



Revista Politécnica

ISSN: 1390-0129

ISSN: 2477-8990

Escuela Politécnica Nacional

Castro, Paula; De los Reyes, Juan-Carlos; González-Andrade, Sergio; Merino, Pedro  
Estimación de Parámetros para un Modelo del SARS-CoV-2 en Ecuador en Presencia de Incertidumbre  
Revista Politécnica, vol. 47, núm. 1, 2021, Febrero-Abril, pp. 7-16  
Escuela Politécnica Nacional

DOI: <https://doi.org/10.33333/rp.v47n1.01>

Disponible en: <https://www.redalyc.org/articulo.oa?id=688772208001>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en redalyc.org



Sistema de Información Científica Redalyc  
Red de Revistas Científicas de América Latina y el Caribe, España y Portugal  
Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso  
abierto

# Estimación de Parámetros para un Modelo del SARS-CoV-2 en Ecuador en Presencia de Incertidumbre

Castro, Paula <sup>1,\*</sup>  ; De los Reyes, Juan Carlos <sup>1</sup>  ; González-Andrade, Sergio <sup>1</sup>  ; Merino, Pedro <sup>1</sup> 

<sup>1</sup>Escuela Politécnica Nacional, Centro de Modelización Matemática en Áreas Clave para el Desarrollo, Quito, Ecuador

**Resumen:** En este artículo presentamos el modelo matemático de tipo compartimental con individuos asintomáticos utilizado por el Centro de Modelización Matemática (MODEMAT) para estudiar la propagación del SARS-CoV-2 en Ecuador, así como el esquema variacional de tipo bayesiano desarrollado para estimar los diferentes parámetros del modelo, en presencia de incertidumbre de los datos observados. Esta estimación permite realizar actualizaciones periódicas del número efectivo de reproducción, así como proyecciones a corto plazo, mediante métodos de ensamble de la incidencia de la epidemia.

**Palabras claves:** Estimación de parámetros, SARS-CoV-2, Problemas inversos, Asimilación de datos, Enfoque bayesiano.

## Parameter Estimation for SARS-CoV-2 Model in Ecuador in Presence of Uncertainty

**Abstract:** In this article we present the mathematical compartmental SARS-CoV-2 model, with asymptomatic patients, used by the Research Center on Mathematical Modelling (MODEMAT) to study the spread of the virus in Ecuador, as well as the developed Bayesian variational approach to estimate the model parameters, under uncertainty of the observed data. Thanks to this framework, we are able to estimate periodically the effective reproduction number and carry out short-term forecasts, using an ensemble method, of the incidence of the epidemic.

**Keywords:** Parameter estimation, SARS-CoV-2, Inverse problems, Data assimilation, Bayesian approach

### 1. INTRODUCCIÓN

Uno de los objetivos principales del Centro de Modelización Matemática (MODEMAT), en el contexto de la pandemia del coronavirus, consiste en monitorear, a través del uso de modelos matemáticos, la expansión del virus SARS-CoV-2, y los casos de la enfermedad COVID-19 provocada por este patógeno en Ecuador. Para esto, utilizamos un modelo epidemiológico compartimental que incluye en su dinámica a los individuos infecciosos asintomáticos (ver detalles en Li et al (2020); Wu et al (2020); Verity et al (2020)), y estimamos periódicamente los diferentes parámetros del mismo, así como el número de reproducción efectivo de la epidemia.

La principal fuente de información para realizar dichas estimaciones son las estadísticas oficiales de casos confirmados reportados por el *Ministerio de Salud Pública de Ecuador*, las mismas que presentan un alto grado de incertidumbre, no sólo por el margen de error de las pruebas de laboratorio o la limitación de la cantidad de tests que se realizan, sino principalmente por los constantes cambios metodológicos (no siempre coordinados)

que han implementado los diversos organismos oficiales para la recolección y el reporte de los casos.

En un escenario simplificado, el proceso de estimación de parámetros se realiza mediante técnicas de ajuste, las cuales consisten en construir una función matemática que se ajusta mejor a una serie particular de datos. En este contexto, el mejor ajuste implica que la distancia entre la curva y los datos, ubicados como puntos en un espacio predeterminado, debe ser minimizada, alcanzando un óptimo aproximado.

De manera general, la técnica de mínimos cuadrados no lineales consiste en identificar los parámetros que minimizan la suma de las diferencias al cuadrado entre los datos observados y el modelo no lineal utilizado. Así, supongamos que los parámetros identificados están dados por  $\Theta := (\theta_1, \theta_2, \dots, \theta_m)$  y los datos observados por  $(y_1, y_2, \dots, y_n)$ , los cuales representan una serie de tiempo en los pasos temporales  $t_i$  en los cuales se realizaron las observaciones. Además, supongamos que la solución del modelo está dada por una función no lineal genérica  $f(t_i, \Theta)$ . El problema, a grandes rasgos, consiste entonces en hallar el conjunto de parámetros  $\hat{\Theta}$  dado por:

\*paula.castro@epn.edu.ec

Recibido: 28/08/2020

Aceptado: 06/01/2021

Publicado: 28/02/2021

10.33333/tp.vol47n1.01

CC BY 4.0

$$\hat{\Theta} := \operatorname{argmin} \sum_{i=1}^n (f(t_i, \Theta) - y_i)^2,$$

donde  $n$  es el número de datos disponibles para hacer la inferencia. Hallada la solución de este problema, podemos considerar que la solución  $f(t_i, \hat{\Theta})$  brinda el modelo mejor ajustado a la serie temporal  $t_i$ , de acuerdo al criterio de mínimos cuadrados (ver, por ejemplo, Chowell (2007)).

Sin embargo, una característica central en la modelización de epidemias y transmisión de enfermedades es que éstos son procesos que obedecen a una cierta dinámica epidemiológica de propagación. El contagio de una persona susceptible se entiende como un encuentro aleatorio con una persona infecciosa, sujeto a varias contingencias, como por ejemplo, la duración o la intensidad de dicho encuentro. De igual manera, las personas ya contagiadas pasan por un periodo de incubación del virus, para luego formar parte de los compartimentos de individuos infecciosos. Todo esto se modeliza a través de un sistema de ecuaciones diferenciales, que debe ser tomado en cuenta a la hora de ajustar datos observados. Así, la principal debilidad del esquema de ajuste de curvas descrito anteriormente es que no necesariamente toma en cuenta dicha dinámica de manera explícita.

Más aún, un proceso de ajuste de mínimos cuadrados de los datos observados puede resultar inadecuado en presencia de incertidumbre en las observaciones. Como se mencionó previamente, en el caso ecuatoriano dicha incertidumbre es sumamente alta en comparación con las estadísticas recogidas por agencias gubernamentales de otros países. Consecuentemente, un proceso de ajuste y estimación de parámetros en el caso ecuatoriano debe tomar en cuenta la incertidumbre como un elemento constitutivo del problema a resolverse.

En ese sentido, la propuesta que presentamos en este artículo consiste en considerar a los parámetros y la solución del modelo epidemiológico del SARS-CoV-2 como variables aleatorias con una cierta distribución de probabilidad y no como valores constantes. Con este cambio en la hipótesis de trabajo, podemos adoptar el paradigma Bayesiano para problemas inversos y estimar los parámetros de manera robusta, buscando minimizar la probabilidad condicional *a-posteriori* de los parámetros, dadas las observaciones.

## 2. ESTIMACIÓN DE PARÁMETROS EN PRESENCIA DE INCERTIDUMBRE

Como mencionamos previamente, una de las principales limitantes de los métodos de ajuste es que no consideran en su construcción la incertidumbre en los datos a ser ajustados, lo cual puede tener efectos importantes en situaciones como la que se presenta actualmente en Ecuador, en relación a la evolución de la COVID-19. Para abordar tal dificultad, optamos por el paradigma bayesiano, con el cual podemos estudiar el problema de estimación de parámetros considerando su carácter intrínsecamente aleatorio.

### 2.1 Estimador de probabilidad máxima a posteriori

En la inferencia bayesiana, a diferencia de la estimación de máxima verosimilitud, se supone que los parámetros (coeficientes, condición inicial, etc.) son una variable aleatoria  $x$  con realizaciones en el espacio multivariable  $\mathbb{R}^m$ , de la cual conocemos, *a priori*, su función de densidad de probabilidad  $p(x)$ . Además, se asume que conocemos la función de densidad de probabilidad condicional (c.p.d.f. por sus siglas en inglés) de las observaciones  $y$ , considerando un conjunto de parámetros  $x$  dado. A esta función de densidad de probabilidad la denotamos por  $p(y|x)$ .

La idea básica de este enfoque consiste en combinar la información de las dos funciones de densidad de probabilidad descritas, utilizando la fórmula de Bayes, con el objetivo de obtener una probabilidad *a posteriori* de los eventos. En este caso, se supone que tanto las observaciones, como la información de fondo (estimaciones previas de los parámetros), tienen errores distribuidos normalmente con matrices de covarianza  $\mathbf{R}$  y  $\mathbf{B}$ , respectivamente. Por lo tanto, la función de densidad de probabilidad condicional de las observaciones viene dada por:

$$p(y|x) = \frac{1}{(2\pi)^{n/2} |\mathbf{R}|^{1/2}} \exp \left\{ -\frac{1}{2} (y - \mathcal{H}(x))^T \mathbf{R}^{-1} (y - \mathcal{H}(x)) \right\}, \quad (1)$$

y la función de densidad de probabilidad de los parámetros por:

$$p(x) = \frac{1}{(2\pi)^{m/2} |\mathbf{B}|^{1/2}} \exp \left( -\frac{1}{2} (x - x^b)^T \mathbf{B}^{-1} (x - x^b) \right), \quad (2)$$

donde  $n$  es el número de datos disponibles para hacer la inferencia,  $m$  es el número de parámetros a estimar,  $x^b$  representa el conocimiento *a priori* o histórico que se tiene del fenómeno, y  $\mathcal{H}(x)$  es un operador que liga las características del fenómeno, en particular los parámetros del modelo, con las observaciones obtenidas del mismo. En general, este es un operador no lineal, el cual está relacionado con el fenómeno observado.

Usando el Teorema de Bayes se obtiene que la distribución de probabilidad *a posteriori* de un evento  $x$ , tomando en cuenta las observaciones  $y$ , está dada por:

$$p(x|y) = \frac{p(y|x)p(x)}{p(y)}. \quad (3)$$

Ahora, dado que nuestro objetivo es hallar el mejor “estimador Bayesiano” utilizando la aproximación del máximo a posteriori, se busca maximizar la distribución de probabilidad *a posteriori* dada por (3). Puesto que  $p(y)$  es la distribución de probabilidad marginal y no depende de  $x$ , obtenemos el siguiente problema de optimización a resolverse:

$$\max_x p(y|x)p(x). \quad (4)$$

Aprovechando el hecho de que la función logarítmica  $\mathbb{R} \ni t \mapsto \ln(t)$  es monótonamente creciente y continua, se tiene que la solución del problema (4) es equivalente a maximizar el logaritmo

natural de la función objetivo de este problema. Esto es, resolvemos el problema

$$\max_x \ln(p(y|x)) + \ln(p(x)).$$

Finalmente, utilizando las funciones de densidad de probabilidad (1) y (2), obtenemos el problema de minimización equivalente

$$\min_x \frac{1}{2}(y - \mathcal{H}(x))^T \mathbf{R}^{-1}(y - \mathcal{H}(x)) + \frac{1}{2}(x - x^b)^T \mathbf{B}^{-1}(x - x^b), \quad (5)$$

el cual constituye un balance robusto entre la información de fondo y los datos observados.

Las matrices  $\mathbf{B}$  y  $\mathbf{R}$  corresponden a las matrices de covarianza de los errores de fondo y de las observaciones, respectivamente. Las matrices de covarianza de los errores se utilizan para describir el impacto de los errores que se acumulan en el proceso de recolección de información. De forma más precisa, la matriz  $\mathbf{R}$  de covarianza de errores de observación describe la desviación de los eventos observados de los eventos esperados. En palabras simples, esta matriz mide la validez de las mediciones frente a la evolución del fenómeno. La matriz  $\mathbf{B}$ , por otro lado, describe la diferencia entre la estimación previa de los parámetros y los parámetros “verdaderos” del fenómeno. En el caso de la evolución de la COVID-19, esta matriz describe cuán confiable es el conocimiento que se tiene acerca de los parámetros de propagación del virus en el momento de iniciar el proceso de estimación.

El problema de optimización (5) guarda estrecha relación con la estimación de máxima verosimilitud de los parámetros del sistema, obtenida a través de *filtros de Kalman* clásicos. En ambos casos, las observaciones son consideradas en un solo instante de tiempo, las cuales son corregidas a través de lo que se conoce como interpolación óptima o, equivalentemente, la resolución del problema variacional (5).

## 2.2 Esquema variacional en el tiempo

En los casos en que las observaciones son tomadas en distintos instantes del tiempo, el esquema bayesiano descrito anteriormente necesita ser adaptado para dar cuenta de la evolución de un sistema, así como de la variación de la incertidumbre. Este tipo de problemas ha sido abordado desde los dos enfoques comentados previamente: filtros de Kalman y métodos variacionales (ver, por ejemplo, Asch et al (2016)). En el primero de ellos, se adapta la construcción del filtro a un problema que varía en el tiempo, obteniéndose los denominados *filtros de Kalman extendidos*, los cuales requieren la solución de las ecuaciones diferenciales de Riccati. En el caso variacional, en cambio, el problema es formulado como uno de *control óptimo* y su solución requiere generalmente del *método adjunto*.

En nuestro caso, el problema de propagación del SARS-CoV-2 está asociado a la solución del sistema de ecuaciones diferenciales que determinan la evolución de los compartimentos poblacionales

utilizados en el modelo. En tal razón, consideramos, de manera general, un sistema de ecuaciones diferenciales del tipo

$$\frac{d\mathbf{u}}{dt} = \mathcal{M}(\mathbf{u}, \mathbf{x}, t),$$

donde  $\mathcal{M}$  es un operador que determina la dinámica del proceso. Además, suponemos que las observaciones están dadas por

$$\mathbf{y} = \mathcal{H}(\mathbf{u}, t) + \epsilon,$$

donde  $\mathcal{H}$  es un operador de observación, que depende del estado  $\mathbf{u}$  y del tiempo  $t$ . Los datos medidos se expresan entonces como el estado observado más errores de observación, dados por  $\epsilon$ .

La idea central detrás del proceso de estimación de parámetros dinámico es buscar un balance entre el conocimiento a priori que se tiene del fenómeno, lo que se conoce como el estado histórico o *background*, y el estado que ha sido observado o medido en el terreno. En el caso de los parámetros en la predicción de la evolución de la COVID-19 en el Ecuador, el estado histórico queda determinado por el conocimiento que se tiene del virus y su comportamiento, algo que ha sido bien establecido en estudios desarrollados, por ejemplo, en Li et al (2020). Por otro lado, el estado medido estaría dado por los resultados arrojados por la cifras nacionales que se obtienen por la toma de muestras, cifras de defunción, etc.

Extendiendo el problem (5), se busca entonces resolver el problema variacional

$$\begin{aligned} \min_x \quad & \frac{1}{2} \int_0^T (y(t) - \mathcal{H}(\mathbf{u}, t))^T \mathbf{R}_t^{-1} (y(t) - \mathcal{H}(\mathbf{u}, t)) d\mu \\ & + \frac{1}{2} (x - x^b)^T \mathbf{B}^{-1} (x - x^b), \\ \text{s.t.} \quad & \frac{d\mathbf{u}}{dt} = \mathcal{M}(\mathbf{u}, \mathbf{x}, t), \end{aligned} \quad (6)$$

donde  $\mu$  puede corresponder a la medida de Lebesgue en caso de considerar observaciones continuas o a la medida de conteo en caso de que las mediciones se obtengan en instantes puntuales del tiempo. Si, adicionalmente, el conjunto de parámetros consiste por una parte, en los coeficientes del modelo y por otra, en la condición inicial del mismo, el segundo término del funcional se puede dividir a su vez en dos, dando lugar al problema variacional

$$\begin{aligned} \min_{\mathbf{u}_0, \Theta} \quad & \frac{1}{2} \int_0^T (y(t) - \mathcal{H}(\mathbf{u}, t))^T \mathbf{R}_t^{-1} (y(t) - \mathcal{H}(\mathbf{u}, t)) d\mu \\ & + \frac{1}{2} (\Theta - \Theta^b)^T \mathbf{B}_{\Theta}^{-1} (\Theta - \Theta^b) \\ & + \frac{1}{2} (\mathbf{u}_0 - \mathbf{u}_0^b)^T (\mathbf{P}_0^b)^{-1} (\mathbf{u}_0 - \mathbf{u}_0^b), \\ \text{s.t.} \quad & \frac{d\mathbf{u}}{dt} = \mathcal{M}(\mathbf{u}, \mathbf{u}_0, \Theta, t), \end{aligned} \quad (7)$$

en el cual la matriz  $\mathbf{B}_{\Theta}$  corresponde a la covarianza de errores de fondo de los coeficientes del modelo y  $\mathbf{P}_0^b$  a la covarianza de errores de fondo de la condición inicial, respectivamente.

Dado que los operadores  $\mathcal{M}$  y  $\mathcal{H}$  son operadores diferenciales continuos y no lineales, se introducen comúnmente, aproximaciones discretas de los mismos, dadas por matrices  $\mathbf{M}$  y  $\mathbf{H}$ , respectivamente. Este proceso de aproximación se realiza, por ejemplo, con la aplicación de métodos de resolución numérica como el método de Euler implícito para la variable temporal y el método de Newton para la variable espacial. Dada esta aproximación en un intervalo temporal, del cual tomamos una sucesión de instantes  $t_k$ , con  $k = 0, 1, \dots$ , se tiene una aproximación de la dinámica del fenómeno dada por:

$$\mathbf{u}_{k+1} = \mathbf{M}_{k+1}(\mathbf{u}_k, \mathbf{x}_k), \quad (8)$$

donde suponemos que no hay errores en el modelo. De la misma forma, consideramos que las observaciones están dadas por:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{u}_k + \epsilon_k. \quad (9)$$

Suponemos que en el instante inicial  $k = 0$  conocemos el estado histórico  $\mathbf{u}_0^b$  y su matriz de covarianza  $\mathbf{P}_0^b$ . Además, suponemos que los errores son no correlacionados con las observaciones en (9). Así, sabemos que una condición inicial  $\mathbf{u}_0$  define una solución única del modelo  $\mathbf{x}_{k+1}$  según (8). Con estos antecedentes, el problema variacional discretizado es el siguiente:

$$\begin{aligned} \min \quad & \frac{1}{2} \sum_{k=0}^T (\mathbf{H}_k \mathbf{u}_k - \mathbf{y}_k)^\top \mathbf{R}_k^{-1} (\mathbf{H}_k \mathbf{u}_k - \mathbf{y}_k) \\ & + \frac{1}{2} (\Theta - \Theta^b)^\top \mathbf{B}_\Theta^{-1} (\Theta - \Theta^b) \\ & + \frac{1}{2} (\mathbf{u}_0 - \mathbf{u}_0^b)^\top (\mathbf{P}_0^b)^{-1} (\mathbf{u}_0 - \mathbf{u}_0^b), \\ \text{s.t.} \quad & \mathbf{u}_{k+1} = \mathbf{M}_{k+1}(\mathbf{u}_k, \mathbf{x}_k) \quad k = 0, 1, \dots \end{aligned} \quad (10)$$

La solución de este problema consiste en el conjunto de parámetros del modelo que maximizan la probabilidad a posteriori, a partir de lo cual es posible hacer una estimación robusta de la evolución del fenómeno. En este problema, nótese que se consideran las observaciones en todo el período temporal para la correcta estimación de los parámetros.

Un elemento esencial en (10) son las matrices de covarianza de errores. Mientras que  $\mathbf{R}$  puede construirse a partir de los errores de instrumentación y la cantidad de mediciones realizadas y  $\mathbf{B}_\Theta$  puede incorporar varianzas de parámetros previamente estudiadas, la matriz  $\mathbf{P}_0^b$  requiere de un tratamiento distinto. Esta matriz cuantifica de cierta manera la sensibilidad del problema con respecto a las condiciones iniciales, con lo cual se vuelve un elemento esencial para realizar pronósticos en modelos no lineales.

Una de las técnicas más utilizadas para actualizar la matriz  $\mathbf{P}_0^b$  consiste en remplazarla por una matriz construida a partir de un conjunto de ensamble, obtenido en base a perturbaciones aleatorias realizadas en los parámetros del modelo en cuestión y las correspondientes simulaciones en un horizonte temporal que permita capturar la dinámica del modelo.

### 3. PROBLEMA INVERSO PARA EL SARS-CoV-2

Uno de los mayores retos en el caso del SARS-CoV-2 es que muchas de las personas infectadas no presentan ningún síntoma, o desarrollan síntomas leves que pueden pasar desapercibidos. Esto hace que exista una importante fracción de la población infectada que no aparece en los registros y que son responsables de una gran cantidad de contagios, lo cual provoca a su vez que la modelización matemática de la propagación del virus y de la COVID-19 sea más compleja. Además, plantea retos a las autoridades sanitarias sobre el diseño eficiente y efectivo de cercos epidemiológicos y, sobre todo, diseño de campañas para la realización de pruebas tempranas a casos sospechosos, para su oportuno aislamiento.

La *Organización Mundial de la Salud (OMS)* sostiene que la mayoría de las personas (alrededor del 80%) se recupera de la enfermedad sin necesidad de realizar ningún tratamiento especial; alrededor de 1 de cada 6 personas que contraen la COVID-19 desarrolla una enfermedad grave; y, las personas mayores o las que padecen afecciones médicas subyacentes tienen más probabilidades de desarrollar una enfermedad grave. En torno al 2% de las personas que han contraído la enfermedad han muerto.

La familia de modelos matemáticos más usada para estudiar la proliferación de enfermedades virales son los modelos SEIR (por las siglas de *Susceptibles, Expuestos, Infectados y Removidos*). En estos modelos se crean compartimentos y se analizan las transiciones desde población susceptible a expuesta, la cual luego de un tiempo pasa a infecciosa, y de ésta, a removida. La remoción de los individuos se lleva a cabo mediante el aislamiento del resto de la población (cuarentena), la inmunización contra la infección (vacunación), la recuperación de la enfermedad con inmunidad total contra la reinfección o la muerte causada por la enfermedad (ver Brauer & Castillo-Chávez (2001)).

Para estudiar matemáticamente las particularidades del SARS-CoV-2, la variable personas infecciosas, en el modelo descrito anteriormente, se debe subdividir en dos compartimentos: personas infecciosas documentadas y no-documentadas. Así, los compartimentos poblacionales resultantes son: población susceptible, expuesta, infecciosa documentada, infecciosa no-documentada y removida. Estos compartimentos se acoplan entre sí mediante un sistema de ecuaciones diferenciales, que describe los flujos entre uno y otro, y que depende fuertemente de dos coeficientes dominantes: el índice de contagio (asociado al distanciamiento social) y la diseminación del virus por parte de los infecciosos no documentados. Además, se incorpora información de movilidad humana entre ciudades o provincias, obteniéndose el siguiente sistema:

$$\frac{dS_i}{dt} = -\frac{\beta_i S_i I_i^r}{N_i} - \frac{\mu_i \beta_i S_i I_i^u}{N_i} + \theta \sum_j \frac{M_{ij} S_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} S_i}{N_i - I_i^r} \quad (11a)$$

$$\frac{dE_i}{dt} = \frac{\beta_i S_i I_i^r}{N_i} + \frac{\mu_i \beta_i S_i I_i^u}{N_i} - \frac{E_i}{Z} + \theta \sum_j \frac{M_{ij} E_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} E_i}{N_i - I_i^r} \quad (11b)$$

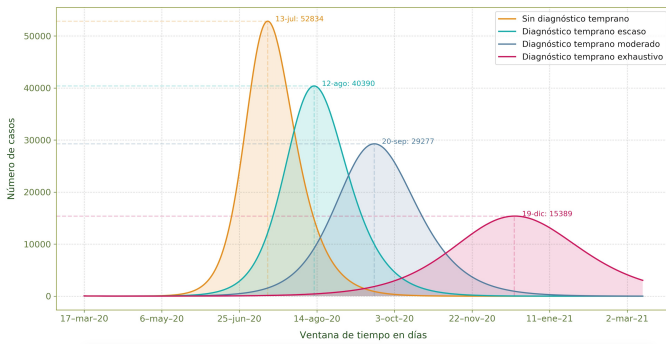
$$\frac{dI_i^r}{dt} = \alpha \frac{E_i}{Z} - \frac{I_i^r}{D} \quad (11c)$$

$$\frac{dI_i^u}{dt} = (1 - \alpha) \frac{E_i}{Z} - \frac{I_i^u}{D} + \theta \sum_j \frac{M_{ij} I_j^u}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} I_i^u}{N_i - I_i^u} \quad (11d)$$

$$\frac{dR_i}{dt} = \rho \frac{I_i^r + I_i^u}{D} \quad (11e)$$

$$N_i = N_i + \theta \sum_j M_{ij} - \theta \sum_j M_{ji}. \quad (11f)$$

Aquí,  $S_i$ ,  $E_i$ ,  $I_i^r$ ,  $I_i^u$ ,  $R_i$  son las soluciones del sistema de ecuaciones diferenciales (11) y representan a la población susceptible, expuesta, infecciosa documentada, infecciosa no documentada y removida, mientras que  $N_i$  representa la población total de la ciudad  $i$ . El sistema se complementa con condiciones iniciales para cada una de las variables del mismo. Las soluciones mencionadas son halladas mediante la aplicación de métodos de resolución numérica como el método de Euler implícito para la variable temporal y el método de Newton para la variable espacial.



**Figura 1.** Escenarios de propagación del virus en pacientes con sintomatología fuerte en la provincia de Pichincha al 17 de marzo del 2020

Se considera que los pacientes con síntomas lo suficientemente fuertes como para ser documentados son los que representan  $I_i^r$  (en principio, la población que debido a sus síntomas es diagnosticada por el sistema de salud), mientras que todas las otras personas infecciosas (en principio no diagnosticadas) representan la variable  $I_i^u$ . La constante  $\beta_i$  representa la tasa de transmisión debida a los casos de personas enfermas documentadas en la ciudad/provincia  $i$ , mientras que la tasa de transmisión de los casos no detectados es la misma  $\beta_i$ , reducida por un factor  $0 \leq \mu_i \leq 1$ . Esto último se debe a que la carga viral en las personas con pocos o ningún síntoma no es tan alta como en las personas que ya han sido documentadas. En Li et al (2020) se sostiene que las personas infectadas no documentadas en Wuhan fueron la mitad de contagiosas que las personas documentadas (antes del 23 de enero de 2020):  $\mu = 0.55$ ; 95 % CI: 0.46-0.62. Los otros parámetros del modelo son:  $\alpha$  que representa la fracción de las infecciones documentadas que desarrollaron síntomas severos,  $Z$  el tiempo promedio de latencia (tiempo de incubación del virus),  $D$  el tiempo promedio de duración del período infeccioso y  $\rho$  la tasa de recuperación. Se debe aclarar que  $R_i$  representa la población que luego de pasar por el período infeccioso, tiene inmunidad, al menos temporal. De acuerdo a los epidemiólogos, una vez que esta población cruce un umbral, la pandemia se da por controlada, al menos temporalmente.

Un parámetro fundamental de la epidemia es el número efectivo de reproducción  $R_e$ , el cual mide el número promedio de casos nuevos que genera un caso confirmado. En este modelo, este indicador se lo calcula, para la ciudad/provincia  $i$ , mediante

$$R_e^i := \alpha \beta_i D + (1 - \alpha) \mu_i \beta_i D.$$

Si este número se sostiene por encima de 1, la enfermedad seguirá expandiéndose (Ridenhour et al (2015)). Así, la meta es bajar este número por debajo de 1.

Finalmente,  $M_{ji}$  representa el número de personas que se desplazan desde la ciudad/provincia  $j$  a la ciudad/provincia  $i$ , mientras que  $\theta$  es un factor superior a uno, el cual se usa para corregir posibles subregistros de la movilidad de las personas. Estos criterios de movilidad son clave en el escenario global actual, donde la población está activamente interconectada y la movilidad es, si no imposible, muy difícil de suspender.

Para la estimación de los diferentes parámetros en el modelo (11) usamos el esquema variacional presentado previamente (ver la subsección 2.2). Para la información de fondo de los parámetros, utilizamos los intervalos de confianza calculados en Li et al (2020) que, aunque obtenidos en un contexto socio-cultural distinto, proveen una primera estimación, así como una medida de la incertidumbre. Estos intervalos son los siguientes:

- Tasa de transmisión de la población infectada reportada:  $0.8 \leq \beta \leq 1.5$ .
- Factor de reducción de la tasa de transmisión de población infectada no reportada:  $0.2 \leq \mu \leq 1$ .
- Factor multiplicativo de corrección de los datos de movilidad de la población:  $1 \leq \theta \leq 1.75$ .
- Tiempo promedio de días de latencia:  $2 \leq Z \leq 5$ , días.
- Fracción de pacientes que desarrollaron síntomas fuertes:  $0.02 \leq \alpha \leq 1$ .
- Tiempo promedio de días de duración de la infección:  $2 \leq D \leq 5$ , días.

Asumiendo una sola provincia aislada e introduciendo el vector de parámetros  $P := (\beta, \mu, \alpha, Z, D)^T$ , el problema inverso de identificación de parámetros está formulado mediante:

$$\begin{aligned} \min_{P \in \mathcal{U}_{ad}} & \sum_{i,j} (I^r(t_i) - I_i^{obs}) \mathbf{C}_{ij}^{-1} (I^r(t_j) - I_j^{obs}) \\ & + (P - P^b)^T \mathbf{B}^{-1} (P - P^b) = J(\bar{P}) \\ \text{s.t.} & \frac{dS}{dt} = -\frac{\beta S I^r}{N} - \frac{\mu \beta S I^u}{N}, & S(t_0) = S_0, \\ & \frac{dE}{dt} = \frac{\beta S I^r}{N} + \frac{\mu \beta S I^u}{N} - \frac{E}{Z}, & E(t_0) = E_0, \\ & \frac{dI^r}{dt} = \alpha \frac{E}{Z} - \frac{I^r}{D}, & I^r(t_0) = I_0^r, \\ & \frac{dI^u}{dt} = (1 - \alpha) \frac{E}{Z} - \frac{I^u}{D}, & I^u(t_0) = I_0^u, \\ & \frac{dR}{dt} = \frac{I^r + I^u}{D}, & R(t_0) = R_0 \end{aligned} \quad (12)$$

donde  $I_i^{obs}$ ,  $i = 1, \dots, n$ , son las observaciones históricas de infectados reportados obtenidas a partir de los registros oficiales,  $\mathbf{C} = (\mathbf{C}_{ij})$  corresponde a la matriz de covarianza de errores de observación,  $\mathbf{B}$  es la matriz de covarianza de errores de fondo y  $\mathcal{U}_{ad}$  es el conjunto admisible conformado por intervalos a priori en los cuales deben hallarse los parámetros. Mientras que la matriz  $\mathbf{B}$  es obtenida, en nuestro caso, a partir de los intervalos de confianza reportados en Li et al (2020), la matriz  $\mathbf{C}$  es estimada tomando en cuenta tanto la confiabilidad de las pruebas serológicas PCR como el número de tests que se realizan en cada provincia a la semana. De manera específica, se tiene la fórmula:

$$\sigma_{ii}^2 = \frac{\kappa}{\eta_i} \sigma_{PCR}^2, \quad (13)$$

donde  $\kappa$  es la cantidad ideal de tests semanales a realizarse por cada millón de habitantes (actualmente tomamos el valor  $\kappa = 4000$ ),  $\eta_i$  es la cantidad de tests por millón de habitantes realizados en la provincia bajo estudio en la semana en cuestión, y  $\sigma_{PCR}^2$  es la varianza propia de las pruebas PCR.

Al tener como restricción un sistema de ecuaciones diferenciales, el problema (12) debe ser tratado como un problema de control óptimo, en el cual la variable de estado es elemento de las funciones absolutamente continuas,  $AC([t_0, t_n], \mathbb{R}^5)$ . Para hallar su condición de optimalidad utilizaremos la técnica del Lagrangeano.

Sean  $\lambda_S, \lambda_E, \lambda_I, \lambda_U, \lambda_R$  los multiplicadores de Lagrange asociados a las restricciones diferenciales del problema (12). Definimos el Lagrangeano del problema por:

$$\begin{aligned} \mathcal{L} = J(P) + \int_{t_0}^{t_n} & \left( \lambda_S(t) \left[ -\frac{dS}{dt} - \frac{\beta SI^r + \mu \beta SI^u}{N} \right] + \lambda_E(t) \left[ -\frac{dE}{dt} \right. \right. \\ & + \left. \frac{\beta SI^r + \mu \beta SI^u}{N} - \frac{E}{Z} \right] + \lambda_I(t) \left[ -\frac{dI^r}{dt} + \alpha \frac{E}{Z} - \frac{I^r}{D} \right] + \lambda_U(t) \left[ -\frac{dI^u}{dt} \right. \\ & \left. \left. + (1 - \alpha) \frac{E}{Z} - \frac{I^u}{D} \right] + \lambda_R(t) \left[ -\frac{dR}{dt} + \frac{I^r + I^u}{D} \right] \right) dt + \ell, \quad (14) \end{aligned}$$

donde

$$\begin{aligned} \ell = & \theta_S(S(t_0) - S_0) + \theta_E(E(t_0) - E_0) + \theta_I(I^r(t_0) - I_0^r) \\ & + \theta_U(I^u(t_0) - I_0^u) + \theta_R(R(t_0) - R_0), \end{aligned}$$

con  $\theta_S, \theta_E, \theta_I, \theta_U, \theta_R$  los multiplicadores de Lagrange asociados a las condiciones iniciales de las variables de estado.

Como primer paso para obtener el sistema adjunto, debemos calcular la derivada direccional del Lagrangeano respecto a la variable de estado,  $x = (S, E, I^r, I^u, R)$ , en una dirección arbitraria  $v = (\sigma, \xi, w, v, r) \in AC([t_0, t_n], \mathbb{R}^5)$ , e igualar el resultado a cero, (ver, por ejemplo, Alekseev et al (1987); De los Reyes (2015)), es decir,

$$\mathcal{L}_x[v] = \mathcal{L}_S[\sigma] + \mathcal{L}_E[\xi] + \mathcal{L}_{I^r}[w] + \mathcal{L}_{I^u}[v] + \mathcal{L}_R[r] = 0.$$

Luego,

$$\begin{aligned} \mathcal{L}_S[\sigma] = & \int_{t_0}^{t_n} \left[ \frac{d\lambda_S}{dt} + \left( \frac{\beta I^r}{N} + \frac{\mu \beta I^u}{N} \right) (\lambda_E - \lambda_S) \right] \sigma dt \\ & - \lambda_S(t_n) \sigma(t_n) + \sigma(t_0) (\lambda_S(t_0) + \theta_S) = 0 \\ \mathcal{L}_E[\xi] = & \int_{t_0}^{t_n} \left[ \frac{d\lambda_E}{dt} + \alpha \frac{\lambda_I}{Z} + (1 - \alpha) \frac{\lambda_U}{Z} - \frac{\lambda_E}{Z} \right] \xi dt \\ & - \lambda_E(t_n) \xi(t_n) + \xi(t_0) (\lambda_E(t_0) + \theta_E) = 0 \\ \mathcal{L}_{I^r}[w] = & \int_{t_0}^{t_n} \left[ \frac{d\lambda_I}{dt} + \frac{\partial J}{\partial I^r} - \frac{\beta S \lambda_S}{N} + \frac{\beta S \lambda_E}{N} - \frac{\lambda_I}{D} + \frac{\lambda_R}{D} \right] w dt \\ & + w(t_n) (C_{nn}^{-1} (I^r(t_n) - I_n^{obs}) - \lambda_I(t_n)) + w(t_0) (\lambda_I(t_0) + \theta_I) = 0 \\ \mathcal{L}_{I^u}[v] = & \int_{t_0}^{t_n} \left[ \frac{d\lambda_U}{dt} - \frac{\mu \beta S \lambda_S}{N} + \frac{\mu \beta S \lambda_E}{N} - \frac{\lambda_U}{D} + \frac{\lambda_R}{D} \right] v dt \\ & - v(t_n) \lambda_U(t_n) + v(t_0) (\lambda_U(t_0) + \theta_U) = 0 \\ \mathcal{L}_R[r] = & \int_{t_0}^{t_n} \frac{d\lambda_R}{dt} r dt - r(t_n) \lambda_R(t_n) + r(t_0) (\lambda_R(t_0) + \theta_R) = 0. \end{aligned}$$

Finalmente, ya que la dirección fue tomada de manera arbitraria, de lo anterior, se sigue que el sistema adjunto está dado por:

$$\begin{aligned} -\frac{d\lambda_S}{dt} = & \left( \frac{\beta I^r}{N} + \frac{\mu \beta I^u}{N} \right) (\lambda_E - \lambda_S), & \lambda_S(t_n) = 0 \\ -\frac{d\lambda_E}{dt} = & -\frac{1}{Z} \lambda_E + \frac{\alpha}{Z} \lambda_I + \frac{(1 - \alpha)}{Z} \lambda_U, & \lambda_E(t_n) = 0 \\ -\frac{d\lambda_I}{dt} = & \frac{\partial J}{\partial I^r} - \frac{\beta S}{N} (\lambda_S - \lambda_E) - \frac{\lambda_I}{D} + \frac{\lambda_R}{D}, & \lambda_I(t_n) = \\ & C_{nn}^{-1} (I^r(t_n) - I_n^{obs}) \\ -\frac{d\lambda_U}{dt} = & -\frac{\mu \beta S}{N} (\lambda_S - \lambda_E) - \frac{\lambda_U}{D} + \frac{\lambda_R}{D}, & \lambda_U(t_n) = 0 \\ -\frac{d\lambda_R}{dt} = & 0, & \lambda_R(t_n) = 0 \end{aligned}$$

El gradiente de la función objetivo se obtiene al derivar el Lagrangeano con respecto a la variable de control. Se procede de esta manera y obtenemos las siguientes ecuaciones:

$$\begin{aligned} \nabla_{\beta} J = & \mathbf{B}_{\beta}^{-1} (\beta - \beta^b) - \int_0^T \frac{S}{N} (I^r + \mu I^u) (\lambda_S - \lambda_E), \\ \nabla_{\mu} J = & \mathbf{B}_{\mu}^{-1} (\mu - \mu^b) - \int_0^T \beta \frac{S}{N} I^u (\lambda_S - \lambda_E), \\ \nabla_{\alpha} J = & \mathbf{B}_{\alpha}^{-1} (\alpha - \alpha^b) + \int_0^T \frac{E}{Z} (\lambda_I - \lambda_U), \\ \nabla_Z J = & \mathbf{B}_Z^{-1} (Z - Z^b) + \int_0^T \frac{E}{Z^2} (\lambda_E - \alpha \lambda_I - (1 - \alpha) \lambda_U), \\ \nabla_D J = & \mathbf{B}_D^{-1} (D - D^b) + \int_0^T \frac{1}{D^2} (I^r \lambda_I + I^u \lambda_U - \lambda_R (I^r + I^u)). \end{aligned}$$

Ya que los parámetros del problema pertenecen a un conjunto de controles admisibles  $\mathcal{U}_{ad}$ , el cual incorpora los intervalos de confianza, la condición de optimalidad está definida por la siguiente desigualdad variacional

$$\nabla J(\bar{P})^T (P - \bar{P}) \geq 0, \quad \forall P \in \mathcal{U}_{ad}. \quad (15)$$

De manera adicional, ya que  $\mathcal{U}_{ad}$  está compuesto por restricciones de caja, la desigualdad variacional puede ser remplazada por una proyección componente a componente. Para la solución del problema, por tanto, usamos un método proyectado de segundo orden, con actualización de la matriz de tipo cuasi-Newton (para más detalles, ver Kelley (1999)).

#### 4. RESULTADOS

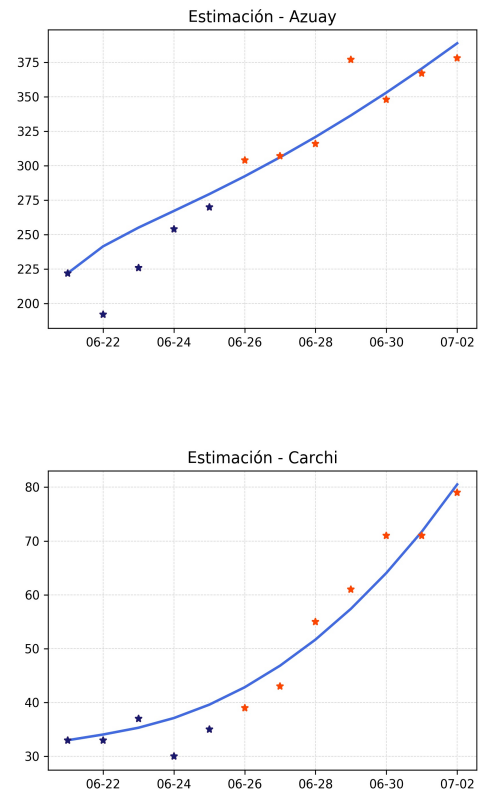
En esta sección presentamos, a manera de ejemplo, los resultados de la estimación de parámetros del modelo (11) para la semana del 26 de junio al 2 de julio de 2020, así como las respectivas proyecciones quincenales de casos reportados de infectados con el coronavirus en diferentes provincias del país. Cabe indicar que este tipo de estadísticas son calculadas por el *Centro de Modelización Matemática* con una periodicidad semanal y puestas a disposición de los tomadores de decisión en materia de políticas de salud.

Para la estimación de parámetros, de lo expuesto en secciones anteriores, consideramos el vector de parámetros a estimar,  $P = (\beta, \mu, \alpha, Z, D)$ , y calculamos sus valores mediante la solución del problema (12). Es importante recordar que este problema variacional toma en cuenta la incertidumbre de las mediciones a través de la matriz de covarianza de errores  $\mathbf{C}$ . Para la construcción de  $\mathbf{C}$  se considera el nivel de confiabilidad de las pruebas de laboratorio *PCR* y el número de tests que se realiza (ver la fórmula (13)). La resolución numérica de (12) es realizada para cada provincia en una ventana de tiempo de 7 días. Asimismo, tanto  $\mathbf{C}$  como los valores de fondo con los cuales se realiza la estimación son actualizados de forma semanal.

En la Figura 2, los puntos representan la información oficial del número de infectados reportados para cada provincia, siendo los de color naranja los correspondientes al periodo de estimación, es decir del 26 de junio al 2 de julio de 2020. Se presenta, además, en color azul, la información del número de personas infectadas en los 5 días anteriores al periodo de estudio, ya que debido al tiempo que transcurre entre la realización y la obtención de los resultados de las pruebas serológicas *PCR*, la condición inicial del problema se considera con 5 días de retraso.

Como resultado de la estimación se obtienen parámetros óptimos para el período de estudio considerado. La curva azul de la Figura 2 corresponde a la solución del modelo (11) para la variable  $I_t^r$ , es decir, la población que debido a sus síntomas es diagnosticada con el virus. Esta curva ajusta las observaciones, tomando en cuenta el modelo matemático y la incertidumbre de la data oficial.

Con los parámetros obtenidos, calculamos adicionalmente el número de reproducción efectivo  $R_e$ . Ya que el modelo es resuelto

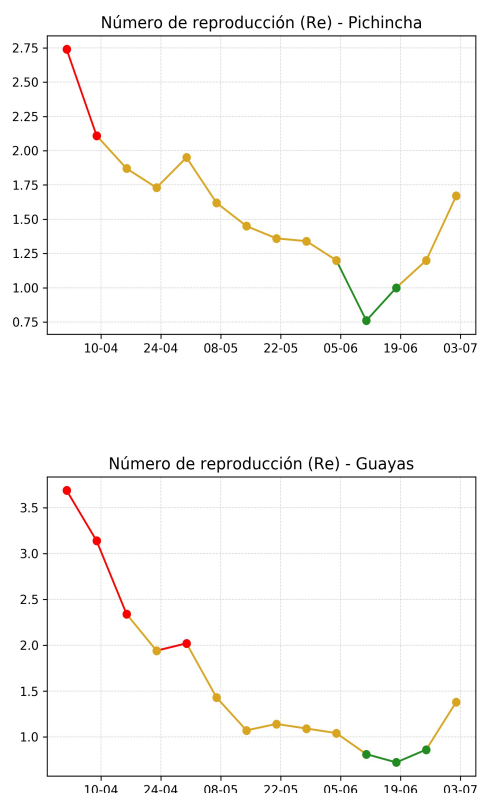


**Figura 2.** Estimación de parámetros en la semana del 26 de junio al 2 de julio de 2020, para las provincias de Azuay (arriba.) y Carchi (abajo.).

semanalmente, podemos también calcular un  $R_e$  actualizado con esa frecuencia. Esto es importante ya que nos permite visualizar la evolución de la epidemia en cada región de estudio, es decir, en cada provincia.

En la Figura 3 se presenta la evolución del número de reproducción efectivo  $R_e$  para las provincias de Pichincha y Guayas, desde la semana del 27 de marzo hasta la semana del 26 de junio de 2020. Los colores de las gráficas nos permiten visualizar rápidamente el estado del  $R_e$  cada semana. En rojo, se presentan aquellos valores de  $R_e \geq 2$ , en amarillo los valores entre  $1 < R_e < 2$ , mientras que en verde los valores de  $R_e \leq 1$ . Recordemos que si este número se sostiene por encima de 1, la enfermedad seguirá expandiéndose, siendo un indicador de que el virus empieza a estar controlado cuando el  $R_e$  se encuentra por debajo de 1.

La segunda parte de los resultados corresponde al pronóstico del número de personas infectadas oficialmente registradas. Estos pronósticos se los realiza de manera semanal para todas las provincias del país y para una ventana de tiempo de 15 días. Los parámetros que se utilizan son los calculados con la metodología descrita en este reporte. Adicionalmente, para resolver el sistema de ecuaciones consideramos aleatoriedad en los parámetros, las condiciones iniciales y los términos en el sistema de ecuaciones (Li et al (2020)). Para los parámetros y las condiciones iniciales consideramos perturbaciones normales, mientras que la estocasticidad de las ecuaciones se refleja a través de distribuciones de probabilidad de Poisson.

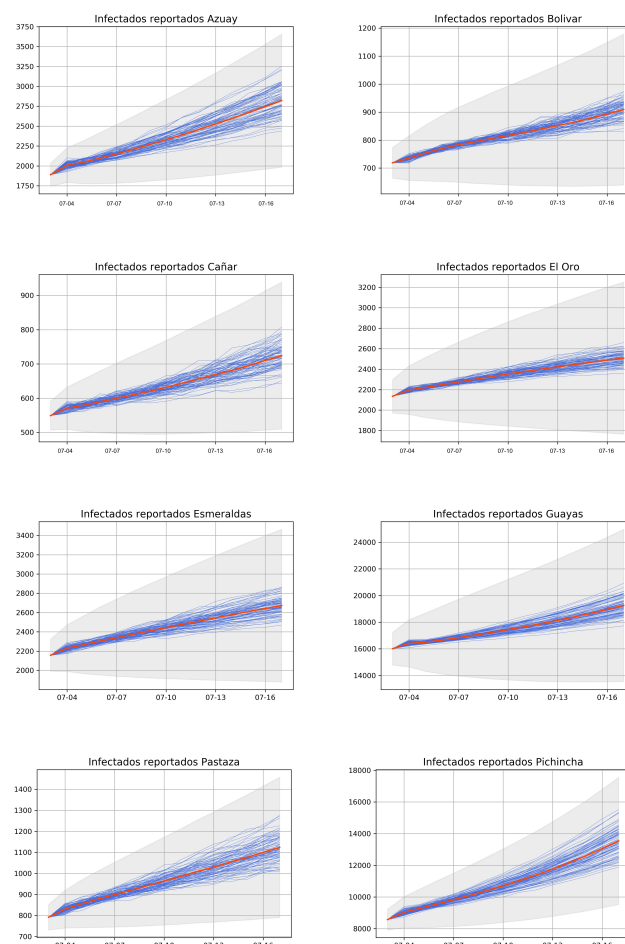


**Figura 3.** Evolución del número de reproducción efectivo, para la provincia de Pichincha (arriba) y la provincia de Guayas (abajo).

Para cada provincia la simulación se ejecuta con un ensamble de tamaño 80. Los resultados de estas simulaciones se presentan en la Figura 4 con líneas de color azul; para cada una de estas corridas la condición inicial es perturbada. En anaranjado se presenta la media de todas las simulaciones realizadas, mientras que el área gris representa el intervalo de confianza al 95 % para el valor medio calculado. Como resultado, de entre todas las simulaciones realizadas presentamos las tres más significativas: escenario optimista (menor número de infectados reportados), escenario pesimista (mayor número de infectados reportados) y el pronóstico promedio.

En la Figura 4 se muestran las proyecciones de los casos reportados de COVID-19 (número de personas infectadas reportadas por las autoridades) de 8 de las 24 provincias del país, para el periodo del 3 al 18 de julio. La condición inicial fue tomada el día 2 de julio de 2020.

Además de los reportes gráficos, para cada provincia se obtienen tablas con los valores numéricos de los escenarios pesimista, optimista y promedio. En la Figura 5 observamos las tablas con los valores de los pronósticos para el periodo del 3 al 18 de julio del 2020 para las provincias de Esmeraldas y Santa Elena.



**Figura 4.** Pronóstico de infectados reportados acumulados para el periodo del 3 al 18 de julio del 2020. 1ra fila: Azuay(izq.), Bolívar(der.). 2da fila: Cañar(izq.), El Oro(der.). 3ra fila: Esmeraldas(izq.), Guayas(der.). 4ta fila: Pastaza(izq.), Pichincha(der.).

## 5. CONCLUSIONES

El monitoreo de la pandemia del SARS-CoV-2 en Ecuador requiere de la estimación rigurosa de diferentes indicadores de interés bajo condiciones de incertidumbre. Uno de las cantidades más importantes en este sentido es el número de reproducción efectivo de la pandemia, el cual requiere de la estimación robusta de los parámetros de un modelo matemático que recoja adecuadamente la dinámica viral. En este artículo presentamos el esquema desarrollado para llevar a cabo dicha tarea de manera periódica en cada región de estudio, proveyendo una herramienta importante de monitoreo y toma de decisiones.

La estimación de parámetros mediante técnicas variacionales, que utilizan un enfoque bayesiano brinda ventajas sobre otros métodos de estimación. Particularmente, en el caso ecuatoriano, esto resulta importante debido al alto nivel de incertidumbre proveniente de la data oficial; un factor determinante que no es considerado en otros enfoques.

Fecha	Escenario pesimista	Pronóstico	Escenario optimista
2020-07-03	2157	2157	2157
2020-07-04	2290	2230	2173
2020-07-05	2319	2266	2218
2020-07-06	2354	2304	2256
2020-07-07	2394	2339	2291
2020-07-08	2454	2376	2297
2020-07-09	2504	2409	2337
2020-07-10	2526	2443	2329
2020-07-11	2597	2478	2353
2020-07-12	2660	2511	2384
2020-07-13	2715	2544	2395
2020-07-14	2757	2578	2420
2020-07-15	2782	2610	2437
2020-07-16	2827	2642	2447
2020-07-17	2865	2672	2470

Fecha	Escenario pesimista	Pronóstico	Escenario optimista
2020-07-03	944	944	944
2020-07-04	960	952	942
2020-07-05	964	955	946
2020-07-06	971	960	951
2020-07-07	976	964	954
2020-07-08	982	968	956
2020-07-09	988	973	960
2020-07-10	1000	977	962
2020-07-11	1008	981	963
2020-07-12	1026	985	962
2020-07-13	1028	990	971
2020-07-14	1035	994	969
2020-07-15	1047	998	970
2020-07-16	1054	1003	973
2020-07-17	1072	1008	974

**Figura 5.** Tablas con proyecciones de infectados reportados acumulados para el periodo del 3 al 18 de julio del 2020 de Esmeraldas(arriba) y Santa Elena(abajo).

La metodología para la construcción de las matrices de covarianza, presentes en el funcional de costo, permite incorporar la mencionada incertidumbre. En efecto, la construcción de  $C$  toma en cuenta la confiabilidad de las pruebas serológicas PCR, así como el número de test que se realizan en una población determinada. Mientras menos tests se realicen, la varianza será más alta.

Las soluciones del sistema de ecuaciones que modelan la propagación del virus, dependen fuertemente de los parámetros del modelo, siendo los dos coeficientes dominantes: el índice de contagio y la diseminación del virus por parte de los infecciosos no-documentados. En este sentido, las soluciones que más se aproximen a la realidad en el caso ecuatoriano serán aquellas cuyos parámetros hayan sido estimados utilizando una metodología que considere las particularidades de la región de estudio, como las descritas en este artículo.

Las proyecciones de propagación del virus consideran aleatoriedad en los parámetros, las condiciones iniciales y los diferentes términos en el sistema de ecuaciones diferenciales del modelo.

Gracias a la aleatoriedad podemos realizar varias simulaciones y presentar distintos escenarios para cada región de estudio.

## REFERENCIAS

- Alekseev, V.M., Tikhomirov, V.M., and Fomin, S.V. (1987). *Optimal Control*. Springer Science and Business Media
- Asch, M., Bocquet, M., and Nodet, M. (2016). *Data assimilation: methods, algorithms, and applications*. Society for Industrial and Applied Mathematics.
- Aster, R., Borchers, B. and Thurber, C.H. (2013). *Parameter Estimation and Inverse Problems*. Academic Press.
- Brauer, F., Castillo-Chávez, C. (2001). *Mathematical Models in Population Biology and Epidemiology*. Springer.
- Chowell, G. (2017). Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecast. *Infectious Disease Modelling*. DOI: 10.1016/j.idm.2017.08.001
- De los Reyes, J.C. (2015). *Numerical PDE-Constrained Optimization*. Springer Verlag.
- Kelley, C.T. (1999). *Iterative Methods for Optimization*. Society for Industrial and Applied Mathematics.
- Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S., Reich, N.G., Lessler, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Annals of Internal Medicine*. DOI: 10.7326/M20-0504
- Li, R., Pei, S., Chen, B., Song, Y., Zhang, T., Yang, W. and Shaman, J. (2020). Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2). *Science*. DOI: 10.1126/science.abb3221
- Ridenhour, B., Kowalik, J.M. and Shay, D.K. (2015). El número reproductivo básico ( $R_0$ ): consideraciones para su aplicación en la salud pública. *PanAmerican Journal of Public Health*. DOI: 10.2105/AJPH.2013.301704
- Verity, R., Okell, L. C., Dorigatti, I., Winskill, P., Whittaker, C., Imai, N., and Dighe, A. (2020). Estimates of the severity of coronavirus disease 2019: a model-based analysis. *The Lancet infectious diseases*. [https://doi.org/10.1016/S1473-3099\(20\)30243-7](https://doi.org/10.1016/S1473-3099(20)30243-7)

Wu, D., Wu, T., Liu, Q. and Yang, Z. (2020). The SARS-CoV-2 outbreak: what we know. *International Journal of Infectious Diseases*. DOI: 10.1016/j.ijid.2020.03.004

diferenciables y modelamiento de problemas provenientes de las biociencias. ORCID: <https://orcid.org/0000-0002-8178-8834>

## BIBLIOGRAFÍAS



**Paula M. Castro C.** Estudiante del Programa de Doctorado en Matemática Aplicada de la Escuela Politécnica Nacional (EPN) e investigadora del centro de Modelización Matemática (MODEMAT). Ingeniera Matemática de la EPN (2015) y Magister en optimización matemática (2017) por la misma casa de estudios. Su área de estudios es el análisis variacional de problemas de asimilación de datos. ORCID: <https://orcid.org/0000-0002-2560-9096>

lisis variacional de problemas de asimilación de datos. ORCID: <https://orcid.org/0000-0002-2560-9096>



**Juan Carlos De los Reyes.** Ph.D. en Matemática. Director del Centro de Modelización Matemática (MODEMAT). Profesor en el área de optimización y control del departamento de matemática de la EPN. Es autor de varios artículos científicos y ha realizado investigaciones y ponencias en universidades y centros de investigación de varios países.

Es miembro activo de la Academia de Ciencias del Ecuador (ACE) y de la Academia Mundial de Ciencias (TWAS). ORCID: <https://orcid.org/0000-0003-2761-5382>



**Sergio González-Andrade.** Ph.D. en Matemática. Director del programa doctoral en Matemática Aplicada de la Escuela Politécnica Nacional (EPN). Investigador del centro de Modelización Matemática (MODEMAT) y profesor del departamento de matemática de la EPN. Sus áreas de investigación son la simulación numérica de fluidos no-Newtonianos

así como el estudio de los métodos multimalla y el análisis de los métodos de Newton generalizados. ORCID: <https://orcid.org/0000-0001-7022-1245>



**Pedro Martín Merino Rosero.** Ph.D. en Matemática. Coordinador del laboratorio nacional de cálculo científico HPCModemat. Investigador del centro de Modelización Matemática (MODEMAT) y profesor del departamento de matemática de la EPN. Actualmente sus intereses científicos se centran en métodos de optimización de problemas no