

Rastreo del número reproductivo efectivo (R_t) de la CoViD-19 en Colombia implementando el Filtro de Kalman

Tracking the effective reproductive number (R_t) of CoViD-19 in Colombia implementing the Kalman Filter

Jeyson David Rosales Mendoza



Alfredo Rafael Roa Narváez



Universidad del Atlántico, Colombia

Jenny María Peralta Polo



Universidad Nacional Abierta y a Distancia, Colombia

OPEN ACCESS

Recibido: 16/03/2022

Aceptado: 09/06/2022

Publicado: 26/09/2022

Correspondencia de autores:
jeyson-1097@hotmail.com



Copyright 2020
by Investigación e
Innovación en Ingenierías

Resumen

Objetivo: Implementar el Filtro de Kalman para estimar el número reproductivo efectivo R_t para la CoViD-19 en Colombia y algunos de sus departamentos y ciudades más afectadas, con base en las observaciones de infectados reportados por el Instituto Nacional de Salud. **Metodología:** El rastreo del R_t se inicia con la calibración y recalibración de la tasa de recuperación Y mediante la resolución de un problema de optimización. Luego, se utiliza una heurística para determinar las covarianzas de perturbaciones ruidosas presentes en el modelo del R_t . Seguido, se aplica el Filtro de Kalman para obtener realizaciones de la estimativa óptima del R_t . Finalmente, se implementa simulación Monte Carlo para obtener la trayectoria media del R_t óptimo con un intervalo de confianza del 95%. **Resultados:** Se realizó el rastreo del R_t de la CoViD-19 en los departamentos de Atlántico, Antioquia y Valle del Cauca, las ciudades de Barranquilla y Bogotá, y en general en Colombia, en los periodos de tiempo determinados por los picos de contagios que experimentó cada territorio en particular, mostrando coherencia con la realidad que se vivió en cada uno. **Conclusiones:** El método para obtener el R_t , usado en esta investigación, brinda una estimación de dicho número más cercana a la realidad al compararlo con el cálculo que se obtiene a partir del modelo epidemiológico SIR determinístico.

Palabras clave: Filtro de Kalman, número reproductivo efectivo, modelo epidemiológico SIR.

Abstract

Objective: To implement the Kalman Filter for estimating the effective reproductive number of CoViD-19 in Colombia and some of its most affected departments and cities, based on the observations of infected people reported by the National Institute of Health. **Methodology:** The tracking of R_t begins with the calibration and recalibration of the recovery rate Y by solving an optimization problem. Then, a heuristic is used to determine the covariances of noisy disturbances present in the model of R_t . Next, the Kalman Filter is applied to obtain realizations of the optimal estimate of R_t . Finally, Monte Carlo simulation is implemented to produce the mean trajectory of the optimal R_t with a confidence interval of 95%. **Results:** The tracking of R_t for the CoViD-19 was carried out in the departments of Atlántico, Antioquia and Valle del Cauca, the cities of Barranquilla and Bogotá, and in general in Colombia, in the periods of time determined by the peaks of infections that each particular territory experienced, showing coherence with the reality that each one suffered. **Conclusions:** The method to obtain the R_t that was used in this research provides an estimate of this number that is closer to reality compared to the calculation obtained from the deterministic SIR epidemiological model.

Keywords: Kalman filter, effective reproductive number, SIR epidemiological model.

Como citar (IEEE): J. Rosales-Mendoza., J. Peralta-Polo., y A. Roa-Narváez. "Rastreo del número reproductivo efectivo (R_t) de la CoViD-19 en Colombia implementando el Filtro de Kalman". Investigación e Innovación en Ingenierías, vol. 10, n°2, 62-77, 2022. DOI: <https://doi.org/10.17081/invinno.10.2.5638>

Introducción

Las enfermedades infecciosas han constituido una amenaza muy grave para la sociedad; el brote de CoViD-19, causada por el nuevo coronavirus SARS-CoV2, tuvo inicio en Wuhan, una ciudad de la provincia de Hubei, China, a finales de 2019, con una rápida propagación y un elevado número de infectados y muertes en todo el mundo. Ésta se convirtió en una pandemia que ha afectado a muchos países [1], entre éstos Colombia. Desde la aparición del primer caso de CoViD-19 en Colombia, el 6 de marzo de 2020, la velocidad de propagación del virus se incrementó de manera exponencial, ubicando al país entre los diez primeros en el mundo con mayor número de contagios durante algunos periodos de ese mismo año, superando el millón de contagios el 24 de octubre y el segundo millón solo tres meses después [2].

Dado esto, en el presente artículo se realiza el rastreo del número promedio de personas que son contagiadas por un individuo en su etapa infecciosa de CoViD-19 en Colombia y algunos de sus departamentos y ciudades más afectadas. Asimismo, consideramos que la importancia de este estudio consiste en proveer de información técnica sobre metodologías alternas a las que hasta el momento vienen usando los entes gubernamentales, en las cuales soportan los ajustes y políticas de salud pública, para un seguimiento y control más robusto de la pandemia en el país y sus regiones. El R_t , en particular, es un elemento fundamental en la toma de decisiones puesto que es a través de este valor que se mide el éxito de las políticas implementadas por el gobierno para disminuir la velocidad en los contagios, lo cual se ve reflejado cuando se logran valores menores a la unidad para dicha tasa [3]. Actualmente, el Instituto Nacional de Salud (INS) hace uso de un modelo Bayesiano para la estimación del número reproductivo de la CoViD-19 en Colombia de manera empírica, a priori, y sin soporte en modelo epidemiológico alguno. Dicho modelo se basa también en la estimación de casos sintomáticos diarios según la fecha de inicio de síntomas, para lo cual se debe contar con evidencia médica, de acuerdo con cada entorno específico y las normas sociales que lo rigen, que conste de información sobre el tiempo que transcurre desde la aparición de síntomas hasta el registro del individuo enfermo. Por tanto, podemos ver que en la práctica esta técnica está condicionada a información disponible del comportamiento biológico de la enfermedad, por ejemplo, tiempos de incubación e infección, que exigen en gran medida a la ya limitada capacidad logística del país.

Ahora bien, a diferencia del INS, en este trabajo el rastreo del número reproductivo efectivo R_t se realizará a través de la estimación del estado de un sistema estocástico derivado del modelo epidemiológico SIR, el cual ha sido aplicado en diversos escenarios de epidemias [4]. Este modelo clasifica, en cada día t , a las personas de una población específica de N individuos en tres categorías disjuntas, a saber: susceptibles de enfermar (S_t : Susceptibles), enfermos (I_t : Infectados) y en personas que superaron la enfermedad o murieron por la infección (R_t : Removidos). El modelo SIR está dado por las siguientes ecuaciones:

$$\begin{aligned} S_t &= S_{t-1} - \beta_t I_{t-1} \frac{S_{t-1}}{N}, \\ I_t &= I_{t-1} + \beta_t I_{t-1} \frac{S_{t-1}}{N} - \gamma I_{t-1}, \\ R_t &= R_{t-1} + \gamma I_{t-1}, \end{aligned} \quad (1)$$

siendo $\beta_t \geq 0$ la tasa de transmisión diaria y $0 < \gamma \leq 1$ la tasa de transición de infectado a removido. Partiendo de este modelo, en [5] se desarrolla el sistema dinámico estocástico para el R_t dado por:

$$\begin{aligned} \mathcal{R}_t &= \mathcal{R}_{t-1} + w_t, \\ y_t &= \gamma(\mathcal{R}_t - 1) + v_t, \end{aligned} \quad (2)$$

donde y_t son las observaciones de la tasa de crecimiento de los infectados activos que se obtienen de una cierta base de datos de infectados diarios, y $\{w_t\}_{t \geq 0}$ y $\{v_t\}_{t \geq 0}$ son procesos estocásticos no correlacionados, idénticamente distribuidos, con media cero y matrices de covarianza definidas positivas. En este trabajo estudiamos el comportamiento del \mathcal{R}_t a través del sistema dinámico estocástico en (2). Ocurre que, dadas las perturbaciones en la variable de estado, se hace necesario realizar una reconstrucción (estimación) de sus valores para cada instante de tiempo t basados en las observaciones ruidosas y_t . Como es sabido, el estimador óptimo para el proceso $\{\mathcal{R}_t\}_{t \geq 0}$, es el conocido Filtro de Kalman [6], el cual está dado también por un sistema dinámico estocástico. El Filtro de Kalman es óptimo en el sentido de que minimiza el error de estimación cuadrático medio y, además, coincide en media con el \mathcal{R}_t , para todo $t \geq 0$. De acuerdo con [6], la dinámica de la estimativa óptima $\hat{\mathcal{R}}_t$ del \mathcal{R}_t , dada por la esperanza condicional $E\{\mathcal{R}_t | Y_t\}$ con $Y_t = \{y_0, y_1, \dots, y_t\}$, está regida por las ecuaciones en diferencias:

$$\begin{aligned} \hat{\mathcal{R}}_t &= \hat{\mathcal{R}}_{t-1} + K_t(y_t - \gamma \hat{\mathcal{R}}_{t-1}) \\ K_t &= P_t^{t-1} \gamma [\gamma^2 P_t^{t-1} + \Sigma^v]^{-1}, \quad t \geq 1, \end{aligned} \quad (3)$$

donde,

$$P_{t+1}^t = P_t^{t-1} (1 - \gamma^2 P_t^{t-1} [\gamma^2 P_t^{t-1} + \Sigma^v]^{-1}) + \Sigma^w, \quad t \geq 0, \quad (4)$$

es la ecuación en diferencias de Riccati que describe la dinámica de la covarianza condicional P_t^{t-1} del error de estimación $\hat{\mathcal{R}}_t = \hat{\mathcal{R}}_t - \mathcal{R}_t$ dado el conjunto de observaciones Y_{t-1} (los valores P_0^{-1} y $\hat{\mathcal{R}}_0$ son la covarianza y la media de \mathcal{R}_0 , respectivamente); el valor K_t es conocido como la ganancia de Kalman y los valores Σ^w y Σ^v son las covarianzas de los procesos w_t y v_t , respectivamente.

Por otro lado, en aplicaciones reales es muy común el desconocimiento de parámetros presentes en las ecuaciones del filtro, lo cual obliga a que, de alguna manera, debamos estimar el valor de estos parámetros. Sin embargo, el rendimiento del filtro puede verse comprometido si no se lleva a cabo una estimación correcta, ocasionando eventualmente divergencia en las estimaciones o resultados con un alto grado de sesgo. El proceso de estimación de parámetros como los mencionados anteriormente se denomina diseño o ajuste del filtro, y se realiza utilizando únicamente las observaciones del modelo. Para el ajuste del filtro existen diversos enfoques basados en técnicas Bayesianas [7], de máxima verosimilitud [8], de coincidencia de covarianza [9], de correlación [10], entre otras. En este trabajo utilizamos el Método Recursivo de Referencia (RRR) presente en [11] para realizar el ajuste del filtro, calculando parámetros tales como: 1. Los valores iniciales $\hat{\mathcal{R}}_0$ y P_0^{-1} realizando simulaciones reiterativas del propio Filtro de Kalman y el suavizador Rauch-Tung-Striebel y, 2. Las covarianzas Σ^w y Σ^v utilizando el algoritmo de esperanza-maximización.

En resumen, haciendo uso de los reportes de contagios diarios provenientes del INS [2] en Colombia, calculamos las observaciones de la tasa de crecimiento de los infectados activos, estimamos los valores iniciales $\hat{\mathcal{R}}_0$ y P_0^{-1} y las covarianzas Σ^w y Σ^v a través del método RRR y, finalmente, implementamos el Filtro de Kalman para obtener las estimaciones óptimas del \mathcal{R}_t en Colombia, como también en las ciudades de Barranquilla y Bogotá D.C., y los departamentos de Atlántico, Antioquia y Valle del Cauca, en un cierto periodo de tiempo que obedece a los picos de contagios que experimentaron cada uno de estos territorios.

Metodología

Para realizar las simulaciones con base en la información ofrecida por el INS, clasificamos primeramente la población de cada territorio considerado de manera particular, en un periodo de tiempo establecido específicamente, en compartimientos de Susceptibles, Infectados y Removidos. Seguido de esto, usamos el software MATLAB [12] para realizar la estimación del \mathcal{R}_t llevando a cabo las siguientes etapas.

Calibración de la tasa de recuperación γ

Primero se calcula la tasa de recuperación óptima γ^* resolviendo el problema de optimización:

$$\begin{aligned} \text{Minimizar} \quad & \sum_{t=1}^{\tau} \|I_t^{\text{obs}} - I_t\| \\ \text{sujeto a:} \quad & 0 < \gamma \leq 1, \end{aligned} \quad (5)$$

donde I_t^{obs} son los infectados activos diarios que se toman de los registros del INS [2] para τ días en una fecha estipulada, y I_t son los infectados diarios generados por la ecuación recursiva,

$$I_{t+1} = (1 - \gamma)I_t + \text{nuevos casos}_t. \quad (6)$$

Para calcular γ^* , se implementa un método iterativo de optimización partiendo de un valor inicial γ_0 , el cual, a su vez, es obtenido por una técnica similar a la presentada en [13], también como solución de un problema de mínimos cuadrados asociado al modelo epidemiológico SIR que considera el parámetro $\beta > 0$ constante. El parámetro γ^* se recalcula para periodos de $\tau = 15$ días y a partir de éste se obtienen secuencias de valores I_t generadas por la ecuación recursiva en (6), con I_0 correspondiendo a los infectados activos de la fecha inicial considerada para cada territorio. Los periodos de tiempo considerados para Colombia, Atlántico, Antioquia, Valle del Cauca, Barranquilla y Bogotá no tienen relación entre sí; estos fueron determinados por los picos de contagios que experimentaron cada territorio en particular.

Ajuste del filtro: Método Recursivo de Referencia

El método RRR [11] es una heurística que realiza una estimación de los parámetros presentes en el modelo del Filtro de Kalman. En su primera iteración, se implementa el Filtro de Kalman dado por las ecuaciones en (3) para calcular $\hat{\mathcal{R}}_t$, K_t y P_{t+1}^t , $t = 0, \dots, \tau$, tomando valores iniciales arbitrarios de $\hat{\mathcal{R}}_0$, P_0^{-1} , Σ^w y Σ^v . Seguido, se estiman los valores $\hat{\mathcal{R}}_0$ y P_0^{-1} por $\hat{\mathcal{R}}_0^r$ y P_0^r , respectivamente, ejecutando el suavizador Rauch-Tung-Striebel (RTS) [14] a tiempo reverso para $t = \tau - 1, \dots, 0$:

$$\begin{aligned} K_t^r &= P_t^r [P_{t+1}^r]^{-1}, \\ \hat{\mathcal{R}}_t^r &= \hat{\mathcal{R}}_t + K_t^r (\hat{\mathcal{R}}_{t+1}^r - \hat{\mathcal{R}}_t), \\ P_t^r &= P_t^r + (K_t^r)^2 (P_{t+1}^r - P_{t+1}^t), \end{aligned} \quad (7)$$

con $P_t^r = (1 - \gamma K_t) P_t^{t-1}$. Los valores K_t^r , $\hat{\mathcal{R}}_t^r$ (definido este como la esperanza condicional $E\{\mathcal{R}_t | Y_t\}$) y P_t^r , para $t < \tau$, son conocidos como la ganancia suavizada, el estado suavizado y la covarianza del error de estimación suavizada, respectivamente. Ahora bien, para estimar las covarianzas Σ^w y Σ^v se utilizó el método de Esperanza-Maximización (EM) presente en [11], el cual utiliza las estimaciones generadas por el suavizador RTS. Así, del método EM se obtiene:

$$\Sigma^v = \frac{1}{\tau} \sum_{t=1}^{\tau} \{(y_t - \gamma \hat{\mathcal{R}}_t^r)^2 + \gamma^2 P_t^r\}, \quad (8)$$

y

$$\Sigma^w = \frac{1}{\tau} \sum_{t=1}^{\tau} \{\hat{w}_t^2 + P_t^\tau + P_{t-1}^\tau - 2P_{t,t-1}^\tau\}, \quad (9)$$

donde $\hat{w}_t = \hat{\mathcal{R}}_t^\tau - \hat{\mathcal{R}}_{t-1}^\tau$, y el valor $P_{t,t-1}^\tau$ es dado por:

$$P_{t,t-1}^\tau = P_t^\tau K_{t-1}^\tau + K_t^\tau (P_{t+1,t}^\tau - P_t^\tau) K_{t-1}^\tau, \quad (10)$$

con $P_{t,t-1}^\tau = (1 - \gamma K_t^\tau) P_{t-1}^{\tau-1}$. Por último, se actualizan las variables $\hat{\mathcal{R}}_t$, K_t y P_{t+1}^τ del modelo implementando nuevamente el Filtro de Kalman, con base en las últimas estimaciones obtenidas para los parámetros $\hat{\mathcal{R}}_0$, P_0^{-1} , Σ^w y Σ^v . Este proceso se hace repetidamente hasta alcanzar, después de unas pocas iteraciones (alrededor de 50 como reportado en [11, 15]), un equilibrio estadístico del filtro con consistencia verificada, en el sentido de que provee una cota de Cramer Rao [16] cuasi-óptima para los parámetros desconocidos y estimativas estables para éstos independientemente de los valores iniciales considerados.

Filtro de Kalman y simulación Monte Carlo

El ajuste del filtro realizado en la etapa anterior para la estimación óptima de los parámetros $\hat{\mathcal{R}}_0$, P_0^{-1} , Σ^w y Σ^v , eliminan cualquier consideración determinista que pueda generar sesgos y garantiza además que los resultados en las estimaciones del \mathcal{R}_t se obtienen por el rendimiento del filtro y no por el ajuste del filtro. En esta etapa se calcula el valor medio del número reproductivo efectivo \mathcal{R}_t usando simulación Monte Carlo con base en 10.000 realizaciones de la trayectoria de estado $\hat{\mathcal{R}}_t$ del Filtro de Kalman, con una región de confianza del 95%.

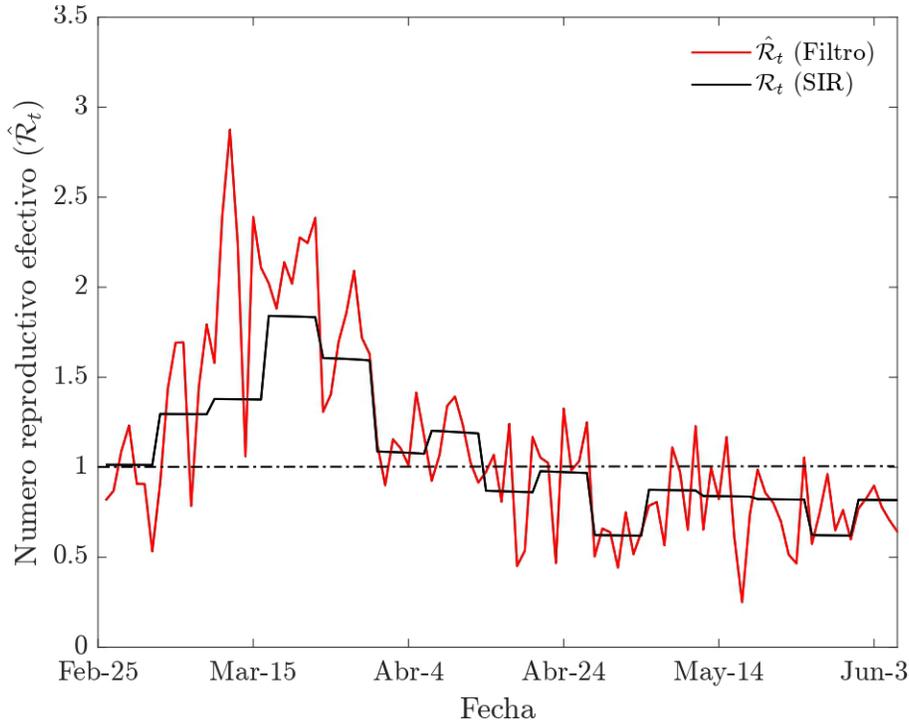
Algunas comparaciones

En lo que sigue, haremos comparaciones de algunos resultados para las estimaciones del \mathcal{R}_t usando la técnica desarrollada en este trabajo, con resultados que se obtienen al aplicar otras técnicas. Para ello, tomaremos el periodo comprendido entre el 25 de febrero de 2021 y el 8 de junio del mismo año en la ciudad de Barranquilla. Una primera comparación consiste en contrastar la estimación $\hat{\mathcal{R}}_t$ del \mathcal{R}_t con el cálculo que se hace de este número según la ecuación:

$$\mathcal{R}_t = \mathcal{R}_t^0 \frac{S_{t-1}}{N}, \quad t \geq 1, \quad (11)$$

donde $\mathcal{R}_t^0 = \beta_t / \gamma$ es el número reproductivo básico obtenido del modelo SIR determinístico. En este caso, los parámetros γ y β_t son calibrados a partir de la solución de un problema de optimización (en el que se considera β_t constante), tal y como se desarrolló en [17] para el cálculo del número reproductivo básico \mathcal{R}_t^0 asociado al modelo SIRD, con periodos de $\tau = 7$ días.

Figura 1. Comparación de las trayectorias del \mathcal{R}_t determinístico obtenido del modelo SIR simplificado y del $\hat{\mathcal{R}}_t$ obtenido del Filtro de Kalman, en la ciudad de Barranquilla del 25 de febrero al 8 de junio de 2021.

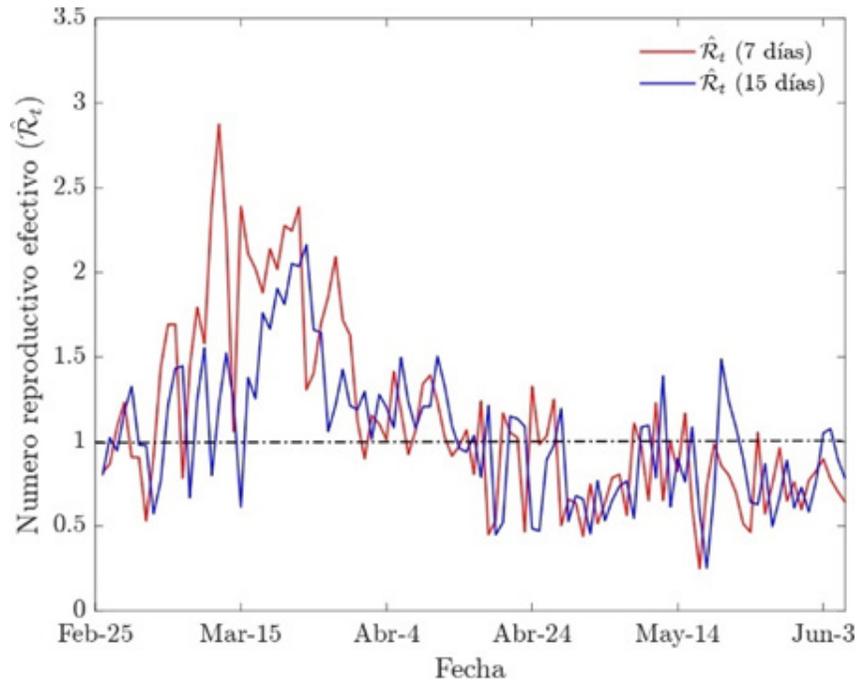


Fuente: Elaboración propia

Puede observarse que la estimación $\hat{\mathcal{R}}_t$ obtenida de la implementación del Filtro de Kalman, en contraste con el cálculo del número reproductivo efectivo de manera determinística, permite que haya una captura más realista del comportamiento día a día de dicho valor. De hecho, el \mathcal{R}_t determinístico presenta comportamientos “controlados” en los periodos de siete días considerados entre cada calibración de los parámetros β y γ (dinámica semiconstante), a diferencia de la estimación obtenida por el Filtro de Kalman, la cual refleja el comportamiento irregular del \mathcal{R}_t día a día, confirmando la influencia que ejerce la variación en el tiempo de la tasa de infección β_t , la cual, además, es considerada una variable de naturaleza estocástica. En todo caso, durante el periodo considerado, ambos cálculos permiten apreciar los momentos críticos en la dinámica del número reproductivo efectivo, dejando ver los periodos en los que el \mathcal{R}_t alcanza sus valores máximos y aquellos donde sus valores disminuyen, vea Figura 1.

En segundo lugar, se presenta la comparación de realizar la estimación del número reproductivo efectivo contrastando la calibración del parámetro γ sugerido para periodos de tiempo de 7 días como en [17, 18], con la evolución que tiene dicho número al actualizar la calibración de la tasa de recuperación cada 15 días.

Figura 2. Comparación entre $\hat{\mathcal{R}}_t$ con γ calibrado en periodos de 7 días y 15 días en Barranquilla del 25 de febrero al 8 de junio de 2021.

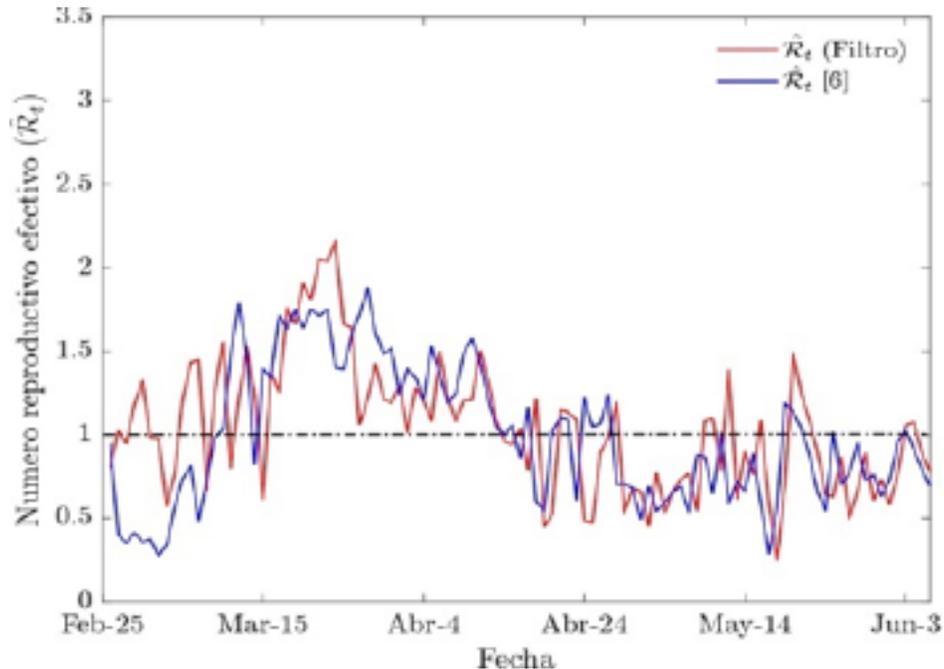


Fuente: Elaboración propia

Como podemos ver en la Figura 2, calibrar γ en periodos de 7 y 15 días no genera mucha discrepancia entre los resultados, por lo que resulta conveniente realizar la actualización de la calibración de γ de 15 en 15 días, aliviando así el gasto computacional para la implementación del filtro toda vez que se quiera hacer seguimiento del \mathcal{R}_t en un intervalo prolongado de tiempo; es por ello que en la metodología presentada anteriormente hemos dispuesto hacer la actualización del valor de γ cada 15 días.

Finalmente, se hace la comparación de la estimación del \mathcal{R}_t realizada en este trabajo con la estimación del \mathcal{R}_t procediendo como en [5], la cual es hecha con base en un valor constante para el parámetro γ . De hecho, como mencionado antes, nuestro modelo para el cálculo del \mathcal{R}_t es aquel extraído de [5]; sin embargo, en el referido artículo se toma como hipótesis que el intervalo serial (tiempo que transcurre entre el inicio de los síntomas en un caso y el inicio de los síntomas en sus casos secundarios) de la CoViD-19, coincidente con el recíproco de γ en el modelo SIR (vea [19]), es igual a 7 días, o mejor, $\gamma = 1/7$; ahora bien, el hecho de asumir un valor fijo para γ está soportado en evidencia médica disponible sobre el comportamiento de la enfermedad, a lo cual no siempre puede tenerse acceso debido a limitación de recursos logísticos.

Figura 3. Comparación del rastreo de las trayectorias del \mathcal{R}_t con $\gamma = 1/7$ (vea [5]) y γ calibrado cada 15 días, en Barranquilla del 25 de febrero al 8 de junio de 2021.



Fuente: Elaboración propia

De esta manera, en la Figura 3 puede verse una comparación del rastreo de la trayectoria del \mathcal{R}_t asumiendo el valor $\gamma = 1/7$ sugerido en [5], y el de la trayectoria obtenida con la metodología presente en este trabajo, tomando una calibración de γ en períodos de 15 días. Puede observarse que, a pesar de tener inicios algo distintos, los comportamientos de las trayectorias acaban siendo muy similares, mostrando esto que nuestra propuesta requiere de la menor información posible para obtener estimaciones satisfactorias del \mathcal{R}_t .

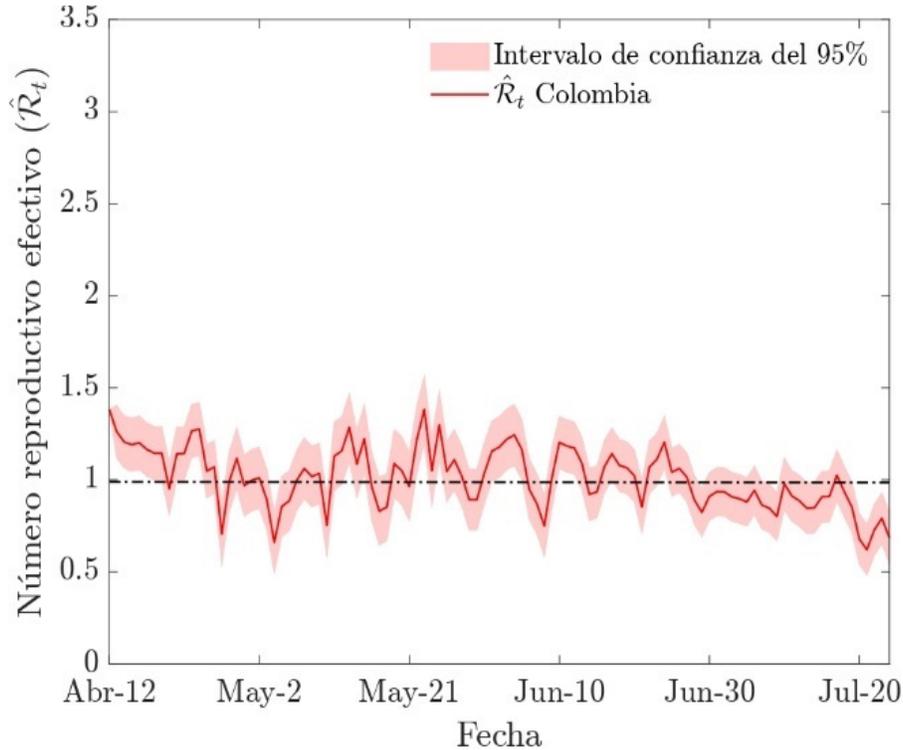
Resultados

A continuación, se presentan los resultados obtenidos al implementar el Filtro de Kalman para la estimación óptima del \mathcal{R}_t , con base en el modelo presentado en [5], en Colombia y algunos de sus territorios. Específicamente, se consideraron los contagios diarios reportados por el Instituto Nacional de la Salud [2] para Colombia, algunos de sus departamentos más afectados como Atlántico, Antioquia y Valle del Cauca, y las ciudades de Bogotá D.C. y Barranquilla.

Colombia

Para la estimación del \mathcal{R}_t de la CoViD-19 en Colombia se tomaron los datos partiendo desde el 12 de abril de 2021, fecha en la que los contagios en el país se mantenían entre los 15.000 a 17.000 contagios diarios, hasta finales del mes de mayo donde hubo un fuerte crecimiento de infección que se extendió hasta finales del mes de julio cuando empezó a descender con gran velocidad.

Figura 4. Rastreo del \mathcal{R}_t en Colombia en el periodo de 12 de abril a 25 de julio de 2021.



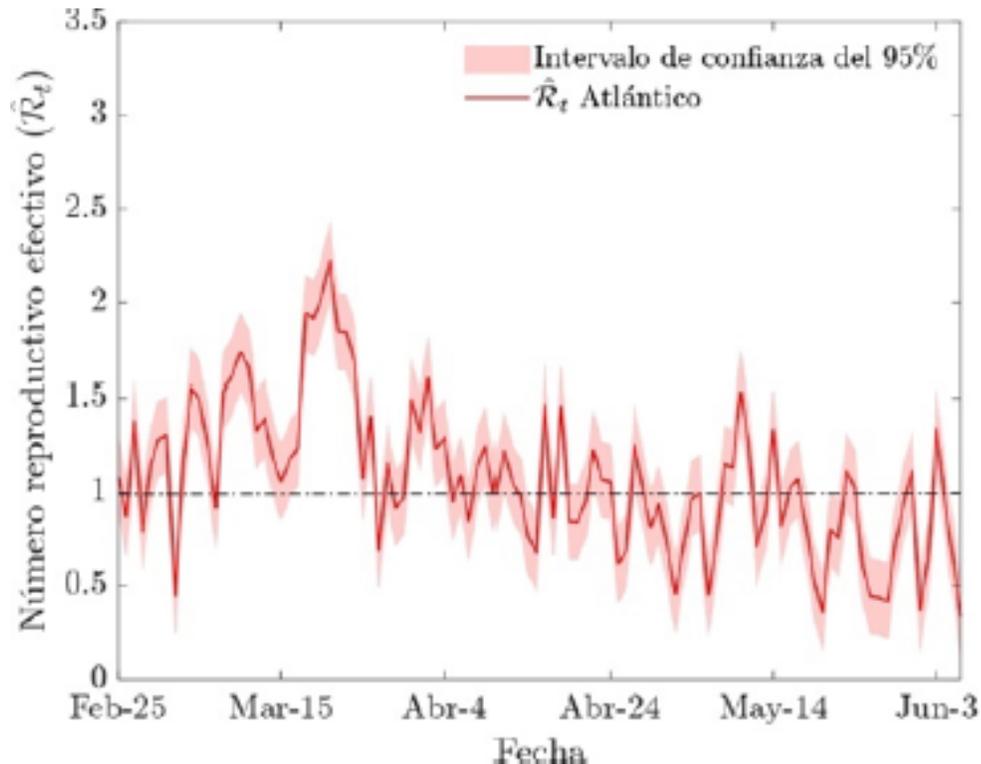
Fuente: Elaboración propia

En los resultados obtenidos podemos ver que, a pesar del crecimiento que se presentó a finales de mayo, la estimación del \mathcal{R}_t presentó un comportamiento controlado, sin picos que excedieran el valor de 1,5, con valores cercanos a 1. Más aún, a finales del mes de junio, con la disminución de los contagios y el aumento de la población inmune por vacunación o recuperación, podemos observar que inicia un decaimiento en la dinámica del \mathcal{R}_t , alcanzando valores por debajo de 1, reflejándose un cierto control de la pandemia a partir del mes de julio.

Atlántico

Para el departamento del Atlántico se consideró el periodo entre el 25 de febrero y el 8 de julio de 2021, teniendo en cuenta que la tercera ola de la pandemia que afectó al departamento se encuentra dentro de este intervalo. Así, podemos ver el comportamiento de las estimaciones del \mathcal{R}_t antes, durante y después del pico de contagio.

Figura 5. Rastreo del R_t en el departamento del Atlántico en el periodo de 25 de febrero a 8 de junio de 2021.



Fuente: Elaboración propia

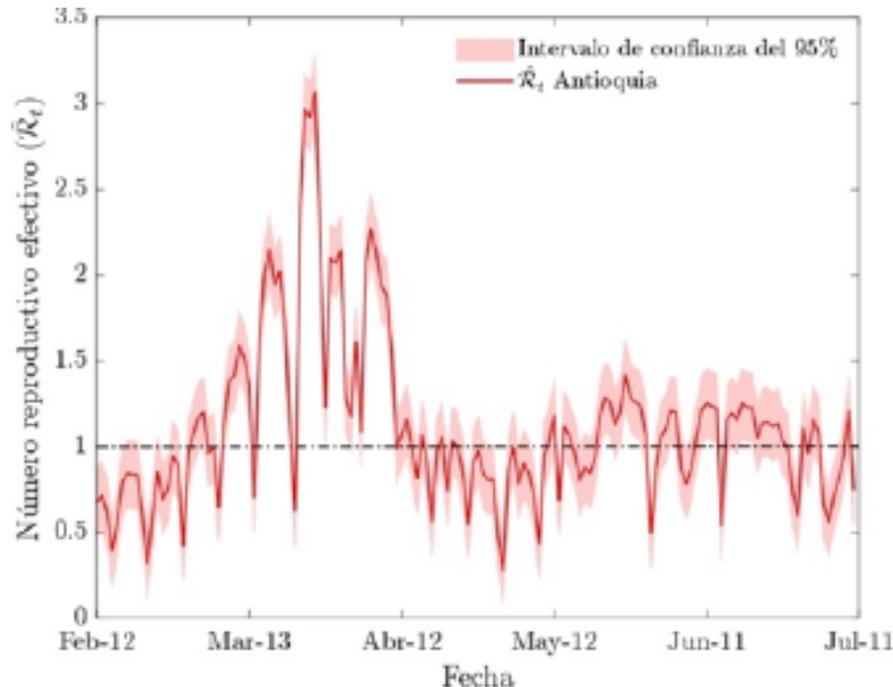
Al inicio del periodo en estudio podemos observar cómo el valor de las estimaciones del R_t se incrementa en sólo un mes, pasando de un valor por debajo de 1 a tomar valores cercanos a 2. Con esto podemos determinar que el periodo crítico que generó el incremento en el número de nuevos casos de infectados activos es el establecido entre los días finales del mes de febrero y los primeros días de marzo. Sin embargo, a pesar de que el departamento del Atlántico experimentó su pico más fuerte de contagios durante este periodo en lo que va de la pandemia, este no se prolongó más allá de un mes, lo cual puede validarse al ver cómo las estimaciones del R_t después de alcanzar el valor de 2, decaen y se mantienen la mayor parte del tiempo por debajo de 1, generando el control sanitario que ha vivido el departamento hasta la última fecha de estudio.

Antioquia

El departamento de Antioquia inició el año 2021 con un nivel bajo de contagios. Muestra de ello, para el seguimiento del R_t se tomó como fecha de partida el 12 de febrero, mes en el que las estimaciones del número reproductivo efectivo se mantuvieron controlado, con $R_t < 1$. Sin embargo, a inicios del mes de marzo el seguimiento del R_t se dispara, alcanzando valores cercanos a 3 en las estimaciones. Lo anterior refleja el incremento acelerado del número de nuevos casos diarios de contagios, el cual se había mantenido inferior a los 1.000 contagios diarios hasta el 15 de marzo y a partir de ese momento inició un incremento significativo generando una tercera ola que puso en crisis sanitaria al departamento. Además, podemos observar cómo las estimaciones del R_t descienden rápidamente en el mes de mayo, donde una persona infectada en promedio contagia a menos de una persona durante su periodo de infección. Lo anterior pudo deberse a la rápida toma de decisiones por parte de los entes de control y la implementación de la vacuna, lo cual también influyó claramente en evitar que el R_t se incrementara nuevamente.

A pesar de que las estimaciones del \mathcal{R}_t han estado por encima de 1 en algunos intervalos de días, éstas se mantuvieron siempre cercanas al valor de 1, provocando una disminución significativa en el número de nuevos contagios y así mismo, de los infectados activos.

Figura 6. Rastreo del \mathcal{R}_t en el departamento de Antioquia en el periodo de 12 de febrero a 11 de julio de 2021

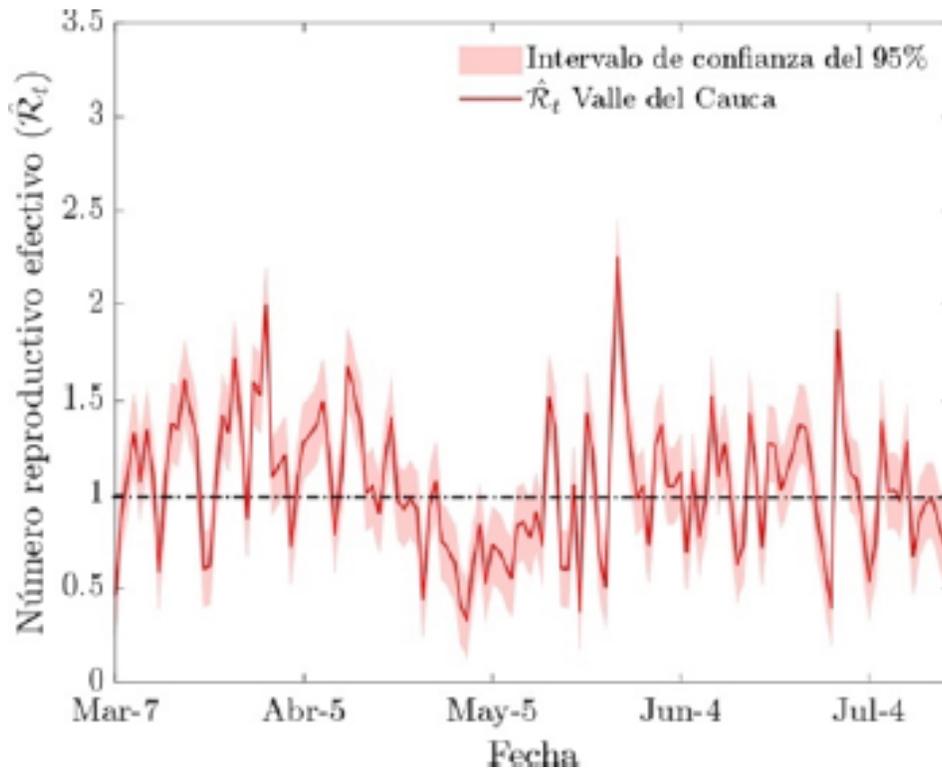


Fuente: Elaboración propia

Valle del Cauca

Desde el inicio de la pandemia por CoViD-19 el departamento del Valle del Cauca en ningún momento había reportado 2.000 o más nuevos casos de contagios en un día; sin embargo, a inicios del mes de mayo inicia la tercera ola de la pandemia, por lo que consideramos hacer el seguimiento del \mathcal{R}_t desde el 7 de marzo, mucho antes del pico, cuando los nuevos contagios rondaban entre 200 a 800 diarios, hasta el 19 de julio cuando el departamento empieza a superar el pico e inicia su descenso. Podemos observar un pequeño aumento en los valores del \mathcal{R}_t en el mes de marzo y una posterior caída de éstos hasta finales del mes de abril cuando inicia un nuevo pico, alcanzando un valor máximo en el mes de mayo con valores por encima de 2 y manteniéndose inestable hasta la fecha final en que se hizo el reporte. Aglomeraciones públicas y la falta de medidas de bio-seguridad han podido ser eventualmente los factores que generaron este tipo de comportamientos en la dinámica del \mathcal{R}_t .

Figura 7. Rastreo del \mathcal{R}_t en el departamento del Valle del Cauca en el periodo de 7 de marzo a 19 de julio de 2021.

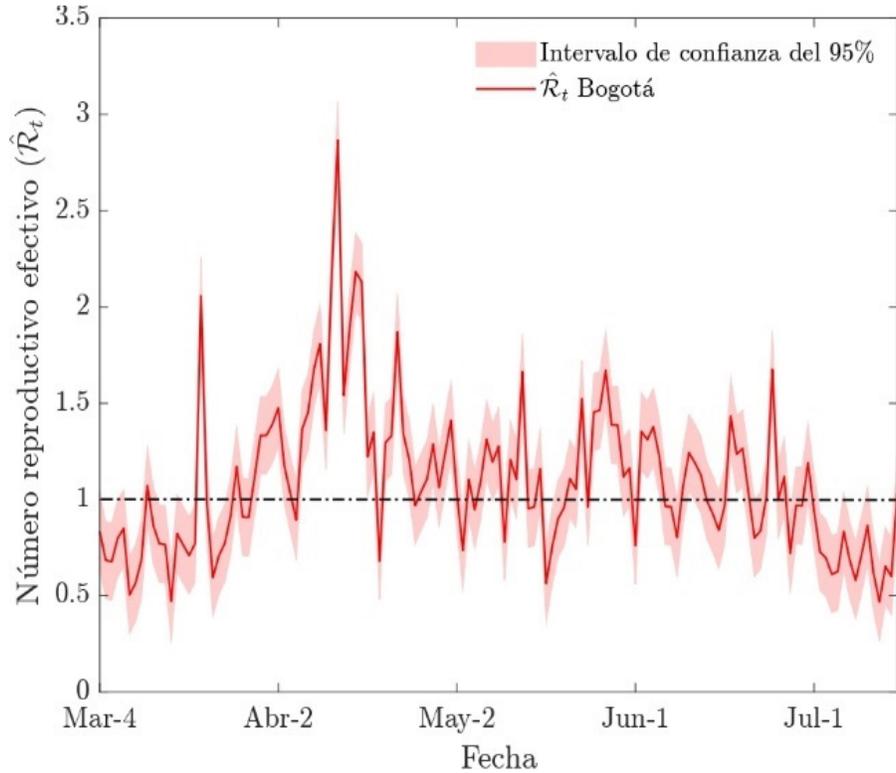


Fuente: Elaboración propia

Bogotá D.C.

En Bogotá D.C. se realizó el seguimiento del \mathcal{R}_t a partir del 4 de marzo de 2021, momento en el que la ciudad se encontraba en con una relativa estabilidad en número de nuevos contagios, con registros inferiores a los 1.000 contagios diarios. Sin embargo, podemos ver cómo al finalizar el mes empezó a ascender el $\hat{\mathcal{R}}_t$, alcanzando un valor máximo cercano a 3 en la primera mitad del mes de abril. Es decir, luego después de iniciar el mes de abril una persona infectada por CoViD-19 llegó a contagiar en promedio entre 2 a 3 personas, durante su periodo infeccioso, en la ciudad de Bogotá. Esto deja en evidencia la realidad que vivió la ciudad, pues fue en este mes que empezó a ascender el grupo de infectados activos alcanzando el mayor pico registrado en lo que va de pandemia. Además, podemos observar cómo el valor de las estimaciones del \mathcal{R}_t se mantiene inestable, con sobresaltos hasta mediados del mes de junio, momento en el que la ciudad veía prolongado su peor momento durante la pandemia. Sin embargo, desde finales del mes de junio hasta el último día considerado en este estudio para la ciudad (16 de julio de 2021), vemos que las estimaciones del número reproductivo efectivo logran mantenerse por debajo de 1, generando una caída rápida del pico de contagio, y en consecuencia registrando a partir de ese momento un menor número de nuevos contagios cada día.

Figura 8. Rastreo del \mathcal{R}_t en la ciudad de Bogotá en el periodo de 4 de marzo a 16 de julio de 2021.

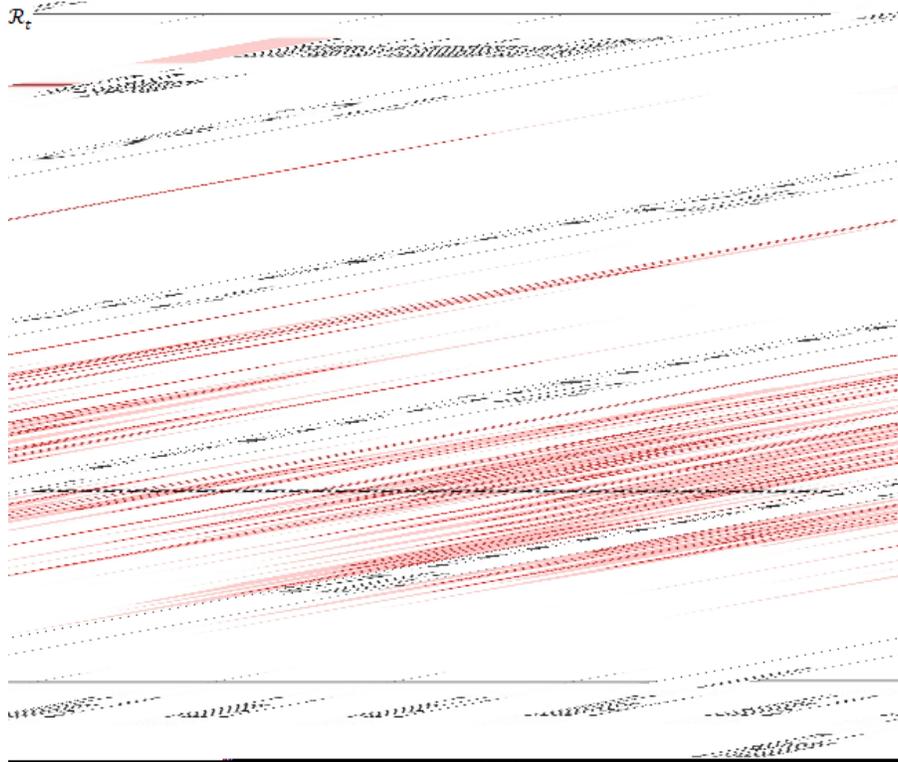


Fuente: Elaboración propia

Barranquilla

Para la ciudad de Barranquilla, capital del departamento del Atlántico, se consideró el mismo periodo que dicho departamento. Teniendo en cuenta que nuestra consideración para tomar los intervalos de tiempo es el pico de contagio, pudimos observar claramente que ambos territorios experimentaron su tercer pico en fechas similares. Además, esto nos permite comparar el comportamiento de las estimaciones del \mathcal{R}_t del departamento con el de su capital. Una vez hecha la simulación para el rastreo del \mathcal{R}_t en la ciudad de Barranquilla, vimos que este experimentó un comportamiento similar al del departamento del Atlántico, presentando un ascenso rápido en menos de un mes, de manera que se alcanzaron valores por encima de 2; como también una caída rápida, compartiendo el patrón de mantenerse por debajo de 1 en la mayoría del tiempo transcurrido desde el mes de abril hasta el último día que se incluyó en este estudio. La ciudad de Barranquilla ha tenido un mantenerse por debajo de 1 en la mayoría del tiempo transcurrido desde el mes de abril hasta el último día que se incluyó en este estudio. La ciudad de Barranquilla ha tenido un mayor control de la pandemia en el segundo trimestre de 2021 en comparación con la ciudad de Bogotá, al mantener el número reproductivo efectivo por debajo de 1 en intervalos prolongados de tiempo y con sobresaltos no mayores a 1,5.

Figura 9. Rastreo del \mathcal{R}_t en la ciudad de Barranquilla en el periodo de 25 de febrero a 8 de junio de 2021.



Fuente: Elaboración propia

Conclusiones

El rastreo del número reproductivo efectivo para la CoViD-19 hecho en este trabajo es un ejercicio de gran relevancia ya que la técnica utilizada para tal fin puede convertirse en un instrumento cuyos resultados servirían de soporte en la toma de decisiones de un gobierno para hacer seguimiento y control a la evolución de una pandemia en tiempo real. El método de estimación del \mathcal{R}_t aquí presentado brinda una estimación de tal valor más cercana a la realidad ya que la dinámica del mismo es representada por un modelo matemático que está dado por un sistema dinámico estocástico, bajo la consideración de que en este tipo de estudio se debe tener en cuenta el comportamiento social de una población. Calcular la tasa \mathcal{Y} considerando el parámetro $\beta_t > 0$ como función del tiempo, permite obtener una estimación de los infectados activos más acertada y cercana a la realidad de su comportamiento. Además, calcularla a través de la solución de un problema de optimización permite mantener un abordaje matemático y no tener que recurrir a consideraciones sobre datos de observaciones médicas como se ha reportado en algunos trabajos que abordan esta problemática. Sin embargo, se debe tener en cuenta que no es recomendable mantener un mismo valor \mathcal{Y} para periodos grandes, vea [18]. De esta manera, hemos trabajado con actualizaciones del valor de \mathcal{Y} entre periodos que comprenden intervalos de tiempo de 15 días, para los cuales las estimativas presentaron un buen comportamiento con relación a periodos más cortos para la actualización de este parámetro.

Por otro lado, en cuanto al desempeño del Filtro de Kalman y garantizar un buen rendimiento de éste evitando que los resultados obtenidos estén contaminados por decisiones arbitrarias sin ningún tipo de justificación, hemos realizado el ajuste de los parámetros a través del Método Recursivo de Referencia (RRR), el cual es una heurística reportada recientemente en la literatura que realiza una estimación de

tales parámetros utilizando el propio Filtro de Kalman. Para este método fue mostrado en [11] que, sin importar el valor inicial que se considere para los parámetros y a pesar de no ser propiamente un método de optimización, se obtienen estimaciones óptimas y estables para los parámetros.

Por último, pudimos observar también que cuando las estimaciones del número reproductivo efectivo \mathcal{R}_t incrementan, tomando valores superiores a 2 o se mantiene por encima de 1,5 en largos periodos de tiempo, esto afecta el número de nuevos casos positivos diarios en los días siguientes, conllevando a un incremento significativo de estos casos y, consecuentemente, incrementándose también el grupo de infectados activos. Esto es una observación importante porque permite a los entes de control conocer el estado de la pandemia en una determinada región en tiempo real para así poder decisiones anticipadas y lidiar de mejor manera con la eventual ocurrencia de un pico de contagio. En Colombia en general, podemos decir que la dinámica del \mathcal{R}_t es menos caótica con relación a las estimaciones en sus departamentos y/o ciudades, presentando menos variaciones en su valor, manteniéndose siempre cerca de 1. A pesar de que las fechas en las que se realizó la estimación muchos de sus departamentos y/o ciudades vivían altos números de contagios, el valor del \mathcal{R}_t en Colombia no se disparó, esto es, aun cuando el número de infectados y de nuevos casos es alto, pudo notarse que su impacto en la población global del país es proporcionalmente bajo en comparación a los departamentos y ciudades estudiadas [20,21].

Referencias bibliográficas

1. Organización Mundial de la Salud, «Organización Mundial de la Salud,» 2020. [En línea]. Available: https://www.who.int/es/health-topics/coronavirus#tab=tab_1. [Último acceso: Julio 2021].
2. Instituto Nacional de la Salud, «Coronavirus,» Ministerio de Salud, 2020. [En línea]. Available: <https://www.ins.gov.co/Noticias/paginas/coronavirus.aspx>. [Último acceso: Julio 2021].
3. G. N. Páez Salamanca, C. F. Restrepo, J. F. Zea, É. C. Rivera, A. Quiroz Salazar, J. F. Cerón, S. E. Cortés, G. Vargas Sandoval y C. Castañeda Orjuela, «Estimación del número reproductivo efectivo de la transmisión del COVID-19 en Colombia,» *Observatorio Nacional de Salud*, 2021.
4. W. Kermack y A. G. Mackendrick, «A contribution to the mathematical theory of epidemics,» *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, vol. 115, nº 772, pp. 700-721, 1927.
5. F. Arroyo Marioli, F. Bullano, S. Kucinskis y C. Rondón-Moreno, «Tracking R of COVID-19: A new real-time estimation using the Kalman filter,» *PLoS one*, vol. 16, nº 1, p. e0244474, 2021.
6. R. E. Kalman, «A new approach to linear filtering and prediction problems,» 1960.
7. D. Alspach, «A parallel filtering algorithm for linear systems with unknown time varying noise statistics,» *IEEE Trans. Auto. Cont.*, vol. 19, nº 5, pp. 552-556, 1974.
8. V. A. Bavdekar, A. P. Deshpande y S. C. Patwardhan, «Identification of process and measurement noise covariance for state and parameter estimation using extended Kalman filter,» *Journal of Process control*, vol. 21, nº 4, pp. 585-601, 2011.
9. Mohamed, A. H. y Schwarz, K. P., «Adaptive Kalman Filtering for INS/GPS,» *Journal of Geodesy*, vol. 73, nº 4, pp. 193-203, 1999.
10. R. K. Mehra, «On the identification of variances and adaptive Kalman filtering,» *IEEE Trans. Automat.*, vol. 15, nº 2, pp. 175-184, 1970.
11. M. a. N. N. a. G. R. a. o. Ananthasayanam, «A Heuristic Reference Recursive Recipe for the Menacing Problem of Adaptively Tuning the Kalman Filter Statistics. Part-1. Formulation and Simulation Studies,» *arXiv preprint arXiv:1505.07201*, 2015.

12. MathWorks, «Matlab the language of technical computing (getting started with matlab) version 7,» Natick, MA, USA, 2005.
13. J. Peralta Polo, M. A. Caro Candezano y L. Niebles Nuñez, «Dos enfoques matemáticos epidemiológicos para modelar el comportamiento de los decesos causados por el COVID-19.,» *Investigación e Innovación en Ingenierías*, vol. 8, nº 2, 2020.
14. E. Rauch Herbert, F. Tung y C. T. Striebel, «Maximum likelihood estimates of linear dynamic systems.,» *AIAA journal*, vol. 3, nº 8, pp. 1445-1450, 1965.
15. M. Ananthasayanam y R. Mudambi , «A reference recursive recipe for tuning the statistics of the Kalman filter,» *Kalman Filters-Theory for Advanced Applications*, 2018.
16. Shao, «Mathematical Statistics.,» *New York: Springer. ISBN 0-387-98674-X.*, vol. Section 3.1.3., 1998.
17. J. P. Polo, «Estudo da CoViD-19 em algumas cidades da Colômbia.,» de 2º *Encontro Fluminense de Mulheres em Biomatemática.*, Rio de Janeiro, RJ, Brasil, 2020.
18. C. E. Rodrigues , P. Natti, N. Romeiro, M. Candezano y J. Polo, «One study of covid-19 spreading at the United States - Brazil – Colombia.,» *Trends in computational and Applied Mathematics TEMA*, vol. 22, nº 3, p. 435–452, 2021.
19. M. Junling, «Estimating epidemic exponential growth rate and basic reproduction number,» *Infectious Disease Modelling*, nº 5, pp. 129-141, 2020.
20. [20] G. E. Chanchí Golondrino, W. Y. Campo Muñoz, y L. M. Sierra Martinez, «Aplicación de la regresión polinomial para la caracterización de la curva del COVID-19, mediante técnicas de machine learning», *Investigación e Innovación en Ingenierías*, vol. 8, nº 2, pp. 87 - 105, jul. 2020.
21. [21] E. J. De la Hoz Domínguez, T. J. Fontalvo Herrera, y A. A. Mendoza Mendoza, "Aprendizaje automático y PYMES: Oportunidades para el mejoramiento del proceso de toma de decisiones", *Investigación e Innovación en Ingenierías*, vol. 8, nº 1, pp. 21-36, 2020. DOI: <https://doi.org/10.17081/invinno.8.1.3506>