



INSTITUTO
NACIONAL DE
SALUD

Estimación del número reproductivo de la transmisión del COVID-19 en Colombia

Observatorio Nacional de Salud

$$\beta(t, \tau) = \phi(t) \omega(\tau)$$

$$\mathcal{R}_c(t) = \sum_{\tau=1}^{\infty} \omega$$

$$(+)\mathcal{R}(t - \tau) (2)$$

$$\mathcal{R}_c(t) = \sum_{\tau=1}^{\infty} \omega$$

$$(+)\frac{\mathcal{R}(t - \tau)}{\rho} (2')$$

Instituto Nacional de Salud
Observatorio Nacional de Salud

FERNANDO RUÍZ GÓMEZ
Ministro de Salud y Protección Social

MARTHA LUCÍA OSPINA MARTÍNEZ
Directora General INS

CARLOS ANDRÉS DURÁN CAMACHO
Secretario General

CARLOS ANDRÉS CASTAÑEDA-ORJUELA
Director Observatorio Nacional de Salud

OFICINA DE COMUNICACIONES INS

Equipo de trabajo

Gustavo Nicolás Páez-Salamanca
Camila Franco-Restrepo
José Fernando Zea
Érica Cruz-Rivera
Adolfo Quiroz-Salazar
Juan Felipe Cerón
Santiago Enrique Cortés
Gina Vargas-Sandoval
Carlos Castañeda-Orjuela

Edición

Juan Camilo Acosta

Diagramación

Claudia Clavijo

2020

Estimación del número reproductivo efectivo de la transmisión del COVID-19 en Colombia

Gustavo Nicolás Páez-Salamanca, Camila Franco-Restrepo, José Fernando Zea, Érica Cruz-Rivera, Adolfo Quiroz-Salazar, Juan Felipe Cerón, Santiago Enrique Cortés, Gina Vargas-Sandoval, Carlos Castañeda-Orjuela¹.

Observatorio Nacional de Salud, Instituto Nacional de Salud

Resumen

En este documento se muestra la metodología implementada por el Instituto Nacional de Salud para determinar el impacto del virus COVID-19 (SARS-CoV-2) en términos de número y tipos de casos, velocidad de transmisión y la cual ha permitido la simulación de escenarios para la toma de decisiones por parte del Gobierno Nacional. Este documento es la primera versión de los modelos y metodologías que el INS ha venido explorando los últimos meses con el fin de mejorar las estimaciones desde la ampliación del modelo SIR hacia un SEIR, pasando por otras metodologías de estimación de la tasa de transmisión, hasta la inclusión de parámetros relacionados con la movilidad y la geografía del contagio. Para este propósito, se hizo uso de los datos de la vigilancia epidemiológica recolectados en los municipios y departamentos del país, así como la implementación de adaptaciones innovadoras de modelos epidemiológicos referenciados en la literatura y que capturan las particularidades del comportamiento del virus en el país.

¹Gustavo Nicolás Páez-Salamanca ha trabajado en el Ministerio de Salud y Protección Social y ahora está vinculado al Myanmar Development Institute. Adolfo Quiroz-Salazar es profesor del departamento de matemáticas de la Universidad de los Andes. Juan Felipe Cerón y Santiago Enrique Cortés están vinculados a la Universidad de los Andes. Los demás autores son miembros del equipo de trabajo del Observatorio Nacional de Salud (ONS) del Instituto Nacional de Salud (INS). Agradecemos a Aval Digital Labs (ADL), su presidenta Ana Albir, a los líderes de este proyecto Tito Neira, Lina Osorio, Liliana Betancourt y sus equipos, quienes dispusieron la plataforma tecnológica para correr de manera más eficiente los códigos diseñados por el equipo técnico del ONS y han participado en la discusión y revisión de códigos para su optimización (disminuir tiempos de ejecución).

Introducción

El coronavirus 2019 (COVID-19) es una enfermedad infecciosa considerada por la Organización Mundial de la Salud (OMS) como un problema de salud pública. Es producida por el coronavirus-2 (SARS-CoV-2) (1), un virus de ARN monocatenario de sentido positivo perteneciente al género Betacoronavirus (2–4). Parece que los reservorios del COVID-19 son los murciélagos (5) y otros animales salvajes (3,6), pero aún no se ha logrado establecer potenciales hospederos intermedios (7). El brote inició en el mercado de mariscos de Hunan en Wuhan (China), donde a menudo se venden animales vivos como murciélagos, serpientes, ranas, entre otros (3,6,8).

El virus se transmite de persona a persona (9) por gotitas, contacto y fómites (10,11). Las estimaciones actuales indican que el COVID-19 tiene un período de incubación medio de tres días (rango de 0 a 24 días) y una alta tasa de transmisión, estudios recientes muestran que la transmisión se da por individuos sintomáticos como individuos asintomáticos (10,11).

La enfermedad presenta síntomas tales como fiebre, tos y mialgias, dolores musculares, o fatiga; es posible que días antes de la fiebre se presente diarrea y náuseas. También puede aparecer cefalea o hemoptisis (10,12) e incluso podría cursar sin síntomas (9). Los adultos mayores que presentan comorbilidad tienen más probabilidad de presentar insuficiencia respiratoria debido al daño alveolar severo que produce el virus (13). La enfermedad, inicialmente, puede mostrar una rápida progresión de leve a grave lo cual causaría insuficiencia orgánica multisistémica y en los casos críticos ocasionaría la muerte (12,14).

Cronología de la pandemia

En diciembre de 2019 se presentó un brote de neumonía en Wuhan (8) causada por el COVID-19, que luego se extendió por China y otros países, afectando considerablemente los sistemas de salud. El 30 de enero de 2020, la OMS declaró el brote como una emergencia de salud pública de preocupación internacional (15). Para limitar y controlar la transmisión se han recomendado medidas relacionadas con detección temprana,

diagnóstico precoz, tratamiento oportuno, seguimiento de posibles contactos y cuarentena, acciones que buscan bloquear la transmisión de persona a persona, así como reducir las infecciones secundarias entre contactos cercanos y trabajadores de la salud (16).

Los modelos epidemiológicos

Su uso es una herramienta imprescindible para el análisis de la transmisión de enfermedades emergentes y reemergentes, puesto que los modelos permiten estudiar los procesos de propagación, detección, prevención, control, así como de posibles tratamientos. En la literatura existe una gran variedad de modelos, el interés se centra en aquellos donde el agente infeccioso de la enfermedad es un virus. Los modelos básicos como SIR (nombre que proviene de las iniciales Susceptible, Infectado y Recuperado) o SEIR (Susceptible, Expuesto, Infectado, Recuperado) son ampliamente utilizados (17,18) por la simplicidad y eficiencia de sus resultados. Estos permiten entender el comportamiento de la enfermedad en una población y, de este modo, hacer la transición a modelos más complejos.

Las propiedades y procesos de la enfermedad COVID-19 son objeto de investigación a nivel mundial. Es claro que Colombia enfrenta grandes retos para la estimación de parámetros de un modelo representativo, que proporcione información para el diseño y ajuste de políticas de salud pública. Uno de los desafíos se debe al subregistro natural causado por la existencia de personas asintomáticas, o con síntomas leves que no consideran necesario reportar la enfermedad. Por otro lado, existe un rezago entre el día en que la persona empieza a sentirse enferma, el momento en el cual el caso es notificado y, finalmente, el día en que la persona es oficialmente diagnosticada con el virus. El segundo reto en Colombia emerge de la idiosincrasia del país, los usos y costumbres de la sociedad que determinan la frecuencia de contactos entre personas (y por ende la facilidad de transmisión del virus). Por otro lado, están las políticas públicas de cuarentena que en mayor o menor medida han sido tomadas en los distintos entes territoriales y las cuales también afectan los patrones de transmisión del virus.

Modelos y métodos utilizados por el INS

El documento se divide en las siguientes secciones: la primera describe la estructura de un modelo SIR con el que se realizan las predicciones para el país y las regiones (Andes Occidentales, Andes Orientales, Caribe, Pacífico, Amazonía y Llanos); en la segunda, se profundiza en la corrección empírica del número de casos generados a partir de métodos de corrección de incurridos y no reportados conocidos como método de Chain Ladder o triángulo de siniestros. A continuación, se detallan dos métodos con los que se está estimando la tasa de transmisión $\mathcal{R}(t)$ en el país. Finalmente, en las conclusiones se presentan los futuros desarrollos.

1. Modelo SIR (Susceptible-Infectado-Recuperado)

Es un modelo compartimental determinístico en donde la población afectada se divide en tres grupos: “S”, el grupo de individuos susceptibles o que pueden contagiarse; “I”, el grupo de individuos infectados o que son capaces de transmitir la enfermedad; y “R”, el grupo de individuos recuperados de la infección (en este modelo el supuesto es que los recuperados adquieren inmunidad) (18). EL modelo SIR se describe mediante el siguiente sistema de ecuaciones diferenciales ordinarias:

Donde β representa la tasa de transmisibilidad que se define como $\beta(t)=\gamma \mathcal{R}(t)$; la cual contiene, el parámetro γ que representa el inverso multiplicativo del periodo de infección de la enfermedad, el cual se estima como 5,8 días (IC 95% 4,3-7,5 días) (19), y $\mathcal{R}(t)$ es la tasa efectiva de reproducción. Desde la visión de política pública, $\mathcal{R}(t)$ es fundamental para comprender la evolución de una epidemia dado que tiene como objeto medir cuántos individuos en estado susceptible pueden ser infectados por una persona enferma durante el tiempo en que esta se encuentra en un estado infeccioso. Este valor incluye tres componentes (20):

1. Duración del periodo de infección.
2. Probabilidad que la infección pase de una persona infecciosa a una persona susceptible.
3. Número promedio de veces, por unidad de tiempo, que un individuo infeccioso se encuentra a un individuo susceptible durante el día.

Como se ampliará más adelante, el $\mathcal{R}(t)$ se estima por fuera del modelo, actualmente a través de dos métodos de series de tiempo que serán explicadas en la tercera parte de este documento. Igualmente, como se explicará a continuación para la estimación de la tasa de transmisión, se realiza un ajuste sobre los datos de número de casos reportados con el fin de corregir el subregistro por medio del método conocido como Chain Ladder.

$$\frac{dS}{dt} = -\beta SI$$

$$\frac{dI}{dt} = \beta SI - \gamma I$$

$$\frac{dR}{dt} = \gamma I$$

2. Ajuste de Chain Ladder o triángulo de siniestros

En esta sección se describe el tratamiento que se da a los datos de entrada del número de casos reportados para hacer una estimación adecuada de los casos sintomáticos que aparecen en un día determinado. El método usado, denominado “Chain Ladder”, o triángulo de siniestros, usa información conocida sobre el tiempo estimado que toma un caso desde que se genera (es decir, desde que aparece un paciente sintomático) hasta que se reporta (bien sea por notificación o por diagnóstico) y lo usa para ajustar aquellos valores donde es posible que exista todavía un rezago en el registro.

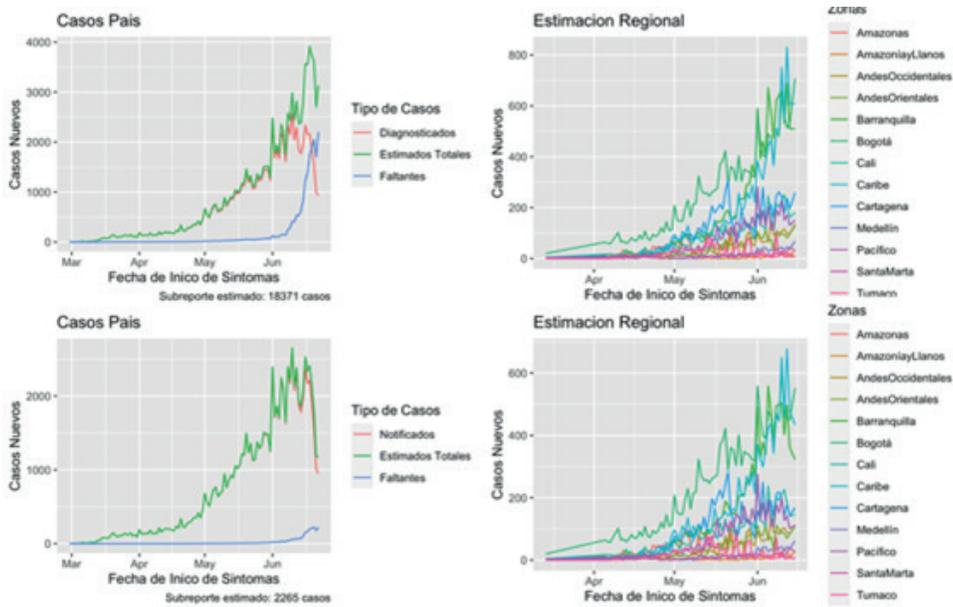
Este método es usado ampliamente en la actualidad con el objetivo de corregir aquellos siniestros incurridos, pero no reportados (IBNR por sus siglas en inglés *-incurred but not reported*). Esta técnica se basa en la elaboración de triángulos de “contingencia” bidimensionales. En este caso particular, tienen como objetivo saber el potencial de contagiados, considerando las diferencias entre los tiempos de aparición de síntomas, la fecha de notificación y la fecha de diagnóstico, por tanto, usando la información disponible, las dimensiones serán la fecha de inicio de síntomas (filas) y los días de diferencia (o rezago) respecto a la fecha de notificación o la fecha de diagnóstico (columnas).

Usualmente, los datos iniciales de reporte tienen problemas de calidad del dato, la cual se da por la diferencia entre la primera fecha en la que se registra síntomas y el día en que se presenta el diagnóstico o notificación, por lo cual se realiza una ponderación a los valores más antiguos (más rezagos), la cual es exponencial de 0,95, esto significa que las proporciones acumuladas del periodo t , van a tener un peso que es el 95% del peso de las proporciones del periodo $t+1$.

La Figura 1 representa los resultados del método de ajuste realizado por medio de fecha de diagnóstico y fecha de notificación del caso. Las técnicas presentan ajustes distintos, dado que ambos rezagos están influenciados por distintos

factores. De forma particular, se sugiere que un ajuste vía fecha de notificación presenta una cota inferior al número de casos esperados, mientras que el de diagnóstico exhibe una cota superior. Esto se debe a que en principio los casos se notifican rápido (a los siete días del evento más de la mitad de los casos ya han sido notificados). Sin embargo, los reportes de vigilancia y laboratorio en salud pública del INS demuestran cómo muchos de los casos sospechosos son negativos al virus, por lo que solo observando el número de casos notificados requiere ajustes adicionales. De otro lado, el diagnóstico por laboratorio toma tiempo, y entre más gente esté infectada, más se demorará el diagnóstico dada la capacidad logística del país. No obstante, todos estos casos son por definición casos verdaderos. Por este motivo, y por ser una cota superior bajo el principio de precaución, en este estudio se usa, y en general se recomienda, el uso del ajuste vía casos diagnosticados.

Figura 1 Ajustes de casos del Covid-19, según notificación y diagnóstico, a corte de 30 de junio de 2020



En consecuencia, es importante notar que el ajuste tiene sentido después de algunos días ya que los valores más recientes presentan una alta volatilidad. Por tal motivo, no se recomienda ajustar los 7 días previos a la fecha actual. Debido a esto, se propone ajustar al 15 de junio y no a la última fecha de actualización de este reporte (22 de junio).

3. Estimación de la tasa efectiva de reproducción ($\mathcal{R}(t)$)

Considerando lo anteriormente mencionado, una de las particularidades del modelo es la estimación “por fuera del SIR” y *a priori* de la tasa efectiva de reproducción. Esta tasa se convierte en uno de los valores más importantes de política pública, puesto que cuando se estima la tasa efectiva esta permite evaluar el efecto de las medidas de política, y tener insumos para la predicción de posibles casos futuros. Como se señalaba anteriormente la tasa efectiva muestra cuantos individuos susceptibles pueden ser infectados por una persona enferma siendo una medida de la “velocidad de contagio” si, por ejemplo, la tasa es igual a 2, significa que un enfermo infectaría a dos personas susceptibles. Intuitivamente, si la tasa es inferior a 1 se entenderá que la enfermedad no se está propagando porque cada enfermo “contagiaría” a menos de una persona. Por esta razón, este indicador permite evaluar el éxito de las medidas de política, es decir, una cuarentena estricta como la que tuvo lugar en Colombia las dos semanas siguientes al 25 de marzo, debería disminuir la velocidad de contagio en la medida que el aislamiento evita la transmisión de los enfermos (incluyendo los asintomáticos). Efectivamente, durante el periodo en mención el $\mathcal{R}(t)$ alcanzó valores medios de 1,1 que luego fue aumentando a medida que la medida se relajó, indicando que en un período de tiempo la velocidad de contagio disminuyó significativamente.

La tasa de reproducción puede estimarse como uno de los parámetros calibrados del modelo teóricamente. Actualmente, el Instituto Nacional de Salud está trabajando en un modelo alternativo, el cual es un modelo epidemiológico compartimental a corto plazo para el estudio de COVID-19 en Colombia, este incluye aspectos que no han sido considerados en los modelos presentados en este documento. Para los análisis actuales y considerando la necesidad de encontrar el número de transmisión en el país, se realiza una aproximación a través de dos modelos basados en series de tiempo. En la primera parte, se realizará una conceptualización de la tasa de reproducción y en la segunda, se describirán los modelos.

3.1 Conceptualización de la tasa de reproducción

Esta descomposición de $\mathcal{R}(t)$, descrita anteriormente, permite evidenciar que la tasa de reproducción es influenciada por factores biológicos de la enfermedad (los cuales determinan por cuanto tiempo la persona es infecciosa y parte de la facilidad en que la enfermedad se contagie entre personas) y factores sociales (que estiman el tipo y frecuencia de las interacciones entre personas). De esta forma, se derivan elementos relevantes que se tienen en cuenta en la modelación de la tasa de reproducción. Por un lado, la tasa varía según el contexto y la idiosincrasia de la población de estudio. Debido a esto, extrapolar valores a zonas geográficas distintas puede ser inapropiado. Por otra parte, la variación también es temporal y es afectada por cambios en la población susceptible y decisiones gubernamentales tales como políticas de cuarentena, distanciamiento social, entre otras.

Autores como Fraser (2007) (21) y Cori et al. (2013)(22) proponen diferenciar dos tipos de tasas de reproducción. Una, identificada como la tasa de reproducción efectiva $\mathcal{R}(t)$, la cual está asociada al número de personas que empiezan su infección en el día t e infectarán si los tres factores previamente mencionados se mantienen invariantes durante el periodo de infección. Otra, es la tasa de reproducción de caso, \mathcal{R}_c , la cual incorpora los cambios que tienen estos tres factores durante el periodo de tiempo en el cual el individuo es infeccioso. De esta forma, el primer indicador tiene un énfasis prospectivo y de característica contrafactual a partir del cual se plantean escenarios tales como: *ceteris paribus* la tasa de reproducción es \mathcal{R} y, por ende, si no se ejercen cambios el número de infectados que se espera es X . Por otro lado, el \mathcal{R}_c presenta un enfoque retrospectivo y de monitoreo a partir del cual se puede visualizar de forma directa la influencia que ejercen los cambios sociales (usualmente impulsados por políticas públicas) en el desarrollo de la enfermedad. Ambos indicadores permiten entender si la epidemia se está expandiendo (cuando estos indicadores son mayores a 1) o se está reduciendo (cuando estos indicadores son menores a 1).

Las siguientes secciones presentan los modelos mediante los cuales estos indicadores se van a medir. Naturalmente, entre modelos la definición matemática presenta pequeñas variaciones. Sin embargo, todas estas se realizan en pro de medir los mismos conceptos bajo diferentes conjuntos de supuestos.

3.2 Modelos de serie de tiempo

En esta sección se presentan dos modelos basados en los cálculos originales desarrollados en el artículo de investigación de Kermack-McKendrick (1927)(23). Dicha estrategia de modelación se elige dado que sus elementos son intuitivos (21) y han sido sistemáticamente usados para la evaluación de las epidemias más recientes (24). Así, el principal modelo de referencia de esta sección es el planteado por Thompson et al. (2019) (24) el cual sintetiza el estado actual de esta serie de modelos. Debido a la forma de estimación, se denomina modelo Bayesiano. Una vez que las componentes principales se describen, se pasará a discutir cuales son las limitaciones de este modelo y en consecuencia a proponer uno alternativo, el cual se considera como un Modelo Estado-Espacio. En el mismo sentido, se analiza las ventajas y las limitaciones del mismo.

El primer modelo, Modelo Bayesiano, se está utilizando para la estimación de la tasa de reproducción nacional, mientras que el Modelo Estado-Espacio es usado para la estimación de la tasa de reproducción regional-distrital. Siendo el primero apropiado para la estimación nacional, pues este asume que no hay migración y se cuenta con un buen número de datos. El segundo modelo

permite realizar ejercicios de migración a través de las matrices de varianzas y covarianzas, y es más robusto frente a la ausencia de datos.

3.2.1. Modelo Bayesiano

Para propósitos de esta sección $I_N(t)$ es el número de individuos que empiezan a ser infecciosos en el período t , e $I(t)$ es el número total de individuos infecciosos en este periodo. En el modelo $I_N(t)$ se considera un valor estocástico (varía con el tiempo o evoluciona en función de otra variable) que depende del número de personas infecciosas en periodos anteriores, ajustado por la diferencia entre los periodos a considerar. De esta forma, se define la ecuación de renovación (*renewal*) como:

$$E[I_N(t)] = \sum_{\tau=1}^{\infty} \beta(t, \tau) I_N(t - \tau) \quad (1)$$

Donde el factor de ajuste es $\beta(t, \tau) = \phi(t) \omega(\tau)$.

En este caso el factor $\phi(t)$ está relacionado con las características sociales vinculadas a la tasa de reproducción, y $\omega(\tau)$ a el intervalo serial de la enfermedad, se define como la densidad de

probabilidad del número de personas que empiezan a ser infecciosas, τ unidades de tiempo después de que la persona que le infectó comenzó a ser infecciosa.

A partir del análisis sugerido por Wallinga y Teunis (2004)(25) y Fraser (2007)(21), se tiene la ecuación (2) el cual se amplía técnicamente y con los diferentes supuestos en el Anexo 2.

$$\mathcal{R}_c(t) = \sum_{\tau=1}^{\infty} \omega(t) \mathcal{R}(t - \tau) \quad (2)$$

Según el análisis de Fraser (2007) (21), las principales ventajas de este modelo son el desarrollo accesible de los cálculos y la relativa simplicidad para encontrar los

intervalos seriales. Sin embargo, el cálculo de estos genera problemas conceptuales para la estimación del $\mathcal{R}(t)$. Tal como lo plantea el autor, el cálculo del intervalo serial se ve afectado por las normas sociales. Debido a esto, presumir intervalos seriales fijos en escenarios donde las

interacciones entre personas son afectadas por políticas públicas, pueden afectar la capacidad del $\mathcal{R}(t)$ de capturar estos cambios. Adicional al problema conceptual, la necesidad de determinar una ventana de estimación w , genera una suavización forzada en la cual diferentes períodos se promedian. Por esta razón, $\mathcal{R}(t)$ no captura exclusivamente la capacidad de infección de un día determinado, sino que incluye, por construcción, una combinación de estos valores en los periodos anexos.

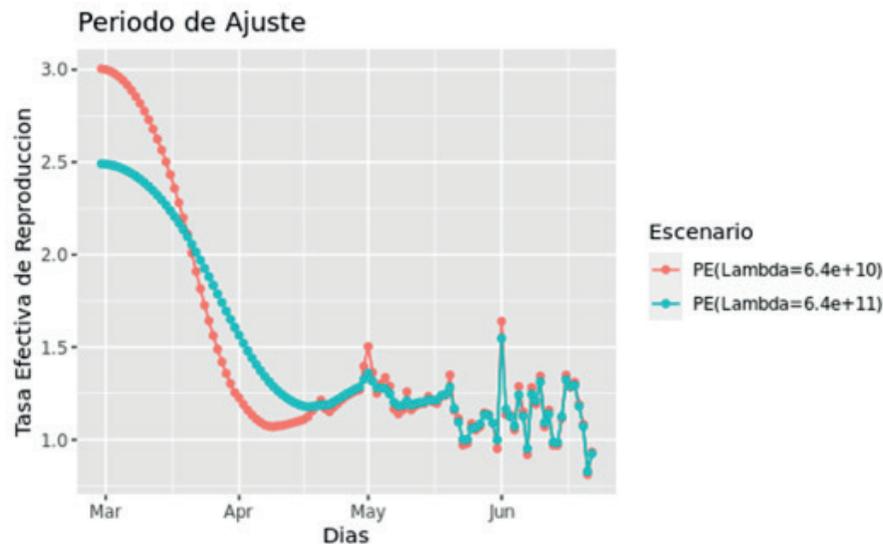
Como se señaló anteriormente con este modelo se realiza la estimación de la tasa efectiva de reproducción. Los valores calculados se suavizan con el objetivo de evitar el sobreajuste a través de un programa de optimización que puede detallarse en el Anexo 2. El parámetro clave sobre el cual se ajusta el suavizamiento es el parámetro λ , el cual en la figura 2 toma dos valores. Un λ menor ajusta mejor los datos respecto a los casos ocurridos. Sin embargo, esto se da porque el modelo presenta un sobreajuste

a la serie y eso hace que sea muy sensible a errores de medición, los cuales son altamente probables (dentro de rangos razonables), dada la característica asintomática del virus. Por otro lado, si bien aumentar el suavizamiento hace el modelo robusto al ruido numérico, también hace que cambios del parámetro en el tiempo no sean visibles. A manera de ejemplo, si $\lambda \rightarrow \infty$, el modelo predice el $\mathcal{R}(t)$ promedio del período, lo cual no permite visualizar cambios de política. Ahora bien, a partir del criterio de expertos y los diferentes

experimentos de estimación realizados, se considera que el valor de λ adecuado a usar está entre $6,4 \times 10^{10}$ y $6,4 \times 10^{11}$.²

De acuerdo con la Figura 2, la tasa efectiva de transmisión nacional inició en niveles cercanos a los 2,5, durante la cuarentena nacional que empezó la última semana de marzo el promedio estuvo en 1,11, y se inició un cambio de tendencia entre la primera y segunda semana de abril. Desde el 27 de abril, se estima un promedio de 1,19. En los últimos periodos se presenta alta volatilidad de la serie³ que oscila alrededor de una pendiente positiva, por lo cual prefiere hacerse una interpretación de un promedio más amplio desde el 27 de abril.

Figura 2. Periodo de ajuste y estimación de la tasa efectiva de reproducción $R(t)$ a corte de 30 de junio de 2020.



3.2.2. Modelo Estado - Espacio

Buscando solucionar las limitaciones planteadas en el modelo anterior, se considera el modelo de Estado-Espacio que redefine el $\omega(\tau)$ con el fin de que capture únicamente los componentes biológicos de la enfermedad

y luego, a partir de la estimación estadística Bayesiana, se conduzca a una estimación basada en series de tiempo con el propósito de evitar los problemas generados por la ventana de estimación.

A partir de esta redefinición, que puede observarse en el Anexo 2 en detalle, se obtiene la siguiente ecuación donde $\mathcal{R}_c(t)$ se determina a través de la ecuación (2) ajustada por Ω que corresponde a el número de días que la persona puede infectar.

² Actualmente, el INS está realizando ajustes para tener un λ optimizado.

³ Esto se da principalmente por la calidad de los datos recién reportados.

$$\mathcal{R}_c(t) = \sum_{\tau=1}^{\infty} \omega(\tau) \frac{\mathcal{R}(t-\tau)}{\Omega} \quad (2')$$

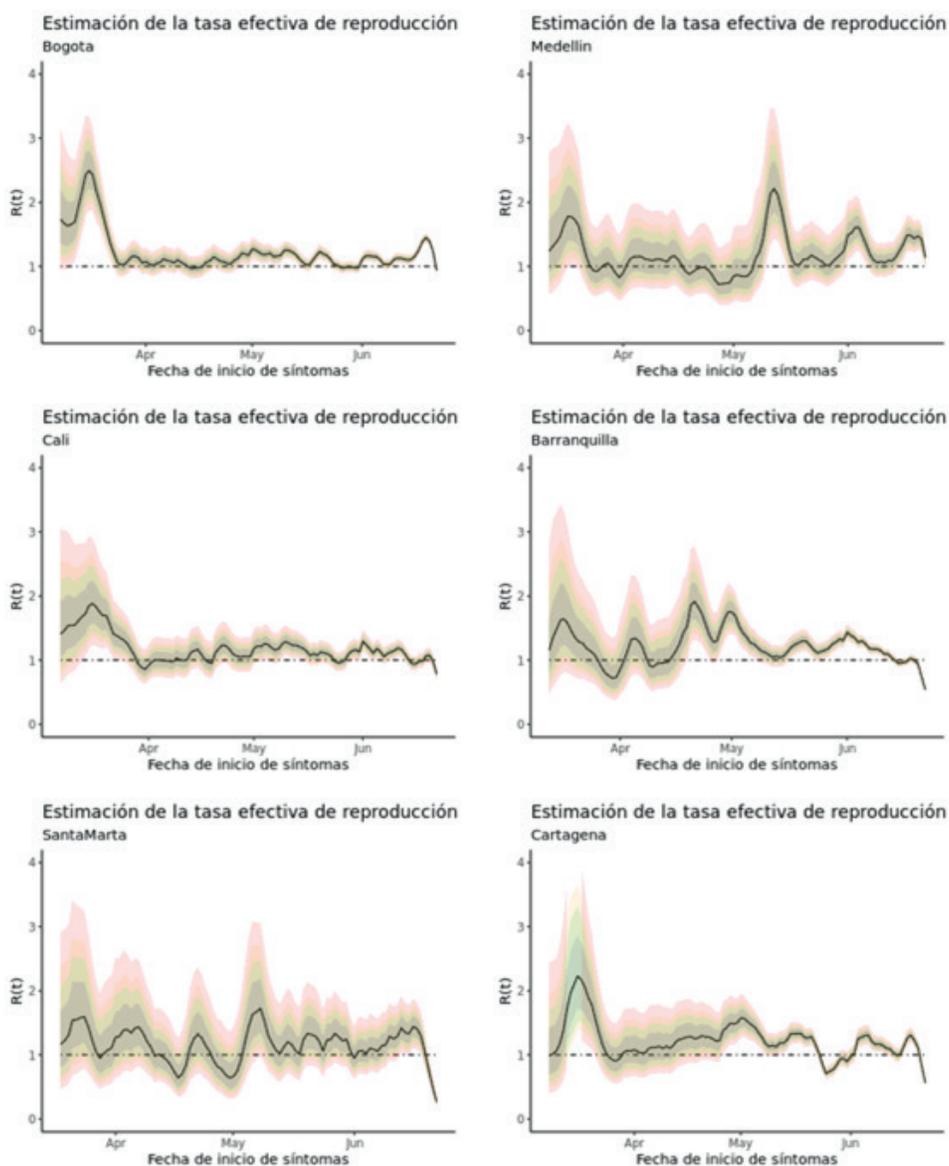
Esta estimación logra resultados similares a el modelo Bayesiano, corrigiendo los dos problemas teóricos mencionados anteriormente. Ahora bien, este proceso trae consigo un nuevo reto ya que cambia el requisito del intervalo serial por el requisito de la distribución de los tiempos de incubación y de infección. Por tal motivo, si bien desde la teoría este modelo presenta mejoras con respecto al modelo anterior, en la práctica su elección está condicionada a información que se tenga sobre estas características de la enfermedad.

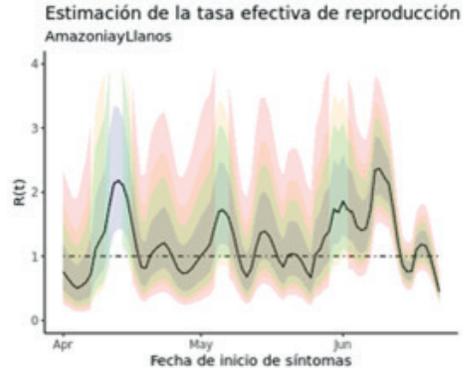
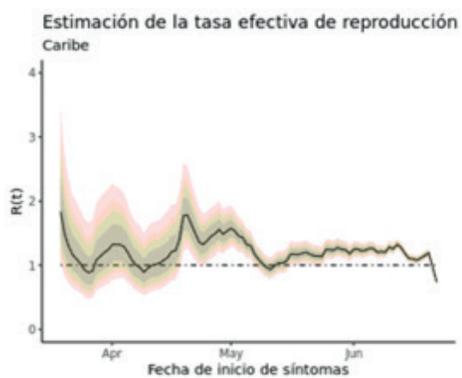
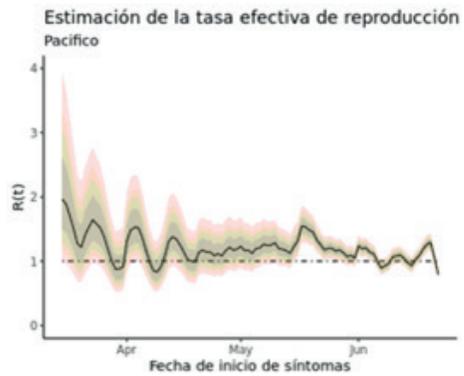
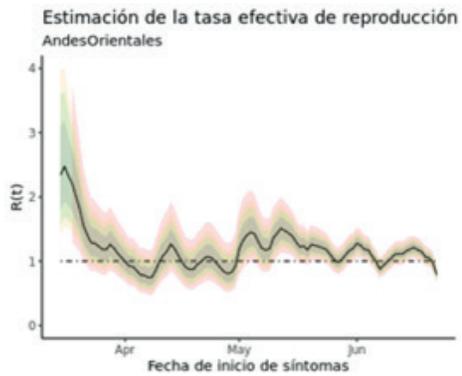
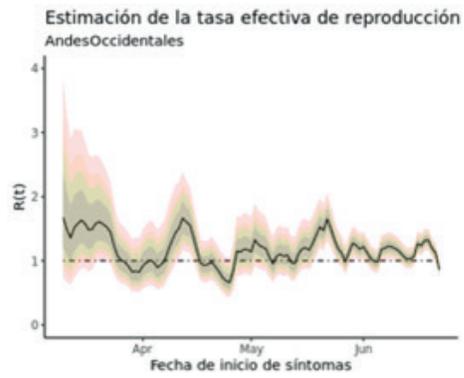
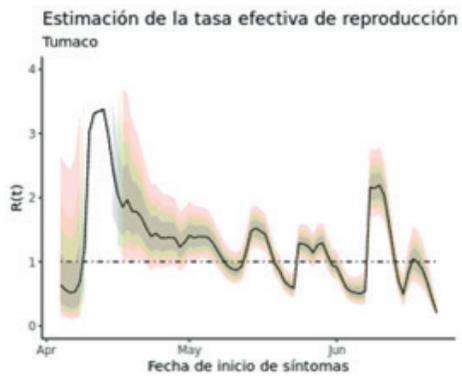
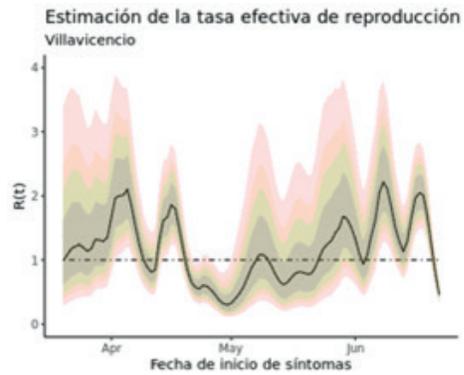
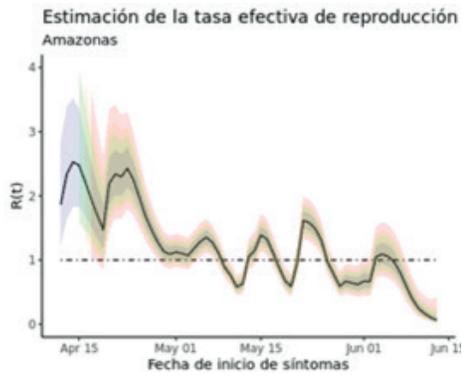
En la Figura 3 y en la Tabla 1, se pueden observar las estimaciones hasta el 30 de junio para las regiones y los distritos elaboradas bajo esta metodología. A modo de ejemplo, Bogotá durante las dos primeras semanas de marzo alcanzó una tasa efectiva de reproducción cercana a 2,4. En la última semana de marzo y primera de abril, después de la medida de cuarentena local y nacional esta tasa decae drásticamente. Ahora, durante el período comprendido entre abril y junio, la tasa efectiva de reproducción se ha mantenido por debajo de 1,5. Este valor ha aumentado en la última semana de junio lo cual se ve reflejado en el actual aumento de contagios. Observamos que las series de Bogotá y Cali tiene un comportamiento similar con una variabilidad pequeña después de darse el pico inicial, en comparación con series de otras ciudades-regiones que muestran una mayor variabilidad en la tasa efectiva de reproducción. Es de notar que Amazonas y Llanos registran altos niveles de volatilidad superando en ocasiones valores de 2,0 en la tasa efectiva de reproducción a lo largo del período.

Tabla 1 $R(t)$ promedio estimado desde el 27 de abril para zonas de Colombia, a corte de 30 de junio de 2020.

Zona	Promedio del $R(t)$ desde 27 de abril
Andes Occidentales	1,18
Andes Orientales	1,18
Caribe	1,23
Pacífico	1,17
Amazonia y Llanos	1,27

Figura 3 Estimación tasa efectiva de reproducción $R(t)$ por regiones a corte de 30 de junio de 2020.





Conclusiones y desarrollos futuros

El presente documento evidencia los métodos que se usan para la estimación de la tasa de reproducción, en consecuencia, el número de casos infectados que explica el comportamiento del contagio de casos de la pandemia Covid-19. Esto se realiza a través de un SIR que toma un parámetro $\beta(t)$ que proviene del cálculo de la tasa de reproducción efectiva, la cual se aproxima por el método Bayesiano a nivel nacional y el modelo Estado-Espacio a nivel regional, como se ve representado en los gráficos.

El uso de modelos matemáticos para el control de las epidemias y pandemias ha fortalecido, significativamente, el desarrollo de políticas públicas informadas. En particular, se resalta la capacidad de estos modelos para estructurar las diferentes variables de interés en el sistema y generar estimaciones de subregistro. De esta forma, se pueden obtener mejores aproximaciones a la situación real de la pandemia y predecir el comportamiento de esta. Por ejemplo, la predicción de la tasa efectiva de reproducción permitió establecer que la cuarentena estricta de finales del mes de marzo redujo significativamente la velocidad de contagio, que las posteriores medidas de relajamiento aumentan suavemente esta velocidad, y que pese a la volatilidad de los datos de las últimas semanas de junio se observa una leve tendencia positiva alertando sobre la dinámica de transmisión. Las estimaciones regionales también permiten a gobiernos locales y nacional observar el efecto diferencial de las medidas que se han tomado. Puede concluirse que los patrones de contagio son significativamente diferentes encontrando regiones donde la velocidad de contagio se ha controlado y mantenido en niveles “bajos” en el periodo de análisis como Bogotá y Cali, mientras muchas otras registran valores tan volátiles que sugieren que las medidas no tuvieron los impactos esperados como es el caso de Tumaco y Villavicencio.

Finalmente, es importante mencionar que si bien la tasa de reproducción efectiva es uno de los parámetros más importantes de análisis de política no es el único en el análisis de la dinámica de contagio. El INS en su trabajo de monitoreo

observa con el mismo interés indicadores como: el tiempo de duplicación de infectados por fecha de inicio de síntomas y tiempo de duplicación de muertes reportadas, la proporción de positividad de muestras diarias procesadas, y el nivel de ocupación diaria de camas UCI y de hospitalización.

Desarrollos futuros

En un ejercicio simultáneo y constante a la observación y monitoreo de los indicadores mencionados anteriormente, el INS trabaja en paralelo sobre el modelo epidemiológico compartimental SEIR ampliado, en el que se extiende la categoría de “Infectado” por los niveles de severidad de la enfermedad, es decir, se observa el número de infectados leves, moderados y críticos con sus respectivos parámetros los cuales son calibrados dentro del modelo partiendo de valores iniciales provenientes de la literatura. Considerando este modelo también se realiza la estimación analítica del $\mathcal{R}(t)$, es decir, la tasa no se calcula con métodos estocásticos sino se estima en función de los parámetros del SEIR.

Igualmente, se trabaja en otros dos modelos de simulación estocástica para calcular el $\mathcal{R}(t)$. Los dos modelos planteados en este documento coinciden en calibrar un proceso de Poisson a partir del cual se modela el proceso de contagio, en contraste, los modelos que se están desarrollando en paralelo modelan el proceso trasladando el reto estadístico a uno probabilístico donde la capacidad computacional es determinante. El primer modelo se enfoca en la descripción del proceso de contagio desde el punto de vista del evento, asumiendo una distribución Poisson y el segundo busca resolver el problema de capacidad computacional a través de soluciones analíticas (asumiendo una distribución exponencial sobre el proceso de incubación e infección).

Por otro lado, la literatura ha sugerido incluir el factor movilidad en los modelos SEIR de metapoblación para robustecer la toma de decisiones sobre medidas de aislamiento. Estos modelos utilizan con éxito en el estudio espaciotemporal de enfermedades como el COVID-19 y están compuestos por dos tipos de dinámica: la

dinámica de la enfermedad gobernada por el modelo compartimentado elegido, SIR o SEIR, y la matriz de movilidad de los individuos a través de las sub-poblaciones (localidades, municipios o regiones) que conforman todo el sistema de meta-población, el país, la cual está en exploración. Este último ingrediente, la movilidad, permite analizar la propagación entre subpoblaciones. Para cada población se simula un modelo epidemiológico (SIR, SEIR) en el que la probabilidad de transmisión del virus considera la afluencia de individuos de otras comunidades en donde la infección ya está presente, por lo que cada población se caracteriza por su transmisión interna y conectividad con otras comunidades. Para esto, actualmente, el INS se encuentra explorando fuentes potenciales de información.

Bibliografía

1. Gorbalenya AE, Baker SC, Baric RS, Groot RJ de, Drosten C, Gulyaeva AA, et al. Severe acute respiratory syndrome-related coronavirus – The species and its viruses, a statement of the Coronavirus Study Group. *bioRxiv*. 2020;2020.02.07.937862.
2. Chan JFW, Kok KH, Zhu Z, Chu H, To KKW, Yuan S, et al. Genomic characterization of the 2019 novel human-pathogenic coronavirus isolated from a patient with atypical pneumonia after visiting Wuhan. *Emerg Microbes Infect*. 2020;9(1):221–36.
3. Lu R, Zhao X, Li J, Niu P, Yang B, Wu H, et al. Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet*. 2020;395(10224):565–74.
4. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*. 2020;727–33.
5. Lam TT-Y, Shum MH-H, Zhu H-C, Tong Y-G, Ni X-B, Liao Y-S, et al. Identification of 2019-nCoV related coronaviruses in Malayan pangolins in southern China. *bioRxiv*. 2020;2020.02.13.945485.
6. Zhang L, Shen F, Chen F, Lin Z. Origin and evolution of the 2019 novel coronavirus. *Clin Infect Dis*. 2020;(ML):2019–20.
7. Li X, Zai J, Zhao Q, Nie Q, Li Y, Foley BT, et al. Evolutionary history, potential intermediate animal host, and cross-species analyses of SARS-CoV-2. *J Med Virol*. 2020 Jun;92(6):602–11.
8. Epidemiology Working Group for NCIP Epidemic Response CC for DC and P. [The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) in China]. *Chinese J Epidemiol*. 2020;41(02):145–51.
9. Chan JFW, Yuan S, Kok KH, To KKW, Chu H, Yang J, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. 2020;395(10223):514–23.
10. Guan W-J, Ni Z-Y, Hu Y, Liang W-H, Ou C-Q, He J-X, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. *N Engl J Med*. 2020;1–13.
11. Zhou P, Yang X-L, Wang X-G, Hu B, Zhang L, Zhang W, et al. A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature*. 2020;
12. Wang D, Hu B, Hu C, Zhu F, Liu X, Zhang J, et al. Clinical Characteristics of 138 Hospitalized Patients with 2019 Novel Coronavirus-Infected Pneumonia in Wuhan, China. *JAMA - J Am Med Assoc*. 2020;1–9.
13. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet*. 2020 Feb;395(10223):507–13.
14. Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet (London, England)*. 2020 Feb;395(10223):497–506.
15. World Health Organization. IHR Emergency Committee on Novel Coronavirus (2019-nCoV). 2020.
16. World Health Organization. Coronavirus disease 2019 (COVID-19): situation report—30. 2020.
17. Uribarri S, Rodríguez Meza M, Cervantes Cota J. Las matemáticas de las epidemias: caso México 2009 y otros. *Cienc Ergo Sum*. 2013;20(3):238–46.
18. Brauer F, Castillo-Chavez C, Feng Z. Correction to: *Mathematical Models in Epidemiology*. Springer; 2019.
19. Chen T, Rui J, Wang Q, Zhao Z, Cui J-A, Yin L. A mathematical model for simulating the transmission of Wuhan novel Coronavirus. *Infect Dis Poverty*. 2020;9(24):1–8.
20. van den Driessche P. Reproduction numbers of infectious disease models. *Infectious Disease Modelling*. 2017.
21. Fraser C. Estimating individual and household reproduction numbers in an emerging epidemic. *PLoS One*. 2007;
22. Cori A, Ferguson NM, Fraser C, Cauchemez S. A new framework and software to estimate time-varying reproduction numbers during epidemics. *Am J Epidemiol*. 2013;
23. A contribution to the mathematical theory of epidemics. *Proc R Soc London Ser A, Contain Pap a Math Phys Character*. 1927;
24. Thompson RN, Stockwin JE, van Gaalen RD, Polonsky JA, Kamvar ZN, Demarsh PA, et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics*. 2019;
25. Wallinga J, Teunis P. Different epidemic curves for severe acute respiratory syndrome reveal similar impacts of control measures. *Am J Epidemiol*. 2004;
26. Osborn DR, Harvey AC. *Forecasting, Structural Time Series Models and the Kalman Filter*. *Economica*. 1991;

ANEXOS

Anexo 1. Método Chain Ladder- IBNR

Para el ajuste por Chain Ladder, los dominios del triángulo estarán dados por la diferencia entre la primera fecha en la que se registran síntomas de toda la base y el último día en el que se presentan diagnósticos o notificaciones. Partiendo de una matriz con la siguiente información, la fecha en la que fueron reportados los síntomas, el tiempo en el que se demoraron en notificar (o diagnosticar) y el número de personas que se cuentan en ambos criterios, como se muestra en la Figura 3. Por ejemplo, una persona presentó síntomas el 15 de abril y notificó al día siguiente (o fue diagnosticado según sea el caso), mientras que otra persona que presentó síntomas el 27 de febrero notificó a los dos días.

Figura 4 Ejemplo de tabla de reporte

sintomas	tiempoNotificar	Freq
2020-04-15	1	1
2020-02-27	2	0

De esta matriz larga (*long*) se crea una tabla de contingencias (*wide*), teniendo en las filas los períodos de aparición de síntomas y en las columnas los rezagos del tiempo de notificación. A partir de esta se crea una matriz de “variación” (proporciones) entre los períodos de notificación.

En el fragmento que se presenta a continuación se puede ver como para el período de reporte de síntomas 15 (fila 15) en el período 5, se presentaron 3 veces el número de casos del período 4. Ahora bien, para el reporte de síntomas 16 (fila 16) en el período 4, se notificaron 1,33 veces más casos que en el período 3.

Figura 5 Ejemplo de tabla de contingencias

	V1	V2	V3	V4	V5
15	NaN	NaN	NaN	NaN	3.000000
16	NaN	NaN	NaN	1.333333	1.250000
17	NaN	NaN	1.000000	2.000000	2.000000
18	NaN	NaN	5.000000	1.200000	1.000000
19	NaN	NaN	NaN	2.000000	1.500000
20	NaN	NaN	1.500000	1.000000	1.333333
21	NaN	1	1.000000	2.000000	3.500000

Una vez construida la anterior matriz se toma el promedio de la variación para cada rezago de notificación y se realiza una multiplicación acumulada por período, es decir, si la proporción para el período 1 es 1,6 y para el período 2 fue 2,0, la proporción acumulada en el período 2 corresponde a $1,6 \times 2 = 3,2$. Estas proporciones acumuladas se aplican al número de casos iniciales para encontrar el número de casos que deberían haberse notificado para cada uno de los períodos. Así, puede obtenerse un estimado de los casos que podrían esperarse para cada fecha de notificación (o diagnóstico) y compararse con los efectivamente observados.

Anexo 2. Modelos de series de tiempo

2.1 Modelo Bayesiano

Como se expresó en la sección 3.2.1, la ecuación de renovación (*renewal*) se define como:

$$E[I_N(t)] = \sum_{\tau=1}^{\infty} \beta(t, \tau) I_N(t - \tau) \quad (1)$$

Donde el factor de ajuste es $\beta(t, \tau) = \phi(t)\omega(\tau)$.

En este caso el factor $\phi(t)$ está relacionado con las características sociales vinculadas a la tasa de reproducción, y $\omega(\tau)$, define un factor dependiente del rezago, asociado directamente al comportamiento biológico de la enfermedad.

A partir del análisis sugerido por Wallinga y Teunis (2004)(25), los autores derivan que el número de personas que pueden ser infectadas por un individuo es:

$$\mathcal{R}(t) = \sum_{\tau=1}^{\infty} \beta(t, \tau) = \phi(t) \quad (2)$$

Al sustituir, se obtiene que:

$$\mathcal{R}(t) = \frac{E[I_N(t)]}{\sum_{\tau=1}^{\infty} \omega(\tau) I_N(t - \tau)} \quad (3)$$

de donde

$$E[I(t)] = \mathcal{R}(t) \sum_{\tau=1}^{\infty} \omega(t) I_N(t - \tau) \quad (4)$$

Finalmente, siguiendo a Fraser (2007)(21), se obtiene que

$$\mathcal{R}_c(t) = \sum_{\tau=1}^{\infty} \omega(t) \mathcal{R}(t - \tau) \quad (5)$$

Ahora bien, suponiendo conocimiento previo del intervalo serial de la enfermedad, el problema de calibración se reduce a la estimación de $E[I_N(t)]$. Para tal fin, la estimación Bayesiana asume que el número de casos infecciosos nuevos en el período $t+1$ sigue una distribución de Poisson de media indicada por la ecuación (6). Adicional a esto, el modelo asume que, condicional a estar en el período t , y por una ventana de tiempo w , la tasa de reproducción no va a cambiar. Bajo estas premisas, se tiene que:

$$\sum_{\tau=1}^w I(t) \sim \text{Poisson} \left(\mathcal{R}(t) \sum_{\tau=0}^{w-1} \sum_{\tau=1}^{\infty} \omega(t) I_N(t-\tau) \right) \quad (6)$$

Ahora, se asume una distribución a priori de $\mathcal{R} \sim \text{Gamma}(\gamma_1, \gamma_2)$ donde γ_1 es el parámetro de forma y γ_2 es el parámetro de razón, se obtiene un modelo Gamma-Poisson en donde:

$$\mathcal{R}(t) \sim \text{Gamma} \left(\gamma_1 + \sum_{\tau=1}^w I_N(t), \gamma_2 + \sum_{\tau=0}^{w-1} \sum_{\tau=1}^{\infty} \omega(t) I_N(t-\tau) \right)$$

Según el análisis de Fraser (2007) (21), las principales ventajas de este modelo son el desarrollo accesible de los cálculos y la relativa simplicidad para encontrar los intervalos seriales. Sin embargo, el cálculo de estos genera problemas conceptuales para la estimación del $\mathcal{R}(t)$. Tal como lo plantea el autor, el cálculo del intervalo serial se ve afectado por las normas sociales. Debido a esto, presumir intervalos seriales fijos en escenarios donde las interacciones entre personas son afectadas por políticas públicas, puede afectar la capacidad del $\mathcal{R}(t)$ de capturar estos cambios. Adicional al problema conceptual, la necesidad de determinar una ventana de estimación w , genera una suavización forzada en la cual, diferentes períodos se promedian. Por esta razón, $\mathcal{R}(t)$ no captura exclusivamente la capacidad de infección de un día determinado, sino que incluye, por construcción una combinación de estos valores en los períodos anexos.

Calibración del modelo (suavizamiento de la serie):

Se obtiene la ecuación (A1):

$$E [I(t+\tau)] \beta(t) N(0) \prod_{h=1}^{\tau} [\beta(h) + e^{-\tau}] \quad (A1)$$

Sea $C(t)$ el número de nuevos casos observados en el período t . Entonces, los $t+2$ parámetros ($\mathbb{P}=\{N(0),\tau,\beta(1),\dots,\beta(t)\}$) del modelo se ajustan a partir del siguiente programa de optimización Ω :

$$\Omega = \min_{\mathbb{P} \in \mathbb{P}} \left[\frac{1}{t} \sum_{h=1}^t \left(\mathbb{E}[I(h)] - C(h) \right)^2 + \frac{\lambda}{t-1} \sum_{h=2}^t \left(\beta(h) - \beta(h-1) \right)^2 \right] \quad (A2)$$

Donde los soportes de cada parámetro respetan las definiciones de la descripción del modelo y λ es un meta parámetro que suaviza la serie $\beta(t)$ para evitar sobreajuste. Nótese que, dada la forma multiplicativa de la ecuación A1, si se asume que el número de casos asintomáticos es una fracción fija del número total de infectados, este modelo permite la estimación de $\beta(t)$ sin necesidad de ajustar por los asintomáticos, recordando que para pasar de $\beta(t)$ a $\mathcal{R}(t)$ se usa la fórmula $\mathcal{R}(t)=\beta(t)/\tau$.

2.2 Modelo estado-espacio

Buscando solucionar las limitaciones planteadas en el modelo anterior, se considera el modelo de Estado-Espacio que redefine las ecuaciones (1)-(7) para generar un cálculo en el cual $\omega(\tau)$ capture únicamente los componentes biológicos de la enfermedad, y luego a partir de la estimación estadística Bayesiana, se conduzca a una estimación basada en series de tiempo con el propósito de evitar los problemas generados por la ventana de estimación.

Sea f_{Inc} una función de probabilidad, donde $f_{\text{Inc}}(t)$ es la probabilidad de que un individuo se tome t días desde que fue expuesto a la enfermedad, hasta el inicio de una etapa infecciosa. De la misma manera, sea F_{Inf} la función de probabilidad acumulada, donde $F_{\text{Inf}}(t)$ es la probabilidad que una persona sea infecciosa por, a lo más, t períodos. A partir de estas definiciones, se considera:

$$\omega(t) = \sum_{\tau=1}^t f_{\text{Inc}}(\tau) (1 - F_{\text{Inf}}(t - \tau))$$

Esta redefinición presenta un valor ω que se relaciona directamente a la probabilidad de que una persona que fue infectada hace t períodos, pueda seguir siendo infecciosa el día de hoy. Mas aún, $\Omega = \sum_{\tau=1}^{\infty} \omega(\tau)$ es el tiempo esperado en que una persona dura siendo infecciosa⁴.

Al realizar estos ajustes, se modifica la ecuación (1), considerando los tiempos de incubación, se define entonces como

⁴Esta afirmación es una conclusión directa de la convolución sobre la función de supervivencia.

$$\mathbb{E} \left[\sum_{\tau=1}^{\infty} \beta I_{nc}(\tau) I_N(t-\tau) \right] = \sum_{\tau=1}^{\infty} \beta(t, \tau) I_N(t-\tau) \quad (1')$$

en donde existe una intuición directa entre el potencial de infectados, reflejado por la parte izquierda de la ecuación, y el número potencial de personas con la capacidad de infectar, representado por la parte derecha de la ecuación.

Al combinar estas ecuaciones, se obtiene que:

$$\mathcal{R}(t) = \frac{\mathbb{E} \left[\sum_{\tau=1}^{\infty} \beta I_{nc}(\tau) I_N(t+\tau) \right]}{\sum_{\tau=1}^{\infty} \omega(\tau) I_N(t-\tau)} \Omega \quad (3')$$

$$\mathbb{E} \left[\sum_{\tau=1}^{\infty} \beta I_{nc}(\tau) I_N(t-\tau) \right] = \frac{\mathcal{R}(t)}{\Omega} \sum_{\tau=1}^{\infty} \omega(\tau) I_N(t-\tau) \quad (4')$$

Nótese que este cambio presenta características análogas al modelo Bayesiano. En este caso, y acorde al concepto, $\mathcal{R}(t)$ está asociado a la tasa de contagio en un día dado,

$$\frac{\mathbb{E} \left[\sum_{\tau=1}^{\infty} \beta I_{nc}(\tau) I_N(t+\tau) \right]}{\sum_{\tau=1}^{\infty} \omega(\tau) I_N(t-\tau)}$$

multiplicada por el número de días que la persona puede infectar Ω .

Ahora bien, para calibrar este modelo se asume, al igual que en el caso anterior, que el proceso de contagio es Poisson de media indicada por la ecuación (6).

$$\sum_{\tau=1}^{\infty} \beta I_{nc}(\tau) I_N(t+\tau) \sim \text{Poisson} \left(\frac{\mathcal{R}(t)}{\Omega} \sum_{\tau=1}^{\infty} \omega(\tau) I_N(t-\tau) \right) \quad (6')$$

Sin embargo, en vez de calibrar este modelo en un marco Bayesiano, se considera el siguiente modelo de estado espacio no Gaussiano (donde la función de enlace es, naturalmente Poisson). Para el caso del modelo Poisson, la función de enlace (*linking function*) es tal que:

$$E \left[\sum_{\tau=1}^{\infty} I_{inc}(\tau) I_N(t+\tau) \right] = e^{z_t \alpha_t}$$

Adicionalmente se definen

$$z_t = \left[I_n \left(\sum_{\tau=1}^{\infty} \omega(\tau) I_N(t-\tau) \right), 1, 0 \right]$$

$$\alpha_t = \begin{bmatrix} \alpha_{1,t} \\ \alpha_{2,t} \\ \alpha_{3,t} \end{bmatrix}$$

$$T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$R = \begin{bmatrix} 0 & 0 \\ \sigma_{\alpha_2} & 0 \\ 0 & \sigma_{\alpha_3} \end{bmatrix}$$

donde las condiciones iniciales están dadas por⁵:

$$\alpha_1 = \begin{bmatrix} 1 \\ \mu_0 \\ \mu_0 \end{bmatrix}, P_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 10^9 & 0 \\ 0 & 0 & 10^9 \end{bmatrix}$$

De esta forma, el modelo se puede definir como:

Ecuación de observación:

$$E \left[\sum_{\tau=1}^{\infty} \omega(\tau) I_N(t+\tau) \right] = e^{\alpha_2 t} \left(\sum_{\tau=1}^{\infty} \omega(\tau) I(t-\tau) \right)$$

Ecuación de transición:

$$\begin{aligned} \alpha_{2,t+1} &= 1 \\ \alpha_{2,t+1} &= \alpha_{2,t} + \alpha_{3,t+1} + \eta_{1,t} \\ \alpha_{3,t+1} &= \alpha_{3,t} + \eta_{2,t} \end{aligned}$$

Lo cual describe un modelo con pendiente estocástica (Harvey, 1990)(26), en donde los parámetros no especificados: σ_{α_2} y σ_{α_3}) se calculan vía máxima verosimilitud.

Ahora, para adaptar esta técnica al modelo epidemiológico, se tiene que,

$$\mathcal{R}(t) = e^{\alpha_2 t} \Omega$$

donde $\mathcal{R}_c(t)$ se determina a través de la ecuación (5) ajustada por Ω :

$$\mathcal{R}_c(t) = \sum_{\tau=1}^{\infty} \omega(\tau) \frac{\mathcal{R}(t-\tau)}{\Omega} \quad (5')$$

⁵Nótese que la varianza del prior de $\alpha_{1,1}$ es 0 porque este parámetro no se desea que cambie en la regresión. De hecho, se desea que sea 1 desde el inicio y se mantenga. La varianza grande de los otros valores se da para resaltar que no se asume información específica sobre el prior, dándole oportunidad al modelo de ajustarse rápidamente a los datos. Finalmente, μ_0 describe \mathcal{R}_0 de la literatura epidemiológica, pero se le da la libertad de variar para ajustarse al contexto específico del área de análisis.