

MODELIZACIÓN Y SIMULACIÓN DE LA PROPAGACIÓN DEL VIRUS SARS-COV-2 EN ECUADOR

Autores (en orden alfabético): Paula Castro, Juan Carlos De los Reyes,
Sergio González, Pedro Merino, Joan Ponce

Centro de Modelización Matemática en Áreas Clave para el Desarrollo
MODEMAT

Escuela Politécnica Nacional de Ecuador
modemat@epn.edu.ec

26 de marzo de 2020

Resumen Ante la inminente llegada de la pandemia de la enfermedad Covid 19, conocida como Coronavirus, al territorio ecuatoriano, el Centro de Modelización Matemática de la Escuela Politécnica Nacional conformó un equipo de trabajo especializado para modelizar y simular la propagación del SARS-CoV-2 (virus causante de la enfermedad), bajo varios escenarios de política pública aplicados a la contención del mismo. En este reporte explicamos los modelos utilizados, las simulaciones obtenidas y las recomendaciones para mejorar los modelos y, sobre todo, para el diseño de políticas públicas que ayuden a minimizar los efectos negativos de la pandemia en la población, mientras se desarrollan las actividades productivas esenciales para el país.

1. ¿QUÉ SON EL SARS-COV-2 Y LA ENFERMEDAD COVID 19?

Los coronavirus son un tipo de virus de ARN monocatenario (*single-stranded ARN viruses*), los cuales afectan animales y humanos, provocando enfermedades respiratorias, hepáticas, gastrointestinales e, incluso, neurológicas. Este tipo de virus CoVs se divide en cuatro clases: alfa-coronavirus, beta-coronavirus, gamma-coronavirus y delta-coronavirus. Actualmente, 6 tipos de coronavirus humanos han sido identificados, incluyendo el SARS-CoV y el MERS-CoV, responsables de pandemias de enfermedades respiratorias reportadas en 2003 y 2012, respectivamente (ver [Di Wu et al]).

Los coronavirus, al parecer, emergen periódicamente en poblaciones humanas, debido a varios factores entre los que se pueden incluir la prevalencia alta y la amplia distribución geográfica de estos virus, y el incremento de las interacciones entre humanos y animales, sobre todo, animales silvestres.

A finales de diciembre de 2019, en la ciudad de Wuhan, China, se identificaron pacientes que presentaban neumonía sin causa aparente. Esto llevó a varios grupos de científicos chinos a identificar un nuevo tipo de coronavirus que estaba detrás de estos casos. Una vez hecha la

identificación, y dada la alta propagación del contagio en la ciudad de Wuhan y en la provincia de Hubei, China, la Organización Mundial de la Salud, el 30 de enero de 2020, bautizó al nuevo virus como SARS-CoV-2, denominó COVID-19 al tipo de neumonía que produce el virus, y declaró la emergencia médica mundial.

El SARS-CoV-2, al igual que el SARS-CoV y el MERS-CoV, es un virus beta-corona de ARN monocatenario positivo y su origen es de orden zoonótico, es decir, saltó de un huésped animal (todavía sin identificar) a uno humano.

Por otro lado, el número básico de reproducción (la cantidad de personas a las que una persona infectada puede contagiar) de este nuevo coronavirus es relativamente alto en comparación a otros virus similares. En efecto, en [Ruiyun et al] se estima que el número efectivo de reproducción R_e es de 2.38 (95 % CI: 2.04 - 2.77), lo que muestra una alta capacidad del virus para sostener la transmisión en el tiempo. Un desafío, por tanto, es estimar este número con la información local y, a partir de estas estimaciones, buscar mecanismos para reducirlo hasta números por debajo de 1, lo que implicaría que la pandemia está controlada.

1.1. MECANISMOS DE CONTAGIO

Una persona puede contraer la COVID-19 por contacto con otra que esté infectada por el virus. La enfermedad puede propagarse de persona a persona a través de las gotículas procedentes de la nariz o la boca que salen despedidas cuando una persona infectada tose, estornuda o exhala. Estas gotículas caen sobre superficies y objetos que rodean a la persona, lo que provoca que otras personas puedan entrar en contacto con el virus al tocar estas superficies u objetos, y pueden contraer la COVID-19 si se tocan la cara, permitiendo al virus ingresar al cuerpo por nariz, boca y ojos. Hay riesgo de contagio también si se inhalan las gotículas desprendidas por las personas enfermas. Esto ha llevado a que las tres principales recomendaciones sean: mantener al menos un metro de distancia entre personas, lavarse frecuentemente las manos y evitar tocarse la cara con las manos sin lavar y/o desinfectar (ver [OMS]).

Por otro lado, la mayoría de estudios sostienen que la principal vía de contagio es mediante el contacto directo con personas enfermas y no hay evidencia suficiente para suponer contagio a través del aire.

1.2. TIEMPO DE INCUBACIÓN

El tiempo de incubación está en estudio todavía. En varias contribuciones se ha estimado que el tiempo promedio para la manifestación de los primeros síntomas está alrededor de 5 días. En [Lauer et al], luego de un estudio sobre una muestra de 180 pacientes confirmados se establece que la mediana del período de incubación se estimó en 5.1 días (95 % CI, 4.5 a 5.8 días). Además, se concluyó que el 97.5 % de los pacientes desarrolló síntomas en 11.5 días (CI, 8.2 a 15.6 días) de infección. En este estudio, los autores suponen que las personas están expuestas al SARS-CoV-2 inmediatamente antes del programa de monitoreo activo y asumen la perfecta verificación de los casos sintomáticos que se desarrollan bajo monitoreo activo. Bajo estas suposiciones, concluyen que 101 de cada 10 000 casos (percentil 99, 482) desarrollarán síntomas después de 14 días de monitoreo activo o cuarentena. En [Di Wu et al] se dan estimaciones similares.

La 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 tiene dificultad para respirar. Las personas mayores y las que padecen afecciones médicas subyacentes, como hipertensión arterial, problemas cardíacos o diabetes, tienen más probabilidades de desarrollar una enfermedad grave. En torno al 2 % de las personas que han contraído la enfermedad han muerto (ver [OMS]).

1.2.1. SUBREGISTRO DE CASOS

El problema, desde el punto de vista sanitario, es que algunas personas se infectan pero no desarrollan ningún síntoma o desarrollan síntomas leves que pueden pasar desapercibidos. Esto hace que haya una importante población infectada que no aparece en los registros y que, al parecer, son responsables de la mayor cantidad de contagios. En efecto, en [Ruiyun et al] se estima que el 86 % de todas las infecciones fueron no documentadas (95 % CI: [82 %-90 %]) antes del 23 de enero de 2020 (fecha de inicio de las restricciones de movilidad) en la ciudad de Wuhan y la provincia de Hubei en China. La tasa de transmisión por persona de infecciones sin documentar fue del 55 % de las infecciones documentadas ([46 %-62 %]). Los autores, sin embargo, estiman que las infecciones sin documentar fueron la fuente de infección de hasta el 79 % de los casos documentados. Este hecho hace que la modelización de la evolución y dispersión del virus y de la COVID-19 sea más compleja. Además, plantea retos a las autoridades sanitarias sobre diseño eficiente y efectivo de cercos epidemiológicos y, sobre todo, diseño de campañas para la realización de pruebas a casos sospechosos, desde los primeros días en que éstos presenten síntomas, para su aislamiento. Esto, como se verá posteriormente, puede ayudar a reducir la tasa de contagio y, en última instancia, a controlar la pandemia.

2. MODELO DE DISPERSIÓN DEL VIRUS Y LA ENFERMEDAD

El estudio cuantitativo del contagio y propagación de una enfermedad de tipo viral conlleva varios retos, como la estimación de las tasas efectivas de transmisión, las condiciones iniciales del contagio, el efecto de medidas de aislamiento, etc. Más aún, se sabe que la propagación de una enfermedad se lleva a cabo mediante encuentros aleatorios entre individuos sanos e infectados, los cuales son muy difíciles de identificar. Ventajosamente, conforme aumenta el tamaño de la población bajo estudio, la ley estadística de los grandes números nos permite predecir de mejor manera el número de infectados en la población. Debido a esto, los modelos determinísticos, basados en *sistemas de ecuaciones diferenciales ordinarias*, son los más usados para el estudio de enfermedades que afectan a poblaciones grandes, como es el caso del SARS-CoV-2.

La familia de modelos matemáticos más usada para estudiar la proliferación de enfermedades virales son los llamados modelos SIR (por las siglas de *Susceptibles, Infectados y Removidos*). En estos modelos se analizan las transiciones desde población susceptible 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-Chavez]).

En general, se considera que un individuo infectado transmite el virus, de acuerdo a un proceso estocástico de Poisson con tasa β ([Miller]). Un proceso de Poisson es un proceso estocástico (sucesión de variables aleatorias que ocurren en el tiempo) de tiempo continuo que se usa para hallar eventos específicos (un contagio en nuestro problema) que se dan en un período de tiempo. La tasa β representa, de modo general, el tiempo transcurrido entre dos eventos consecutivos [Last-Penrose]. La carga viral es recibida por un individuo elegido al azar de la población. Si el receptor es susceptible, entonces se vuelve infectado. Los individuos infectados se recuperan siguiendo un proceso de Poisson de un solo paso con tasa γ y, una vez recuperados, son inmunes a futuras infecciones. Este modelo se expresa mediante el siguiente sistema de ecuaciones diferenciales ordinarias:

$$(2.1) \quad \begin{aligned} \frac{dS}{dt} &= -\beta IS \\ \frac{dI}{dt} &= \beta IS - \gamma I \\ \frac{dR}{dt} &= \gamma I, \end{aligned}$$

A estas ecuaciones de la dinámica de propagación viral se agregan condiciones iniciales dadas por

$$\begin{aligned} S(0) &= 1 - r \\ I(0) &= r \\ R(0) &= 0, \end{aligned}$$

donde r es la proporción de población enferma al inicio del estudio. La primera ecuación muestra la evolución en el tiempo de la población susceptible. Se considera que una persona promedio de la población susceptible hace contacto suficiente para transmitir la infección con βN otras personas por unidad de tiempo, donde N representa el tamaño total de la población. A este efecto se lo conoce como incidencia de acción masiva. La población susceptible decrece de forma exponencial de acuerdo con la tasa de transmisión β y el incremento de la población infectada I . La segunda ecuación muestra la evolución de la población infectada, la cual crece exponencialmente de acuerdo con la tasa de transmisión β , pero, dado que a esta población se le va retirando la proporción de la población infectada que se recupera de acuerdo a la tasa de recuperación γ , alcanza un máximo (pico de la epidemia) y decrece exponencialmente. Finalmente, la última ecuación muestra la evolución de la población removida, la cual crece de acuerdo con la tasa de recuperación γ . Una característica de este modelo es que considera que la población N es constante.

En estos modelos, el número efectivo de reproducción inicial R_0 se calcula como $R_0 = \frac{\beta}{\gamma}$ (ver [Ridenhour et al]). De esta forma, si la tasa de recuperación es mayor a la tasa de transmisión, R_0 es menor que 1 y se considera que la epidemia puede ser contenida en un período corto de tiempo. Por el contrario, si la tasa de transmisión es mayor a la de recuperación, la enfermedad tiene la capacidad de extenderse, tanto en el tiempo como geográficamente, de forma acelerada.

A partir de estos conceptos, es posible dar una explicación a la estrategia que se ha discutido mucho en estos días: aplanar la curva. La idea detrás de esta estrategia es disminuir la tasa de contagio β y aumentar la tasa de recuperación γ , mediante la aplicación de varias medidas

que ayuden a disminuir velocidad de contagio de la población, por un lado, y a aumentar la proporción de la población que se recupera o que deja de ser susceptible (desarrollo de medicina y vacunas, etc.).

El comportamiento de estos modelos se puede apreciar en la figura 1, donde se ve que la población infectada alcanza un pico y descende (curva a ser aplanada), la población susceptible descende de forma acelerada y la población recuperada crece proporcionalmente.

Estos modelos, con sus variantes, se han utilizado de forma eficiente en el análisis cualitativo de varias epidemias, y han dado luces sobre la evolución de una infección y la proyección del daño que puede causar en determinadas poblaciones.

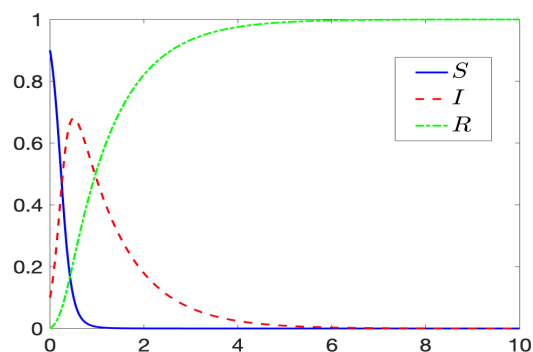


Figura 1: Simulación con el modelo SIR, considerando $\beta = 10$ y $\gamma = 1$. Aquí, la tasa de infección es alta respecto a la tasa de recuperación, lo que explica el crecimiento acelerado de la población infectada, llegando al pico de la epidemia en un corto período de tiempo. Además, se ve el rápido descenso de la población susceptible.

Una extensión de los modelos SIR son los llamados modelos SEIR (*Susceptibles, Expuestos, Infectados y Removidos*). La diferencia respecto de los modelos SIR es la incorporación de una variable que mide la población expuesta a la infección. De este modo, se puede medir la evolución de la población susceptible dividiéndola en personas no necesariamente expuestas a la infección y aquellas que sí lo están, por ejemplo, gente del círculo íntimo de personas ya infectadas o gente que ha tenido interacción con personas infectadas en sistemas de transporte público, etc. La idea central es modelizar el porcentaje de la población que ha sido expuesta a la infección y que, por algún tiempo, puede ser removida (porque no contrajo la infección) o ya pasa a ser parte de la población infectada. Con esta información, se puede entender de mejor forma la importancia de los cercos epidemiológicos en la propagación del virus. Estos modelos tienen la siguiente formulación:

$$\begin{aligned}
 \frac{dS}{dt} &= -\beta IS \\
 \frac{dE}{dt} &= \beta IS - \alpha E \\
 \frac{dI}{dt} &= \alpha E - \gamma I \\
 \frac{dR}{dt} &= \gamma I.
 \end{aligned}
 \tag{2.2}$$

Aquí, α es una tasa de exposición de la población (ver [Chowell]). Es sencillo ver que las

ecuaciones para la población susceptible S y la recuperada R son las mismas que en el caso del modelo SIR. El cambio central en este modelo es la introducción de la ecuación que mide la evolución de la población expuesta E . Esta ecuación mide la evolución de E , considerando la diferencia entre la población susceptible y la población expuesta mediada por la tasa α . Por otro lado, la ecuación para la población infectada I mide la evolución de la población infectada considerando la diferencia entre los expuestos que eventualmente desarrollarán síntomas y los infectados que se recuperan con tasa γ . Por la forma de estas ecuaciones, las dos poblaciones, tanto la expuesta como la infectada, crecen en una primera etapa de forma exponencial, alcanzan un pico, y descienden de forma también exponencial.

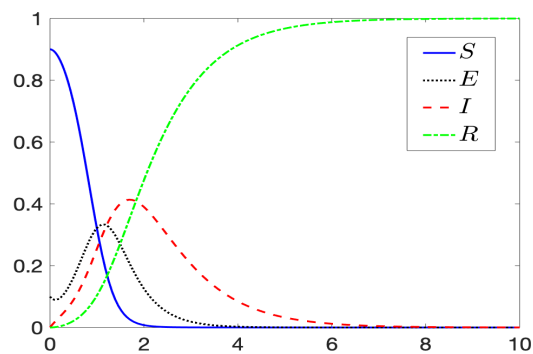


Figura 2: Simulación con el modelo SEIR, considerando $\beta = 10$, $\gamma = 1$ y $\alpha = 2$. Al igual que el modelo SIR, la tasa de infección es alta respecto a las tasas de exposición y recuperación, generando el crecimiento acelerado de la población expuesta y, posteriormente, de la población infectada. Aquí trabajamos con una condición inicial $E(0) = r$ e $I(0) = 0$, es decir, suponemos que la proporción de población expuesta es conocida, pero que al inicio de la epidemia no hay población infectada [Brauer-Castillo-Chavez].

Aún cuando este tipo de modelos resultan adecuados y suficientes para enfermedades virales como la influenza, en el caso del SARS-CoV-2 resultan poco confiables debido a las particularidades de esta nueva enfermedad viral. En particular, la gran cantidad de casos que no se detectan a tiempo de la enfermedad (Ver sección 1.2.1), demandan variantes importantes en el modelo y en su aplicación.

2.1. MODELO CON POBLACIÓN INFECTADA NO DETECTADA

Como se discutió antes, en [Ruiyun et al] los autores identificaron altos porcentajes de personas infectadas que no fueron detectadas como tales, y que contribuyeron a una alta propagación de la enfermedad en Wuhan. Considerando que éste parece ser el común denominador en la pandemia del COVID-19, los autores introducen un modelo que divide la variable de personas infectadas en dos compartimentos: personas infectadas documentadas y no documentadas. Además, incorporan información de movilidad humana entre ciudades o provincias, obteniendo

el siguiente sistema:

$$\begin{aligned}
\frac{dS_i}{dt} &= -\frac{\beta_i S_i I_i^r}{N_i} - \frac{\mu_i \beta_i S_i I_i^u}{N_i} + \theta \sum_j \frac{M_{ij} S_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} S_i}{N_i - I_i^r} \\
\frac{dE_i}{dt} &= \frac{\beta_i S_i I_i^r}{N_i} + \frac{\mu_i \beta_i S_i I_i^u}{N_i} - \frac{E_i}{Z} + \theta \sum_j \frac{M_{ij} E_j}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} E_i}{N_i - I_i^r} \\
\frac{dI_i^r}{dt} &= \alpha \frac{E_i}{Z} - \frac{I_i^r}{D} \\
\frac{dI_i^u}{dt} &= (1 - \alpha) \frac{E_i}{Z} - \frac{I_i^u}{D} + \theta \sum_j \frac{M_{ij} I_j^u}{N_j - I_j^r} - \theta \sum_j \frac{M_{ji} I_i^u}{N_i - I_i^r} \\
\frac{dR_i}{dt} &= \rho \frac{I_i^r + I_i^u}{D} \\
N_i &= N_i + \theta \sum_j M_{ij} - \theta \sum_j M_{ji}.
\end{aligned}
\tag{2.3}$$

Aquí, S_i , E_i , I_i^r , I_i^u , R_i y N_i representan a la población susceptible, expuesta, infectada documentada, infectada no documentada, recuperada y población total de la ciudad i .

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 la severidad de sus síntomas fue diagnosticada por el sistema de salud), mientras que todas las otras personas infectadas (en principio no diagnosticadas) representan la variable I_i^u . La constante β_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 β_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 [Ruiyun et al] 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: α 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 de la infección y ρ la tasa de recuperación. Se debe aclarar que R_i representa la población que luego de pasar la enfermedad se recuperó y, en principio, tiene cierta inmunidad, al menos temporal. De acuerdo a los epidemiólogos, una vez que esta población cruce un umbral y represente la mayoría de la población total, la pandemia se da por controlada, al menos temporalmente.

Al igual que en los modelos SIR y SEIR, un parámetro fundamental de la epidemia es el número efectivo de reproducción R_e . En este modelo se lo calcula, para la ciudad/provincia i , como

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

Si este número se sostiene por encima de 1, la enfermedad seguirá expandiéndose. 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 θ es un factor más grande que 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.

3. IDENTIFICACIÓN DE PARÁMETROS

El modelo (2.3) posee una serie de parámetros que deben ser estimados de acuerdo a características de la enfermedad viral, así como formas de propagación en el país bajo estudio. Un objetivo central del trabajo que realiza el Centro de Modelización Matemática consiste en identificar los parámetros mediante técnicas de ajuste (*fitting techniques*), utilizando la información proporcionada por organismos oficiales del Estado. En particular, gracias a una estrecha colaboración con el Instituto Nacional de Salud Pública e Investigación - INSPI, organismo rector en temas de diagnóstico de enfermedades y epidemias en el Ecuador, hemos podido obtener información de primera mano para realizar tal estimación.

Estudios previos realizados con gran cantidad de información (ver [Ruiyun et al]), aunque en un contexto socio-cultural distinto, establecieron intervalos para los parámetros que deben ser observados en el modelo (2.3), cualquiera sea la realidad local. Estos intervalos son los siguientes:

- Tasa de transmisión de la población infectada reportada, β : $0,8 \leq \beta \leq 1,5$.
- Factor de reducción de la tasa de transmisión de población infectada no reportada, μ : $0,2 \leq \mu \leq 1$.
- Factor multiplicativo de corrección de los datos de movilidad de la población, θ : $1 \leq \theta \leq 1,75$.
- Tiempo promedio de días de latencia, Z : $2 \leq Z \leq 5$, días.
- Fracción de pacientes que desarrollaron síntomas severos, α : $0,02 \leq \alpha \leq 1$.
- Tiempo promedio de días de duración de la infección, D : $2 \leq D \leq 5$, días.

Como se mencionó previamente, la intención de nuestro estudio es utilizar el modelo (2.3), estimando los parámetros de la expansión del COVID-19 de acuerdo a la realidad ecuatoriana. La idea fundamental de las técnicas de ajuste utilizadas consiste en identificar los parámetros que minimizan la suma de las diferencias al cuadrado entre los datos observados y la solución correspondiente del modelo. Así, supongamos que los parámetros identificados están dados por $P := (\theta_1, \theta_2, \dots, \theta_m)$ y los datos observados por $Y_{t_i}^{obs} := (y_{t_1}^{obs}, y_{t_2}^{obs}, \dots, y_{t_n}^{obs})$, 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 $f(t_i, P)$. El problema, a grandes rasgos, entonces consiste en hallar el conjunto de parámetros \hat{P} dado por

$$\hat{P} := \arg \min \sum_{i=1}^n (f(t_i, P) - y_{t_i}^{obs})^2,$$

donde n es el número de datos disponible para hacer la inferencia. Hallada la solución de este problema, podemos considerar que la solución $f(t_i, \hat{P})$ da el modelo mejor ajustado a la serie temporal t_i , de acuerdo al criterio de mínimos cuadrados (ver [Chowell]).

Este esquema, sin embargo, tiene como inconveniente el no tomar en cuenta la incertidumbre en los parámetros. Considerando a estos, así como a la solución, como variables aleatorias con distribución gaussiana, podemos adoptar el paradigma Bayesiano para problemas inversos y estimar los parámetros buscando minimizar la probabilidad condicional *a-posteriori* de los parámetros, dadas las observaciones.

En el caso del SARS-CoV-2, asumiendo una sola provincia aislada e introduciendo el vector de parámetros $P := (\alpha, \beta, \mu, \rho)^T$, el problema inverso de identificación de parámetros está formulado mediante:

$$\begin{aligned}
 \min_P & \sum_{i,j} (I^r(t_i) - I_i^{obs}) C_{ij}^{-1} (I^r(t_j) - I_j^{obs}) + (P - P^b)^T B^{-1} (P - P^b) \\
 \text{s.t.} & \frac{dS}{dt} = -\frac{\beta S I^r}{N} - \frac{\mu \beta S I^u}{N}, \\
 & \frac{dE}{dt} = \frac{\beta S I^r}{N} + \frac{\mu \beta S I^u}{N} - \frac{E}{Z}, \\
 & \frac{dI^r}{dt} = \alpha \frac{E}{Z} - \frac{I^r}{D}, \\
 & \frac{dI^u}{dt} = (1 - \alpha) \frac{E}{Z} - \frac{I^u}{D}, \\
 & \frac{dR}{dt} = \rho \frac{I^r + I^u}{D}
 \end{aligned}
 \tag{3.1}$$

donde I_i^{obs} , $i = 1, \dots, n$, son las observaciones históricas de infectados reportados, $C = (C_{ij})$ corresponde a la matriz de covarianza de errores de observación y B es la matriz de covarianza de errores de fondo. Mientras que la matriz C es estimada partir de la forma, más o menos, adecuada de realizar el registro de infectados, la matriz B es obtenida, en nuestro caso, a partir de un esquema de aprendizaje *ensemble*, realizando perturbaciones aleatorias en los parámetros de referencia.

4. ESCENARIOS

En la primera etapa de este trabajo estimamos los parámetros necesarios para el modelo utilizando la información real obtenida de los entes respectivos del país en días anteriores y utilizando estos parámetros junto con condiciones iniciales actualizadas para obtener predicciones en distintos escenarios. Estos escenarios contemplan distintos porcentajes de la población en aislamiento social: desde el 0% hasta el 90%. Cabe mencionar que la estimación de los parámetros está hecha, en una primera etapa, de forma empírica a partir de los estimadores obtenidos en [Ruiyun et al]. Posteriormente, conforme se incorpora mayor información real, se procede a realizar el ajuste y la estimación de parámetros adaptados a la realidad nacional.

En la Figura 3 se pueden ver las predicciones para las personas infectadas en Pichincha, a partir del 22 de marzo de 2020. Como se aprecia en las figuras, por ejemplo, hacia el 27 de marzo con una cuarentena del 75% de la población, se reportarían entre 100 y 150 infectados.

En la Figura 4 se muestran las proyecciones de casos reportados (personas infectadas reportadas por las autoridades) en Guayas (izquierda) y Pichincha (derecha), para la semana del 23 al 28 de marzo de 2020.

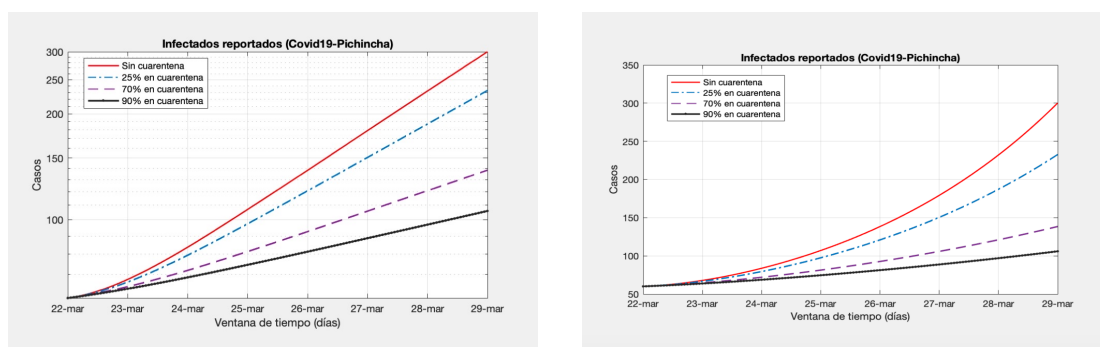


Figura 3: Personas infectadas reportadas, a partir del 22 de marzo de 2020 para la provincia de Pichincha: escala logarítmica (izq.) y escala lineal (der.).

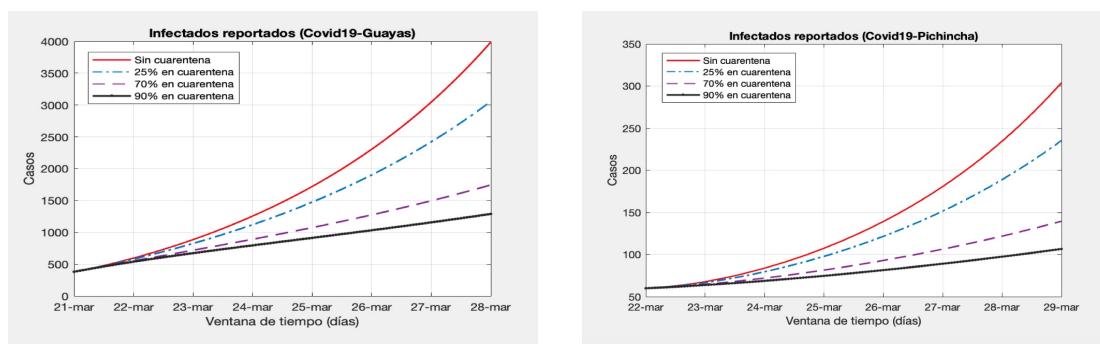


Figura 4: Proyección de personas infectadas reportadas en la semana del 23 al 28 de marzo de 2020, en escala logarítmica, para la provincia del Guayas (izq.) y la provincia de Pichinhca (der.).

En el caso de la provincia del Guayas se usaron los parámetros $\beta = 1,2$ y $\alpha = 0,3154$. El parámetro α , que muestra la proporción de personas infectadas que desarrollan síntomas severos, se obtuvo mediante técnicas de ajuste con la información oficial. Por otro lado, la tasa de capacidad de contagio β se ha aproximado de forma empírica, a partir de los parámetros del modelo de [Riuyun et al]. La condición inicial para esta simulación fue el dato oficial del 21 de marzo de 2020: 381 infectados. Con estos parámetros, se ha logrado una proyección acertada. Por ejemplo, para el 24 de marzo de 2020, con un 70 % de la población en cuarentena se esperaban 891 infectados, y el dato oficial fue de 807 infectados. Esta discrepancia es esperada y se puede explicar por la incertidumbre en el porcentaje real de la población aislada.

En el caso de la provincia de Pichincha los parámetros son $\beta = 1,2$ y $\alpha = 0,29$. Nuevamente, el parámetro α se obtuvo mediante técnicas de ajuste con la información oficial y β a partir de los parámetros mencionados por el modelo en [Riuyun et al]. La condición inicial para esta simulación fue el dato oficial del 22 de marzo de 2020: 60 infectados. En este caso, la proyección era que para el 24 de marzo de 2020, la cantidad de personas infectadas fuera de 71 con 80 % de la población en cuarentena y ese fue, precisamente, el dato oficial. En este caso, la incertidumbre

sobre el porcentaje de personas aisladas es menor, debido a que en Quito, particularmente, se han impuesto condiciones de cuarentena más duras que en otras ciudades del país.

De las simulaciones se puede apreciar el efecto positivo que tienen las medidas de aislamiento social, pero también que éstas no frenan el contagio definitivamente, si no se determina de forma temprana a la población infectada. La estrategia debe mudar hacia el aislamiento de todas, o de la gran mayoría, de las personas infectadas y al relajamiento de medidas de aislamiento paulatino para el resto de la población. Esto tendrá éxito siempre que se detecte eficiente y efectivamente a la mayor cantidad posible de infectados de forma temprana. Por este motivo, la recomendación sigue siendo mantener condiciones de aislamiento social, pero emprender campañas agresivas de diagnóstico (aplicación de pruebas) a toda persona con indicios de síntomas, por leves que parezcan.

Para reforzar lo señalado en el párrafo anterior, mostramos un posible escenario de salida de las medidas de aislamiento. En la Figura 5 se muestran las proyecciones del número de personas infectadas reportadas en la provincia de Pichincha en un único escenario de salida de la cuarentena: 15 días de cuarentena para el 90 % de la población, los siguientes 30 días la cuarentena aplica al 75 % de la población, y posteriormente solamente al 50 % de la población. Las distintas proyecciones (representadas por las distintas curvas en el gráfico) sin embargo dependen de un parámetro clave: μ , el factor de reducción de la tasa de contagio de la población infectada no reportada. El modelo, como quedó explicado, considera que las personas infectadas no reportadas presentan una tasa de transmisión atenuada de $\mu\beta$.

Así, la idea detrás de esta simulación es mostrar el efecto real que tiene la reducción de este factor mediante el aislamiento eficiente y efectivo de la población infectada: cuantos más test diagnósticos se realicen, más rápido será el aislamiento de los enfermos y se romperá la cadena de contagio respectiva. En conclusión, la población I_i^u se va reduciendo a cero. En la simulación se puede ver que manteniendo las políticas de aislamiento (cuarentena) igual: 15 días al 90 % de la población, 30 días más al 75 % y luego aislamiento al 50 %, pero sin reducir la tasa de contagio de las personas infectadas no reportadas ($\mu = 1$), el pico de la epidemia llegaría aproximadamente en 3 meses con cerca de 55.000 personas infectadas reportadas. Por otro lado, con la reducción de la tasa de contagio de las personas infectadas en un 50 % ($\mu = 0,5$), el pico de la pandemia se demoraría hasta 6 meses y tendría, aproximadamente, 30.000 enfermos reportados. El efecto de los test diagnósticos masivos es claro y esto está en concordancia con la recomendación de la OMS.

Esta simulación se hizo con los parámetros $\beta = 1,5$ y $\alpha = 0,23$. Este último parámetro se obtuvo por técnicas de ajuste con la información oficial.

5. RECOMENDACIONES

El objetivo principal de la lucha contra una epidemia es reducir el factor efectivo de transmisión de la enfermedad R_e . Este número depende de la misma naturaleza de la enfermedad y de la virulencia del virus. Sin embargo, las simulaciones muestran que hay factores complementarios en los cuales se puede incidir decisivamente para esta reducción: *aislamiento social y detección temprana*.

En un determinado escenario de aislamiento social, el principal factor de reducción es la

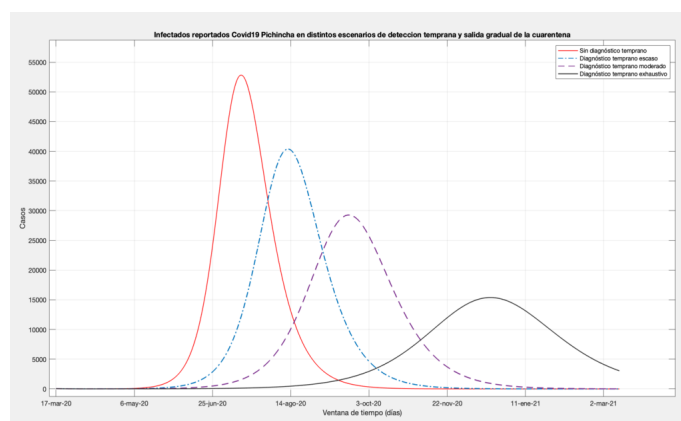


Figura 5: Escenario de salida de cuarentena progresiva con implementación de test diagnósticos diferentes para la provincia de Pichincha.

identificación temprana de la población infectada, reduciendo la variable I_i^u , y aislando estos pacientes para evitar la diseminación agresiva del virus. Esto solo es posible con la **aplicación masiva de pruebas diagnósticas** a todas las personas que presenten algún síntoma de enfermedad respiratoria y no esperar a que la gente desarrolle síntomas severos. Esto redundará en una reducción del contagio “ciego” que provocan las personas que tienen síntomas leves o inexistentes y, además, al aplicar tratamientos tempranos a los pacientes se puede incidir en la carga viral de esas personas, reduciendo la cantidad de personas en las unidades de cuidados intensivos.

REFERENCIAS

- [1] FRED BRAUER AND CARLOS CASTILLO-CHÁVEZ. (2001). *Mathematical Models in Population Biology and Epidemiology*. New York: Springer.
- [2] GERARDO CHOWELL. (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.
- [3] DI WU, TIAN TIAN WU, QUN LIU AND ZHICONG YANG. (2020). *The SARS-CoV-2 outbreak: what we know*, *International Journal of Infectious Diseases*, DOI: 10.1016/j.ijid.2020.03.004.
- [4] GÜNTER LAST AND MATHEW PENROSE. (2018). *Lectures on the Poisson Process*. Cambridge: Cambridge University Press.
- [5] STEPHEN A. LAUER, KYRA H. GRANTZ, QIFANG BI, FORREST K. JONES, QULU ZHENG, HANNAH R. MEREDITH, ANDREW S. AZMAN, NICHOLAS G. REICH, JUSTIN LESSLER. (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.

- [6] JOEL C. MILLER. (2017). *Mathematical models of SIR disease spread with combined non-sexual and sexual transmission routes*, Infectious Disease Modelling, DOI: 10.1016/j.idm.2016.12.003.
- [7] Organización Mundial de la Salud. (2020). Preguntas y respuestas sobre la enfermedad por coronavirus (COVID-19) <https://www.who.int/es/emergencias/diseases/novel-coronavirus-2019/advice-for-public/q-a-coronaviruses>
- [8] BENJAMIN RIDENHOUR, JESSICA M. KOWALIK Y DAVID K. SHAY. (2015). *El número reproductivo básico (R_0): consideraciones para su aplicación en la salud pública*, Pan American Journal of Public Health, DOI: 10.2105/AJPH.2013.301704.
- [9] RUIYUN LI, SEN PEI, BIN CHEN, YIMENG SONG, TAO ZHANG, WAN YANG AND JEFFREY SHAMAN. (2020). *Substantial undocumented infection facilitates the rapid dissemination of novel coronavirus (SARS-CoV2)*, Science, DOI: 10.1126/science.abb3221.