

# Informe Proyecciones COVID-19 UdeC

Guillermo Cabrera-Vives (guillecabrera@inf.udec.cl)

Cristóbal Donoso-Oliva

Montserrat Martínez

Roberto Molina

Alexis Sánchez

27 de Abril, 2020

## Resumen

El siguiente reporte presenta los resultados de ajustar un modelo epidemiológico SEIR modificado a datos COVID-19 de las regiones de Ñuble, Biobío y a Chile. Los modelos SEIR son modelos de compartimentos donde la población se divide en Susceptibles-Expuestos-Infectados-Removidos. Nuestro modelo modificado considera que el número de infectados reportados no es real, debido a la capacidad de muestreo y a los infectados asintomáticos. Además, considera que la tasa de contagio de la población varía en el tiempo debido, por ejemplo, a distintas medidas de mitigación. **Es MUY IMPORTANTE notar que los resultados de nuestro modelo no se pueden interpretar como finales, son solamente aproximaciones cualitativas de la situación ACTUAL. Esto debido a que los datos no son completamente confiables y el comportamiento de la población cambia día a día.** Este reporte ha sido escrito bajo la urgencia de la pandemia COVID-19 con el objetivo de informar a tomadores de decisión del país. Nuestro código y datos se encuentran disponibles en <https://github.com/guille-c/COVID-19>.

## 1. Modelo SEIR considerando muestreo y estrategias de contención

Hemos modelado la epidemia del COVID-19 utilizando un modelo de compartimentos SEIR modificado. El modelo SEIR [KM27, Het00] es un modelo matemático que modela a la población en distintos compartimentos dependiendo de su estado de infección. En este modelo se divide a la población en cuatro grupos:

**Susceptibles ( $S(t)$ ):** Número de personas en el tiempo  $t$  que no han sido infectadas por el virus, y que corren el riesgo de serlo.

**Expuestos ( $E(t)$ ):** Número de personas en el tiempo  $t$  que se han visto expuestas al virus y que están en etapa de incubación. Esta población no está infectando.

**Infectados ( $I(t)$ ):** Número de personas contagiadas en el tiempo  $t$  que son capaces de contagiar al resto.

**Removidos ( $R(t)$ ):** Número de personas recuperadas o fallecidas en el tiempo  $t$  que ya no contagian y que no volverán a contagiar.

Además, consideramos el número acumulado de contagiados a la fecha como  $C(t)$ . El número de personas en cada estado es modelado mediante las siguientes ecuaciones diferenciales:

$$\frac{dS}{dt} = -\beta SI, \quad (1)$$

$$\frac{dE}{dt} = \beta SI - \sigma E, \quad (2)$$

$$\frac{dI}{dt} = \sigma E - \gamma I, \quad (3)$$

$$\frac{dR}{dt} = \gamma I, \quad (4)$$

donde  $\beta$  es la tasa de contagio,  $\sigma^{-1}$  es el período de incubación, y  $\gamma^{-1}$  es el período de contagio. Además, si consideramos que  $C(t) = I(t) + R(t)$ , entonces tenemos que

$$\frac{dC}{dt} = \sigma E. \quad (5)$$

Tanto el período de incubación como el período de contagio son variables propias del virus (y de las condiciones inmunológicas de la población). Entonces, la única variable que podemos modificar de este modelo es la tasa de contagio  $\beta$ . En nuestro modelo, consideramos que la tasa de contagio  $\beta(t)$  puede variar en el tiempo. En particular, la ajustamos a distintos escenarios de medidas de contención (por ejemplo, cuarentenas o cercos sanitarios).

Dentro de los datos entregados por el gobierno de Chile, tenemos el número de contagiados acumulados, el número de recuperados y el número de fallecidos. El número de recuperados no se calcula directamente y el criterio para hacerlo ha cambiado en el tiempo, por lo que no es un dato confiable. El número de fallecidos, es un dato más confiable, pero en nuestro modelo actual está siendo agregado con los recuperados, por lo que no podemos utilizarlo directamente. Planeamos considerarlo explícitamente en nuestro modelo en el futuro. El número de contagiados entregados por el gobierno solamente toma en consideración la población a la cuál se le ha realizado un examen PCR, por lo que el modelo no puede ser ajustado a este valor. Si consideramos que el número de contagiados por día reportados  $C'(t)$  es un porcentaje  $\alpha(t)$  de los contagiados reales diarios  $C(t)$ , entonces

$$\frac{dC'(t)}{dt} = \alpha(t) \frac{dC(t)}{dt}, \quad (6)$$

$$= \sigma \alpha(t) E(t). \quad (7)$$

Para modelar  $\alpha(t)$ , asumimos que si el número de contagiados diarios es bajo, entonces, hay suficientes exámenes para toda la población i.e. el número de contagiados diarios reportados es igual al número de contagiados diarios reales ( $\frac{dC'(t)}{dt} = \frac{dC}{dt}$ ). Por el otro lado, asumimos que si el número de contagiados diarios es muy alto, solamente se podrá medir como máximo a un porcentaje  $a$  de la población. De esa forma, modelamos  $\alpha(t)$  como una sigmoide invertida

$$\alpha(t) = 1 + \frac{a - 1}{1 + e^{-k(\frac{dC}{dt} - a_C)}}, \quad (8)$$

donde  $k$  describe la tasa de decrecimiento, y  $a_C$  representa el punto medio. Es importante notar que este modelo agrega tres nuevos parámetros:  $a$ ,  $k$ , y  $a_C$ .

### 1.1. Ajuste de parámetros

El virus SARS-CoV-2 causante de la enfermedad por coronavirus COVID-19 aún está bajo estudio y no se conoce con exactitud los períodos de incubación y contagio<sup>1</sup>. De la misma forma, tampoco conocemos la tasa de contagio  $\beta$  ni el efecto exacto de las distintas medidas de contención. Debido a esto decidimos ajustar cada uno de los parámetros del modelo ( $\beta(t)$ ,  $\sigma$ ,  $\gamma$ ,  $a$ ,  $k$  y  $a_C$ ). Para esto, consideramos que  $\beta(t)$  es constante para cada escenario de medidas de contención.

El ajuste fue realizado utilizando una búsqueda exhaustiva sobre una grilla de parámetros adaptable. Para ajustar el valor de los parámetros utilizamos un ajuste de mínimos cuadrados sobre  $C'(t)$ . Los rangos de los períodos de incubación y contagio fueron definidos como  $\sigma^{-1} \in [1, 7]$  y  $\gamma^{-1} \in [6, 18]$  basado en informes de la Organización Mundial de la Salud [OO+20]<sup>2</sup>.

## 2. Resultados

Ajustamos el modelo a las regiones de Ñuble, Biobío y al territorio nacional. Los valores de los respectivos parámetros ajustados para cada experimento se describen en el Cuadro 1. De acuerdo a nuestros resultados, el período de incubación es de 3-4 días y el período de infección es de 8-12 días.

### 2.1. Ñuble

Para la región de Ñuble consideramos tres escenarios: desde el 21 de Marzo del 2020 (sobre 50 infectados) al 2 de Abril (inicio de cuarentena en Chillán), desde esa fecha al 23 de Abril (fin de cuarentena), y desde esa fecha hasta el fin de este estudio el 27 de Abril. Podemos ver en el Cuadro 1 que hay una disminución importante en  $\beta(t)$  entre el primer y segundo periodo teniendo el segundo

<sup>1</sup>ver, por ejemplo, [https://github.com/midas-network/COVID-19/tree/master/parameter\\_estimates/2019\\_novel\\_coronavirus](https://github.com/midas-network/COVID-19/tree/master/parameter_estimates/2019_novel_coronavirus)

<sup>2</sup><https://www.who.int/docs/default-source/coronaviruse/who-china-joint-mission-on-covid-19-final-report.pdf>

	Ñuble	Biobío	Chile
Población <sup>3</sup>	511.551	1.663.696	19.458.310
$\beta$	21/3-02/4: $8 \times 10^{-7}$ 02/4-23/4: $1 \times 10^{-7}$ 23/4-27/4: $3 \times 10^{-7}$	24/3-06/4: $1,51 \times 10^{-7}$ 06/4-27/4: $8,42 \times 10^{-8}$	02/3-15/3: $4,63 \times 10^{-8}$ 15/3-24/3: $3,70 \times 10^{-8}$ 24/3-27/4: $7,70 \times 10^{-9}$
$1/\sigma$	3	3	4
$1/\gamma$	8	11	12
$a$	0.5	0.153	0.2
$k$	0.01	0.151	0.001
$a_C$	60	50	20
<i>RMSE</i>	12.81	10.19	169.35

Cuadro 1: Parámetros del modelo SEIR modificado ajustados por mínimos cuadrados a los contagiados medidos (reportados por el Ministerio de Salud).

periodo una tasa de contagio de aproximadamente un 12.5 % del primero. Luego de eliminar la cuarentena, la tasa de contagio es aproximadamente un 37.5 % de la tasa previa a la cuarentena lo cual podría ser debido a que la población sigue intentando evitar el contacto. Otra razón podría ser que nuestro modelo esta sobre estimando el porcentaje de infectados que se están midiendo. Evaluar la verdadera razón de esto es demasiado complejo, y requeriría de muestreo aleatorio de la población, por lo que está fuera del alcance de este estudio. De acuerdo a nuestro modelo, el porcentaje mínimo de infectados medidos sobre el total es de  $a = 50 \%$ .

La Figura 1 muestra el ajuste de nuestro modelo a los datos, así como también la proyección hasta Julio del 2021. Si las condiciones siguen siendo como las post-relajación de cuarentena (por ejemplo, la población sigue comportándose igual), entonces el peak se produciría en Febrero del 2021, llegando a aproximadamente un 2 % de la población de Ñuble infectada.

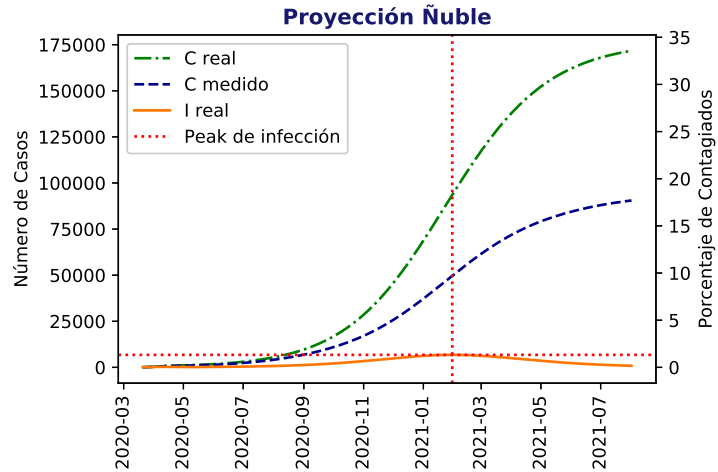
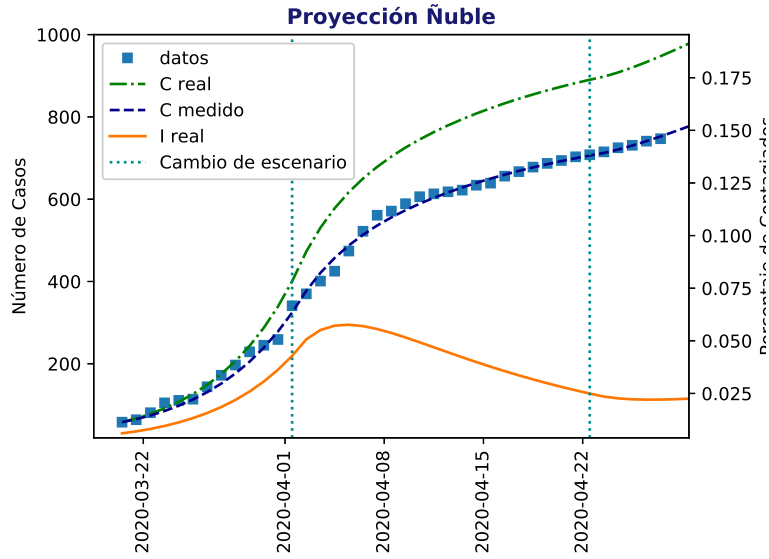


Figura 1: Proyecciones para la región de Ñuble. Esta figura muestra los contagiados acumulados reales  $C(t)$  (de acuerdo a nuestro modelo), los contagiados acumulados medidos  $C'(t)$  (ajustados a los datos entregados por el Ministerio de Salud) y los infectados simultáneos  $I(t)$ . El gráfico superior muestra el modelo ajustado a los datos entre Marzo del 2020 y Abril del 2020. El gráfico inferior muestra las proyecciones hasta Julio del 2021.

## 2.2. Biobío

Para la región del Biobío es más difícil considerar distintos escenarios basados en medidas puntuales que en Ñuble, ya que se han tomado distintas medidas en distintas comunas. Debido a esto, definimos dos escenarios de manera arbitraria mirando la curva de contagiados acumulados e intentando reducir el error del modelo: entre el 24 de Marzo (sobre 50 contagiados) y el 6 de Abril, y entre esta última fecha y el 27 de Abril. Podemos ver que en el segundo escenario se reduce a aproximadamente un 55.8 % la tasa de contagio. Al igual que antes, evaluar la verdadera razón de esta disminución está fuera del alcance de este estudio. De acuerdo a nuestro modelo, el porcentaje mínimo de infectados medidos sobre el total es de  $a = 15,3\%$ .

La Figura 2 muestra el ajuste de nuestro modelo a los datos, así como también la proyección hasta Enero del 2021. Si las condiciones siguen siendo como las actuales, entonces nuestro modelo proyecta que el peak se produciría en Octubre del 2020, llegando a aproximadamente un 7 % de la población del Biobío infectada al mismo tiempo.

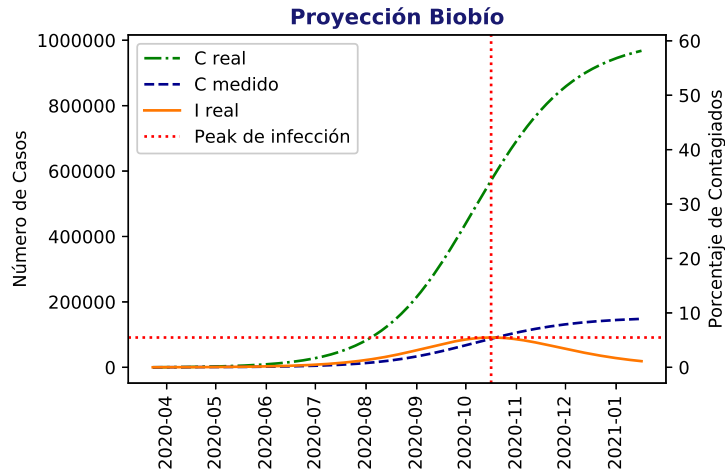
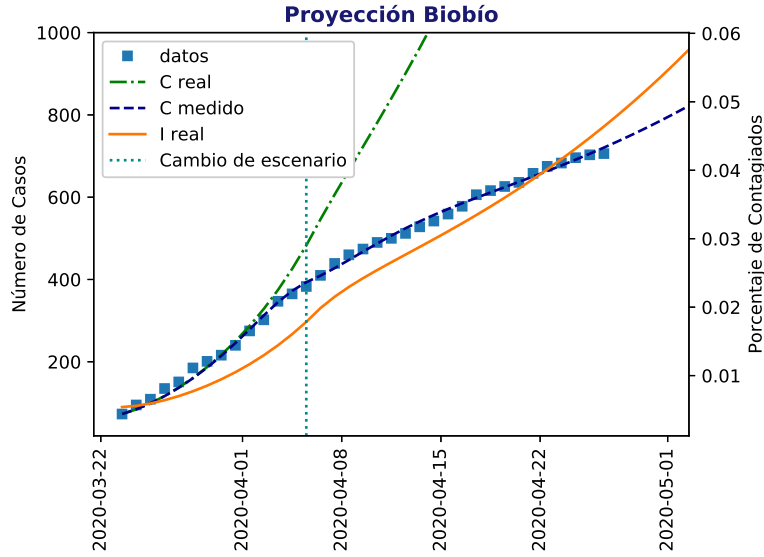


Figura 2: Proyecciones para la región de Biobío. Esta figura muestra los contagiados acumulados reales  $C(t)$  (de acuerdo a nuestro modelo), los contagiados acumulados medidos  $C'(t)$  (ajustados a los datos entregados por el Ministerio de Salud) y los infectados simultáneos  $I(t)$ . El gráfico superior muestra el modelo ajustado a los datos entre Marzo del 2020 y Abril del 2020. El gráfico inferior muestra las proyecciones hasta Enero del 2021.

### 2.3. Chile

Para Chile también es difícil considerar distintos escenarios basados en medidas puntuales, ya que se han tomado distintas medidas en distintas comunas y regiones. Esta vez, definimos tres escenarios de manera arbitraria mirando la curva de contagiados acumulados e intentando reducir el error del modelo: entre el 2 de Marzo (primer contagiado) y el 15 de Marzo, entre esta última fecha y el 24 de Marzo, y entre esa fecha y el 27 de Abril. Podemos ver que en el segundo escenario se reduce a aproximadamente un 79.9% la tasa de contagio y en el tercer escenario se reduce a un 16.6% de la tasa de contagio del primer escenario. Al igual que antes, evaluar la verdadera razón de esta disminución está fuera del alcance de este estudio. De acuerdo a nuestro modelo, el porcentaje mínimo de infectados medidos sobre el total es de  $a = 20\%$ .

La Figura 3 muestra el ajuste de nuestro modelo a los datos, así como también la proyección hasta Diciembre del 2020. Si las condiciones siguen siendo como las actuales, entonces nuestro modelo proyecta que el peak se produciría en Septiembre del 2020, llegando a aproximadamente un 9% de la población nacional infectada al mismo tiempo.



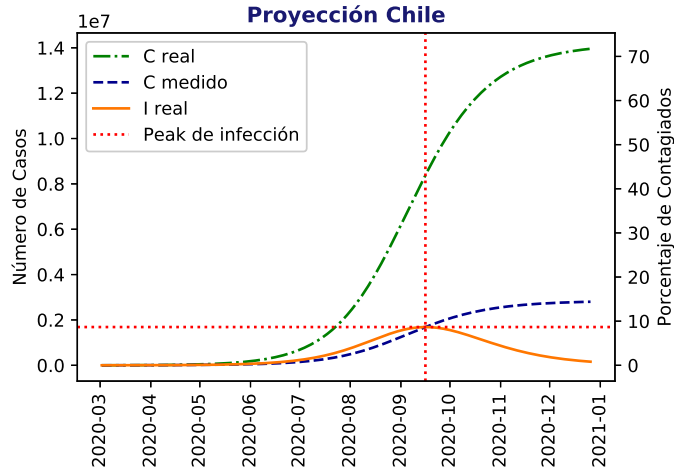
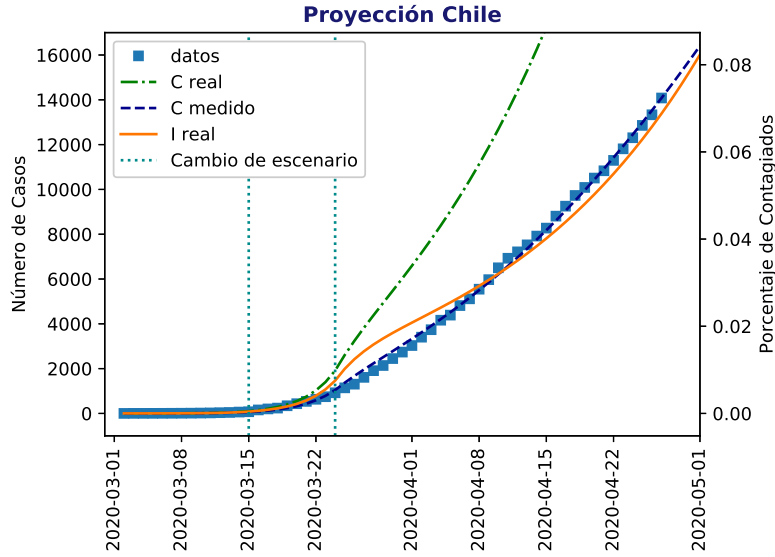


Figura 3: Proyecciones para Chile. Esta figura muestra los contagiados acumulados reales  $C(t)$  (de acuerdo a nuestro modelo), los contagiados acumulados medidos  $C'(t)$  (ajustados a los datos entregados por el Ministerio de Salud) y los infectados simultáneos  $I(t)$ . El gráfico superior muestra el modelo ajustado a los datos entre Marzo del 2020 y Abril del 2020. El gráfico inferior muestra las proyecciones hasta Julio del 2021.

## 2.4. Ejemplo de cuarentenas alternadas: Ñuble

En esta sección consideramos el efecto de hacer cuarentenas alternadas para Ñuble. Para hacer esto, consideramos dos escenarios que se van alternando: escenario pre-cuarentena ( $\beta = 8 \times 10^{-7}$ ) y escenario de cuarentena ( $\beta = 1 \times 10^{-7}$ ). Es importante notar que esta simulación, al suponer la tasa de contagio  $\beta = 8 \times 10^{-7}$  (equivalente a la tasa pre-cuarentena) para períodos sin cuarentena puede sobre-estimar el número de infectados proyectados. Esto se puede notar al comparar esta tasa de contagiados con la estimada post cuarentena en el Cuadro 1 de  $\beta = 3 \times 10^{-7}$ . Sin embargo, creemos que este experimento es importante, ya que puede ilustrar de manera cualitativa el efecto de realizar cuarentenas alternadas.

La Figura 4 muestra el caso de estas cuarentenas alternadas cada dos semanas y el caso sin cuarentena (considerando  $\beta = 8 \times 10^{-7}$ ) después del 6 de Abril. Se puede ver que efectivamente las cuarentenas progresivas son capaces de disminuir el número de infectados simultáneos. A su vez, las cuarentenas alternadas producen el efecto de extender la epidemia.

La Figura 5 muestra el caso de estas cuarentenas progresivas cada tres semanas y el mismo caso anterior sin cuarentena después del 6 de Abril. Se puede ver que en este caso se llega a un máximo de infectados simultáneos de aproximadamente 63.256 (12,37% de la población de Ñuble). Esto es mayor a los 50.085 infectados simultáneos al simular cuarentenas cada dos semanas (9.8% de la población de Ñuble).

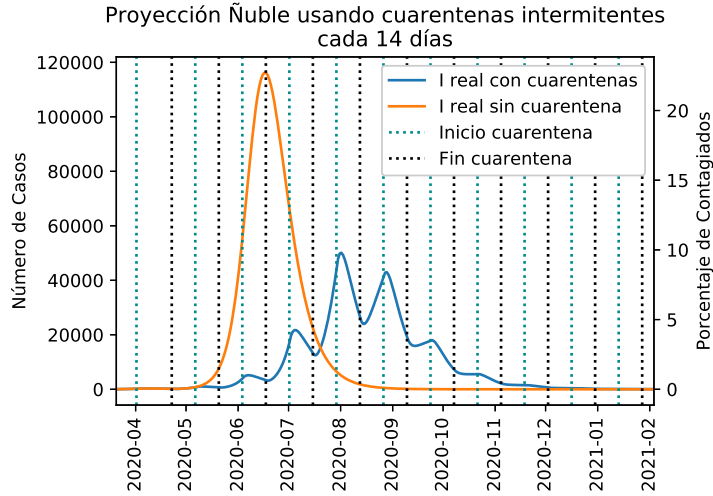


Figura 4: Proyecciones para Ñuble con y sin cuarentenas progresivas de dos semanas.

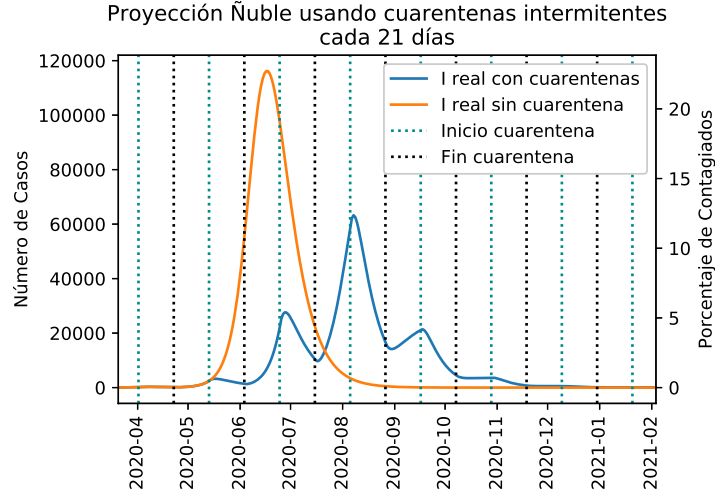


Figura 5: Proyecciones para Ñuble con y sin cuarentenas progresivas de tres semanas.

### 3. Observaciones finales

Cualquier tipo de modelo epidemiológico es muy difícil de ajustar, y sus proyecciones no se pueden tomar como definitivas, principalmente debido a que los datos son poco confiables, y el comportamiento de la población cambia en el tiempo. En ese sentido, los resultados de este y otros informes deben ser entendidos de manera cualitativa y no cuantitativa. En otras palabras, el número de infectados o fechas exactas de peaks en el largo plazo no deben tomarse como reales.

Nuestro estudio nota que a medida que se toman nuevas medidas de contención, es posible disminuir el número de contagiados totales simultáneos. A su vez, al hacer esto, la duración de la epidemia aumenta.

Nuestro modelo no considera el detalle de cada medida tomada en cada región o comuna (excepto en Ñuble). Tampoco considera el transporte entre distintas zonas. Hemos desarrollado un modelo que si lo hace, pero no hemos sido capaces de ajustarlo con una baja tasa de error (menor al 10% de los contagiados acumulados) a los datos reales por comuna.

Posterior a las fechas consideradas en este modelo se comenzaron a medir de distinta manera los casos contagiados diarios, lo cual plantea un nuevo problema para re-ajustar los parámetros. Actualmente estamos trabajando en esto.

Existen muchas aproximaciones al modelamiento epidemiológico. Algunos trabajos consideran hospitalizados, hospitalizados críticos, transporte entre zonas, segmentación por edades, segmentación por nivel socioeconómico, y micro-simulaciones [FLNG<sup>+</sup>20], entre otros. Estamos considerando varias de estas alternativas para modelar en más detalle el COVID-19 en Chile. Es importante

notar que al ir complejizando el modelo, se van agregando nuevos parámetros que ajustar o intentar fijar desde evidencia científica. Por ejemplo, el modelo presentado en [CCdW<sup>+</sup>20] considera aproximadamente 15 parámetros (algunos ajustados, otros fijados desde valores obtenidos de la literatura). Hay que tener un especial cuidado de no sobre-ajustar el modelo a los datos a medida que éstos se van complejizando.

## Agradecimientos

El equipo de proyecciones agradece a los académicos de la Universidad de Concepción que participan activamente en modelar el COVID-19 desde distintas aristas: Andrea Rodríguez, Sebastián Astroza, Juan Carrasco y Diego Seco. Agradecemos también a la Seremi de Ciencia de la macrozona centro sur, Paulina Assmann por el apoyo brindado al equipo. Agradecemos a Mauricio Lima y a Pedro Gajardo por compartir los códigos de sus modelos que han dado nuevas luces sobre las distintas direcciones que podríamos seguir.

## Referencias

- [CCdW<sup>+</sup>20] Alonso Cancino, Carla Castillo, Taco de Wolff, Pedro Gajardo, Rodrigo Lecaros, Claudio Munoz, Jaime Ortega, Héctor Ramirez, Nicolás Valenzuela, and Santa Maria. Report# 4: Estimation of maximal icu beds demand for covid-19 outbreak in some chilean regions and the effects of different mitigation strategies. 2020.
- [FLNG<sup>+</sup>20] Neil Ferguson, Daniel Laydon, Gemma Nedjati Gilani, Natsuko Imai, Kylie Ainslie, Marc Baguelin, Sangeeta Bhatia, Adhiratha Boonyasiri, ZULMA Cucunuba Perez, Gina Cuomo-Dannenburg, et al. Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid19 mortality and healthcare demand. 2020.
- [Het00] Herbert W Hethcote. The mathematics of infectious diseases. *SIAM review*, 42(4):599–653, 2000.
- [KM27] William Ogilvy Kermack and Anderson G McKendrick. A contribution to the mathematical theory of epidemics. *Proceedings of the royal society of london. Series A, Containing papers of a mathematical and physical character*, 115(772):700–721, 1927.
- [OO<sup>+</sup>20] World Health Organization, World Health Organization, et al. Report of the who-china joint mission on coronavirus disease 2019 (covid-19), 2020.