s,t) = r(\bullet,t) \times n(s,t)$$

Donde $r(\bullet,t)$ es la tasa de incidencia acumulada de COVID-19 la población de la Entidad Federativa correspondiente y $n(s,t)$ es la población total del municipio s , ambos al tiempo t (Gomez-Rubio, 2020). Es importante señalar que la $Tiae$ se presenta en escala logarítmica debido a que permite manejar de mejor forma la dispersión presente en las tasas estimadas. Una

$Tiae(s,t) > 1$ es un indicador de que el riesgo en el municipio s , al tiempo t , es mayor que el riesgo presente en la población estatal.

El predictor lineal en escala logarítmica se define como:

$$\eta_{st} = E[G(Y(s,t))] = \ln(Tiae(s,t))$$

Donde la función de enlace empleada para el caso de distribuciones *Poisson* es el logaritmo natural (Wang, Yue y Faraway, 2018).

Modelo Clásico Paramétrico

La formulación paramétrica clásica de los modelos espacio-temporales para el predictor lineal η_{st} se expresa como:

$$\eta_{st} = \beta_0 + u_s + v_s + (\beta + \delta_s)t \quad (1)$$

Donde β_0 representa la tasa promedio de incidencia acumulada de COVID-19 en todos los municipios del país, u_s representa los efectos aleatorios espacialmente estructurados y v_s los efectos espacialmente no estructurados, de manera que, cada unidad espacial presenta una tendencia temporal con intercepto espacial ($\beta_0 + u_s + v_s$) y pendiente ($\beta + \delta_s$) (Blangiardo y Cameletti, 2015). El vector de parámetros a estimar está dado por $\theta = \{\beta_0, u, v, \beta, \delta\}$, mientras que el vector de hiperparámetros se define como $\psi = \{\tau v, \tau u, \tau \delta\}$.

Modelo Dinámico No Paramétrico

Los modelos de tipo paramétrico (1) imponen restricciones sobre la linealidad del efecto temporal global (β) y sobre la tendencia espacial diferencial (δ_s), mientras que los modelos no paramétricos, evitan dichas restricciones al utilizar una forma dinámica (Bivand, Gómez-Rubio y Rue, 2015), que permiten modelar tendencias temporales a través de efectos estructurados ρt y efectos no estructurados φt :

$$\eta_{st} = \beta_0 + u_s + v_s + \rho_t + \varphi_t \quad (2)$$

La tendencia temporal estructurada ρt se modelada dinámicamente utilizando una caminata aleatoria de orden uno ($rw(1)$) cuya distribución corresponde a:

$$\rho_t | \rho_{t-1} \sim N(\rho_{t-1}, \sigma^2)$$

La tendencia temporal no estructurada φ_t se modela utilizando una distribución previa gaussiana intercambiable $N(0, 1/\tau\varphi)$ (Schrödle y Held, 2011), mientras β_0 , u_s y v_s se definen de la misma forma que en el modelo (1). El modelo (2) asume una tendencia temporal no paramétrica cuyos parámetros e hiperparámetros a estimar se especifican a partir de los vectores $\theta = \{\beta_0, u, v, \rho, \varphi\}$ y $\psi = \{\tau u, \tau v, \tau\rho, \tau\varphi\}$.

Modelo No Paramétrico con Interacción Espacio-Temporal

El modelo no paramétrico de interacción espacio-temporal estima la interacción entre efectos espaciales y temporales al introducir al modelo (2) el término γ_{st} , con lo que se tiene la especificación:

$$y_{st} = \eta_{st} = \beta_0 + u_s + v_s + \rho_t + \varphi_t + \gamma_{st} \quad (3)$$

El efecto de interacción espacio-tiempo γ_{st} sigue una distribución gaussiana dada por $\gamma_{st} \sim Normal(0, 1/\tau_\gamma)$. Los elementos restantes se definen como en el modelo (2) (Sharafifi et al., 2018). Los vectores de parámetros e hiperparámetros a estimar se definen como $\theta = \{\beta_0, u, v, \rho, \varphi, \gamma\}$ y $\psi = \{\tau u, \tau v, \tau\rho, \tau\varphi, \tau\gamma\}$. Para la implementación del modelo (3) en la aproximación integrada anidada de Laplace en lenguaje R (R-INLA, por sus siglas en inglés, *Integrated Nested Laplace Approximation*), se especificó el término γ_{st} como la interacción espacio-temporal del efecto espacial no estructurado v_s y el efecto temporal estructurado φ_t (interacción tipo II) (Blangiardo y Cameletti, 2015) bajo las especificaciones expuestas en el siguiente apartado.

Efectos espaciales y temporales

El componente espacial estructurado (u_s) sigue una distribución condicional dada por:

$$u_s | u_{-s} \sim N\left(\frac{1}{\#N(s)} \sum_{j=i}^n c_{sj} u_j, \sigma_s^2\right) \quad (4)$$

para toda $i \neq j$, donde c_{ij} representa el criterio de vecindad; si $c_{sj} = 1$ las áreas s y j son vecinas, en caso contrario $c_{sj} = 0$, mientras que la varianza $\sigma_s^2 = \sigma_u^2 / \#N(s)$ del área s , depende del número de vecinos $N(s)$, conocido como efecto autorregresivo condicional intrínseco *ICAR*(1) por sus siglas en inglés (Rue, Martino y Chopin, 2009).

La especificación Besag – York – Mollié (BYM) dada por la ecuación (4) supone además que el componente espacial no estructurado v_s tiene una distribución previa $Normal(0, \sigma v^2)$, donde la varianza σv^2 representa la dispersión de efectos aleatorios espaciales no estructurados (Riebler *et al.*, 2016). En general, los modelos espacio-temporales suponen la existencia de términos espaciales y temporales separables en el predictor lineal, por lo que la estructura de su matriz de covarianza puede ser separada como el producto de una matriz de efectos espaciales y otra de efectos temporales (Martino y Rue, 2008).

Calidad de ajuste

El criterio de información de Akaike (AIC) es una medida de bondad de ajuste de un modelo, la cual se obtiene al ponderar entre el ajuste de los datos y el número de parámetros estimados. Se construye tomando como base el logaritmo de la función de verosimilitud:

$$AIC = -2l(\hat{\theta}) + 2k$$

Donde $l(\hat{\theta})$ es la función de máxima log-verosimilitud y k es el número de parámetros del modelo (Moraga, 2019). El modelo con mejor ajuste debe tener un AIC menor que el resto de los modelos en competencia, mientras que el número efectivo de parámetros estimados debe ser menor. Se presenta también el criterio DIC (*Deviance Information Criterion*) con la finalidad de confirmar la selección del modelo, dado que este incorpora el número de observaciones.

Tabla 1. Criterios de información AIC y DIC de los modelos (1) a (3)

	Modelo	AIC	DIC
(1)	$y_{it} = \eta_{it} = \beta_0 + u_i + v_i + (\beta + \delta_i)t$	66,848.71	64,523.66
(2)	$y_{st} = \eta_{st} = \beta_0 + u_i + v_i + \rho_t + \varphi_t$	81,239.71	78,370.27
(3)	$y_{st} = \eta_{st} = \beta_0 + u_i + v_i + \rho_t + \varphi_t + \gamma_{it}$	56,352.33	57,462.68

Fuente: elaboración propia con base en estimaciones de R-INLA.

La Tabla 1 muestra que el Modelo No Paramétrico de Interacción Espacio-Temporal presenta el mejor ajuste para las tasas de COVID-19, con una ventaja muy amplia sobre los modelos en competencia.

Resultados

La aplicación de R-INLA (R Core Team, 2016) hizo posible la estimación con éxito de los modelos espacio-temporales propuestos. Esto se debe a que la aproximación integra importantes avances en estadística computacional, como el derivado de la implementación de un algoritmo determinístico, que permite obtener estimaciones bayesianas de manera rápida y exacta (Martins *et al.*, 2012). De esta forma R-INLA se ha convertido en una herramienta práctica y accesible para la estimación de modelos bayesianos espacio-temporales.

Con el objetivo de modelar el comportamiento espacio-temporal de las tasas de incidencia acumulada de COVID-19 en los municipios de México, se ajustaron tres modelos: (1) Modelo Clásico Paramétrico; (2) Modelo Dinámico No Paramétrico; y (3) Modelo No Paramétrico de Interacción Espacio-Temporal. Según el criterio AIC de selección de modelos en competencia, el modelo que ajusta mejor los datos analizados resultó ser el Modelo de Interacción Espacio-Temporal (Tabla 1). Para fines de comparación, se presentan las estimaciones de los tres modelos ajustados.

Tabla 2. Estimaciones posteriores de efectos fijos, modelos espacio-temporales

Modelo Clásico Paramétrico	Media	Desv. Est.	Q1	Q2	Q3	Moda
Constante	-1.173	0.026	-1.225	-1.173	-1.122	-1.173
Año	0.015	0.003	0.010	0.015	0.021	0.015
Modelo no paramétrico						
Constante	-1.044	0.016	-1.075	-1.044	-1.012	-1.044
Modelo de interacción						
Constante	-0.308	0.046	-0.399	-0.308	-0.216	-0.308

Fuente: elaboración propia con base en estimaciones de R-INLA.

Los resultados obtenidos de las estimaciones posteriores de los parámetros para cada uno de los modelos incluyen la estimación de efectos fijos presentados en la Tabla 2. La estimación de efectos aleatorios espaciales y temporales presentados en la Tabla 3 y en la Figura 2; y la estimación de efectos temporales estructurados y no estructurados presentados en la Tabla 3 y en la Gráfica 3, mientras que la Figura 1 muestra la distribución espacio-temporal de las tasas de incidencia acumulada de COVID-19 de los municipios de México para los meses analizados.

Las estimaciones posteriores de los efectos fijos para la media, desviación estándar y cuartiles de los modelos (1) a (3) pueden verse en la Tabla 2. El parámetro β_0 puede interpretarse como el logaritmo natural del riesgo de contagio por COVID-19 para México durante los meses analizados. Lo que implica, para el caso del Modelo No Paramétrico de Interacción Espacio-Temporal, que el riesgo de contagio a nivel nacional para los meses estudiados corresponde a $0.734 = \exp(-0.308)$. Se debe señalar que la tasa de incidencia acumulada observada para México se estimó en 0.99 (Gráfica 2).

Tabla 3. Estimaciones posteriores de efectos aleatorios, modelos espacio-temporales

Modelo Clásico Paramétrico		Media	Desv. Est.	Q1	Q2	Q3
u_i	Efecto espacial estructurado	2.899	0.396	2.224	2.861	3.778
v_i	Efecto espacial no estructurado	0.451	0.045	0.366	0.449	0.543
β	Efecto temporal global	82.441	3.940	75.005	82.335	90.482
Modelo no paramétrico						
u_i	Efecto espacial estructurado	3.564	0.346	2.938	3.545	4.294
v_i	Efecto espacial no estructurado	0.638	0.054	0.537	0.636	0.751
ρt	Efecto temporal estructurado	39465.504	25696.660	8271.943	33701.107	104712.691
ϕt	Efecto temporal no estructurado	41507.936	26814.281	8923.013	35473.575	110047.838
Modelo de interacción						
u_i	Efecto espacial estructurado	2353.085	2019.316	261.700	1807.879	7685.023
v_i	Efecto espacial no estructurado	2087.713	1940.527	204.351	1538.452	7258.929
ρt	Efecto temporal estructurado	363.377	729.147	22.775	167.712	1912.272
ϕt	Efecto temporal no estructurado	13.070	7.273	3.666	11.579	31.328
γ_{it}	Efecto int. espacio-temporal	14.254	0.488	13.328	14.244	15.242

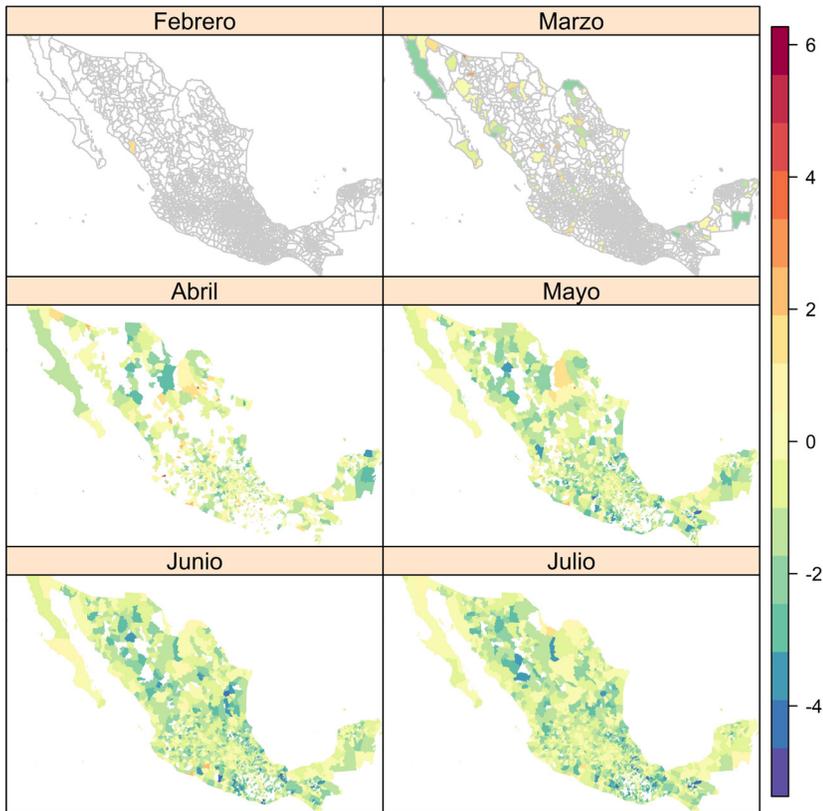
Fuente: elaboración propia con base en estimaciones de R-INLA.

La Tabla 3 presenta el valor medio, desviación estándar y cuartiles estimados para los efectos temporales estructurados y no estructurados de los modelos (1) a (3). En particular puede observarse que los modelos (1) y (2) asignan mayor peso a los efectos temporales con relación a los efectos

espaciales, con lo que sobrestiman el efecto del tiempo. Por el contrario, el modelo (3) da mayor peso a las estimaciones de los efectos espaciales (estructurados y no estructurados), por lo que las relaciones espaciales y de vecindad son consideradas en este modelo como un factor explicativo fundamental para modelar la evolución de los contagios por COVID-19 en los municipios de México.

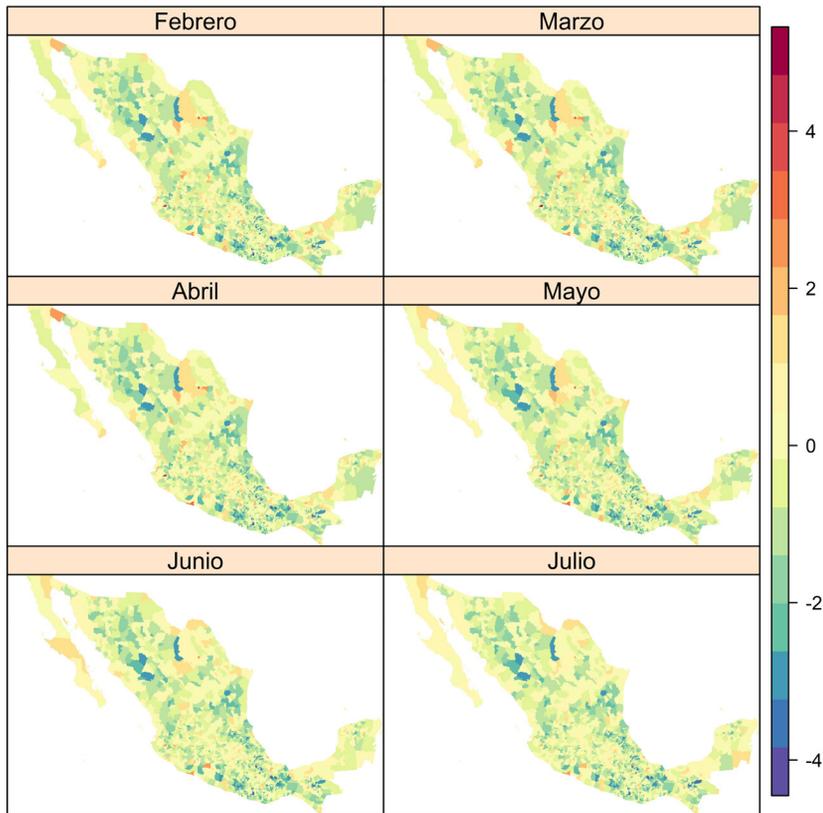
La Figura 1 presenta las variaciones espacio-temporales de la tasa de incidencia acumulada de COVID-19 en los municipios de México, las tasas se presentan en escala logarítmica para evitar problemas de sobredispersión de datos. Asimismo, en la Figura 1 es posible observar la evolución mes a mes de las tasas de COVID-19. Mientras que en febrero se registraron contagios en solo 13 municipios (los municipios en blanco no presentan contagios), para el mes de marzo se registraron casos confirmados en 352 municipios, la cifra de municipios afectados continuó creciendo hasta el mes de julio, cuando se contabilizaron 1,847 municipios afectados (Figura 1).

Figura 1. Tasa de incidencia acumulada de COVID-19 por municipio, febrero-julio 2020



Fuente: elaboración propia con base en estimaciones de R-INLA.

Figura 2. Media posterior de índice estandarizado de incidencia de COVID-19, modelo (3)



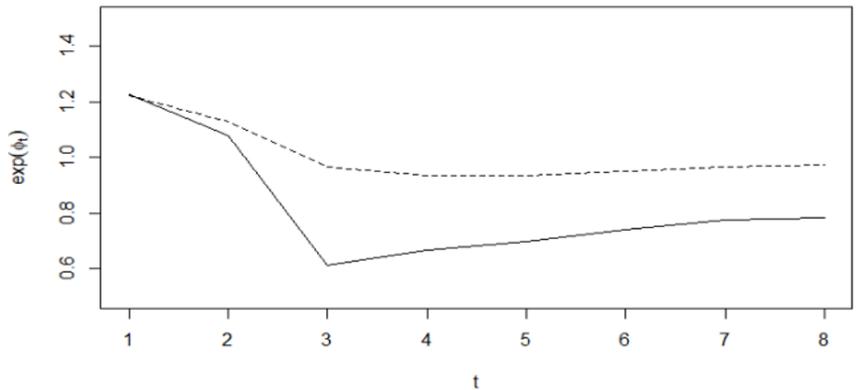
Fuente: elaboración propia con base en estimaciones de R-INLA.

La Figura 2 presenta información relativa a la evolución de la media posterior estimada por el modelo (3) de interacción espacio-temporal, para la Tiae de COVID-19 de los municipios de México entre los meses de febrero a julio, las tasas se presentan en escala logarítmica para evitar problemas de sobredispersión. En la Figura 2 se muestra un ajuste del comportamiento de las tasas de incidencia acumulada de COVID-19, a la par que estima correctamente el valor medio de las tasas de COVID-19 en los municipios donde no se cuenta con información. Sin embargo, también se observa un claro efecto de sobreajuste de las tasas en los meses de febrero y marzo.

La evolución de las tasas muestra la ausencia de contagios en un importante número de municipios durante los primeros dos meses de la pandemia (febrero y marzo), como puede verse en la Figura 1. El modelo (3) ajusta este hecho asignando un riesgo de contagio bajo o muy bajo a los municipios donde no se registraron contagios, como se observa en la

Figura 2. En general, la estimación de las tasas presenta un efecto espacial más fuerte que el temporal, por lo que la evolución de los contagios tiene un fuerte componente espacial que se “dispersa” a lo largo del tiempo.

Gráfica 3. Tendencia posterior de los efectos temporales estructurados ρt y no estructurados φt del Modelo de Interacción Espacio-Temporal



Efecto temporal no estructurado línea sólida φt

Efecto temporal estructurado línea punteada ρt

Fuente: elaboración propia con base en estimaciones de R-INLA.

La tendencia estimada por el Modelo de Interacción Espacio-Temporal para los efectos temporales estructurados ρt y no estructurados φt puede verse en la Gráfica 3. Se observa una reducción continuada de ambos efectos temporales entre febrero y marzo, posteriormente, se observa un ligero incremento; sin embargo, el efecto en ambos casos se mantiene por debajo de la unidad, pero en todo momento el efecto estructural se mantiene de forma sostenida por arriba del efecto no estructural, lo que implica que el efecto de las variaciones temporales, producto de variaciones aleatorias, tiene una menor incidencia sobre la evolución de los contagios de COVID-19, que los efectos, producto de cambios temporales estructurados.

El valor estimado del término de interacción espacio-temporal γ_{st} , del modelo (3), captura adecuadamente la dependencia de la relación espacio-temporal de las tasas de incidencia acumuladas de COVID-19, debido a: 1) un elevado nivel de significancia estadística —presenta una desviación estándar pequeña en relación con el valor medio estimado—, y 2) logra una importante reducción del valor estimado AIC en relación con los modelos en competencia.

Se puede concluir que el riesgo medio de contagio de COVID-19 de un alto porcentaje de municipios de México es bajo. Sin embargo, existe una amplia variación en la incidencia del riesgo entre municipios. En general, puede observarse una gran cantidad de municipios con baja densidad poblacional que presentan un riesgo de contagio por abajo del promedio nacional, mientras en el otro extremo se observa otro importante número de municipios con tasas de contagio altas o muy altas, pero que en general se encuentran densamente poblados. Por ejemplo, en el estado de Oaxaca, en los municipios de Santa Cruz Zenzontepec, Santa María Chilchotla y Santiago Yosondúa, la Tiae estimada por el modelo (3) al mes de julio fue de 0.030, 0.036 y 0.040, respectivamente; y todos los municipios cuentan con poblaciones menores a 350,000 habitantes; mientras que la capital del estado, el municipio de Oaxaca de Juárez, presentó una tasa de 6.35 con 4.2 millones de habitantes.

Discusión

El modelo (3) supone que la variación de la Tiae en cada municipio se obtiene como la suma de efectos aleatorios espaciales, temporales y la interacción de efectos espacio-temporales, por lo que no es posible aislar el comportamiento específico de un municipio en particular, situación que puede causar un sobreajuste del modelo, para algunos municipios durante algún periodo de tiempo (Gomez-Rubio, 2020), rasgo que se presenta con mayor intensidad cuando existe una gran cantidad de datos faltantes.

El modelo (3) presenta altos niveles de correlación espacial, es decir, similitudes entre municipios vecinos, lo que genera estimaciones similares en áreas cercanas, producto de la estimación ICAR; se presenta un alto grado de suavizado que se transfiere en el tiempo al interactuar con el efecto temporal estructural definido, como una caminata aleatoria de orden 1.

La variación observada entre el efecto espacial estructurado u_s y el efecto temporal estructurado ρ_t , donde la estimación media de la densidad posterior de la precisión de los efectos espaciales estructurados es significativamente mayor que la estimada para los efectos temporales estructurados, implica que la capacidad explicativa del modelo (3) debe tomar en cuenta la interacción de los efectos espaciales y temporales (Schrödle y Held, 2011). Esto se debe a que las variaciones entre efectos espaciales y efectos temporales son relativamente grandes, comparadas con la variación entre efectos temporales estructurados y no estructurados (Tabla 3).

Como puede verse en la Gráfica 3, la evolución de los efectos temporales estructurados y no estructurados parece indicar que la tasa de contagios tenderá a estabilizarse en los próximos meses (siempre que se continúen las políticas sanitarias y de distanciamiento social aplicadas hasta el momento). Los efectos espaciales estructurados y no estructurados parecen ejercer una influencia mayor que los efectos temporales, al momento de modelar el comportamiento de los contagios por COVID-19, estos efectos muestran también una reducción en la intensidad de sus interacciones (Figura 2), debido en principio a la reducción en el ritmo de crecimiento de los contagios. Sin embargo, el efecto producido por las interacciones espaciales sobre la tasa de contagio del COVID-19 es en extremo alto, como lo muestran las estimaciones del Modelo de Interacción presentado en la Tabla 3, por lo que es recomendable aplicar políticas de salud orientadas a reducir los niveles de interacción espacial entre municipios y localidades.

Conclusiones

A lo largo de este trabajo se aplicaron tres modelos lineales generalizados bayesianos espacio-temporales, mismos que fueron ajustados con el paquete R-INLA, como un primer ejercicio de estimación que tiene por objeto analizar la evolución espacial y temporal de la tasa de incidencia acumulada de COVID-19 en los municipios de México, incorporando la interacción de efectos espacio-temporales.

El reto de reducir y eliminar los contagios por COVID-19 en el territorio mexicano, implica romper la cadena de contagios, para lo cual resulta fundamental entender la interacción del virus con el territorio y a lo largo del tiempo. Así, esta investigación busca aportar evidencias que ayuden a generar estrategias para reducir la tasa básica de reproducción R_0 del COVID-19, aportando elementos para reforzar las acciones de prevención y contención de contagios con el objetivo de interrumpir el ciclo de transmisión del coronavirus.

Si bien, la evolución de la tasa de incidencia acumulada de COVID-19 para México indica un crecimiento paulatino y constante en el número de contagios (Gráfica 1), así como en el número de municipios afectados por la enfermedad, el cual ha crecido también de forma sostenida durante los meses analizados a causa de una mayor dispersión espacial del COVID-19 (Figuras 1 y 2). No obstante, la tendencia temporal ajustada por el modelo

(3) indica que la tasa de contagio tenderá a estabilizarse a finales del mes de agosto de 2020 (Gráfica 3), siempre que se mantengan las condiciones sanitarias presentes durante el estudio.

Como se ha expuesto, la tasa de incidencia acumulada de COVID-19 en un alto porcentaje de municipios es baja. Sin embargo, presenta también un alto grado de heterogeneidad, afectando de manera diferencial a municipios densamente poblados o con altos niveles de interacción, mientras que protege a aquellos con mayores niveles de aislamiento geográfico o social. La heterogeneidad de la tasa de incidencia combinada con el efecto potencial de la tasa básica de reproducción del COVID-19, implica que es necesario continuar con la aplicación de políticas sanitarias dirigidas a romper la cadena de contagios.

Referencias

- Bivand, R., Gómez-Rubio, V. y Rue, H. (2015). Spatial Data Analysis with R-INLA with Some Extensions. *Journal of Statistical Software*, 63(20), 1-31. DOI: 10.18637/jss.v063.i20
- Blangiardo, M. y Cameletti, M. (2015). *Spatial and Spatio-Temporal Bayesian Models with R-INLA*. Chichester: John Wiley & Sons.
- Castañeda, J. y Garrido, S. (2020). ¿Cómo entender los datos de defunciones por COVID-19 en México? Taller de Datos. Nexos. Recuperado de: <https://datos.nexos.com.mx/?p=1351>
- Comisión Económica para América Latina y el Caribe. (2020). *América Latina y el Caribe ante la pandemia del Covid-19: efectos económicos y sociales* (Informe especial). Recuperado de: <https://www.cepal.org/es/publicaciones/45337-america-latina-caribe-la-pandemia-covid-19-efectos-economicos-sociales>
- Gomez-Rubio, V. (2020). *Bayesian Inference with INLA*. New York: Chapman & Hall/CRC Press.
- Martínez, S. J. y Vargas, A. F. (2020). Covid-19: evolución y estimaciones de las curvas epidémicas. *Temas Estratégicos* (78). Recuperado de: <http://bibliodigitalibd.senado.gob.mx/handle/123456789/4877>
- Martino, S. y Rue, H. (2008). *Implementing Approximate Bayesian Inference using Integrated Nested Laplace Approximation: a manual for the inla program*. Recuperado de Statistical Laboratory, University of Cambridge.

- Martins, G., Simpson, D., Lindgren, F. y Rue, H. (2012). *Bayesian computation with INLA: new features. Computational Statistics & Data Analysis*, 67, 68-83. DOI: 10.1016/j.csda.2013.04.014
- Moraga, P. (2019). *Geospatial Health Data: Modeling and Visualization with R-INLA and Shiny*. EE. UU.: Chapman & Hall/CRC Biostatistics Series.
- Organización Panamericana de la Salud. (2020). *Plan estratégico de preparación y respuesta para la enfermedad por coronavirus 2019 (COVID-19). Pautas para la planificación operativa de la preparación y la respuesta de los países (Llamamiento de la OPS/OMS)*. Recuperado de: <https://www.paho.org/es/documentos/respuesta-al-brote-covid-19-region-americas>
- Partida, B. V. y García, G. V. (2018). *Proyecciones de la población de México y de las entidades federativas 2016-2050*. México: CONAPO. Recuperado de: https://www.gob.mx/cms/uploads/attachment/file/390959/Proyecciones_de_la_poblacion_de_Mexico_2016_2050_primera_parte.pdf
- R Core Team. (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Ramos, C. (2020). Covid-19: la nueva enfermedad causada por un coronavirus. *Salud Publica de México*, 62(2), 225-227. DOI: 10.21149/11276
- Riebler, A., Sørbye, S. H., Simpson, D. y Rue, H. (2016). An intuitive Bayesian spatial model for disease mapping that accounts for scaling. *Statistical Methods in Medical Research*, 25(4), 1145-1165. DOI: 10.1177/0962280216660421
- Rue, H., Martino, S. y Chopin, N. (2009). Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 71(2), 319-392. DOI: 10.1111/j.1467-9868.2008.00700.x
- Schrödle, B. y Held, L. (2011). Spatio-temporal disease mapping using INLA. *Environmetrics*, 22(6), 725-734. DOI: 10.1002/env.1065
- Secretaría de Salud. (2020). *Datos Abiertos Dirección General de Epidemiología*. Recuperado de Gobierno de México: <https://www.gob.mx/salud/documentos/datos-abiertos-152127>

- Sharafifi, Z., Asmarian, N., Hoorang, S. y Mousavi, A. (2018). Bayesian spatio-temporal analysis of stomach cancer incidence in Iran, 2003-2010. *Stoch Environ Res Risk Assess* (32), 2943-2950. DOI: 10.1007/s00477-018-1531-3
- Wang, X., Yue, Y. R., y Faraway, J. J. (2018). *Bayesian Regression Modeling with INLA*. Reino Unido: Chapman and Hall/CRC.