

Optimización de un modelo de transmisión de COVID-19 para su aplicación a múltiples periodos de intervención

Trabajo Fin de Máster

Máster Universitario en Bioinformática y Bioestadística

Área 1, Subárea 13: Modelado molecular y biología computacional en biomedicina

Autor	Javier Blecua Pérez
Consultor	Juan Fernández Recio
Profesor responsable	David Merino Arranz
Fecha de entrega	23 de junio de 2021



Esta obra está sujeta a una licencia de Reconocimiento-NoComercial-CompartirIgual
[3.0 España de Creative Commons](https://creativecommons.org/licenses/by-nc-sa/3.0/es/)

Ficha del trabajo final

Título del trabajo:	<i>Optimización de un modelo de transmisión de COVID-19 para su aplicación a múltiples periodos de intervención</i>
Nombre del autor:	<i>Javier Blecua Pérez</i>
Nombre del consultor:	<i>Juan Fernández Recio</i>
Nombre del PRA:	<i>David Merino Arranz</i>
Fecha de entrega:	<i>06/2021</i>
Titulación:	<i>Máster Universitario en Bioinformática y Bioestadística</i>
Área del Trabajo Final:	<i>Modelado molecular y biología computacional de biomedicina</i>
Idioma del trabajo:	<i>Español</i>
Número de créditos:	<i>15</i>
Palabras clave	<i>COVID-19, Intervenciones no farmacológicas, Análisis bayesiano</i>

Resumen del Trabajo:

El COVID-19 (*coronavirus disease 2019*) es la enfermedad causada por un nuevo tipo de coronavirus llamado SARS-CoV-2.

Con este Trabajo Fin de Máster, se ha optimizado y extendido un modelo existente de transmisión del COVID-19 para el análisis de múltiples periodos de diferente velocidad de transmisión, permitiendo incluir, como parte del análisis, la información sobre un número arbitrario de medidas o intervenciones no farmacológicas.

La metodología consiste en la aplicación de un modelo mecanístico que describe la evolución de la enfermedad a partir del efecto de las distintas medidas de intervención y que, aplicando métodos de análisis bayesiano, permite obtener estimaciones sobre este efecto para cada medida.

Los buenos resultados obtenidos para un total de 32 países europeos, con datos que siguen patrones muy distintos, permiten concluir, por un lado, que el modelo utilizado es un buen método para describir la evolución de la enfermedad.

Por otro lado, a partir de los resultados, se han extraído conclusiones sobre el efecto de las distintas medidas que son consistentes con la información conocida. Por ejemplo, los resultados para la primera ola de la pandemia indican que las medidas más efectivas fueron las relacionadas con el confinamiento, las medidas aplicadas conjuntamente en marzo de 2020 y el uso obligatorio de mascarillas.

Sin embargo, algunos resultados para fechas más recientes han puesto de manifiesto algunas limitaciones del modelo. Entre otras, que en la versión actual no se considera el efecto de la vacunación en la evolución de la enfermedad.

Abstract:

COVID-19 (coronavirus disease 2019) is the disease caused by a new type of coronavirus called *SARS-CoV-2*.

With this Master thesis, an existing *COVID-19* transmission model has been optimized and extended for the analysis of multiple periods of different transmission rate, making possible the inclusion, as part of the analysis, of information on an arbitrary number of non-pharmacological measures.

The methodology consists in the application of a mechanistic model that describes the evolution of the disease based on the effect of the different intervention measures and that makes the estimation of this effect possible by applying Bayesian analysis methods.

The positive results obtained for a total of 32 European countries, with data following very different patterns, lead to the conclusion, on the one hand, that the model used is an appropriate method for describing the evolution of the disease.

On the other hand, the results lead to conclusions regarding the effect of the different measures that are consistent with available information. For example, the results for the first wave of the pandemic indicate that the most effective measures were those related to the stay-at-home orders and recommendations (lockdown), the measures applied jointly in March 2020 and the mandatory use of masks.

However, some results for more recent data have highlighted some limitations of the model. Among others, the current version does not take into consideration the effect of vaccination on the evolution of the disease.

Ficha del trabajo final	2
Índice.....	4
Lista de figuras.....	7
Lista de tablas.....	8
1 Resumen	9
2 Introducción	11
2.1 Contexto y justificación del Trabajo	11
2.2 Objetivos del Trabajo	12
2.3 Enfoque y método seguido	13
2.4 Planificación del Trabajo	14
2.5 Breve resumen de los productos obtenidos	16
2.6 Breve descripción de los otros capítulos	16
3 Estado del arte	17
3.1 Número de reproducción (R_t)	17
3.2 Estudios previos y métodos existentes	18
3.3 Aportación de este Trabajo Fin de Máster	22
4 Metodología	23
4.1 Número de reproducción (R_t)	24
4.2 Planteamiento	25
4.3 Fundamento estadístico - Modelo mecanístico	26
4.4 Metodología - Análisis bayesiano	28
4.5 Datos necesarios para el análisis	31
4.6 Funcionalidad	34
4.6.1 Herramienta informática - Componentes	35
4.6.2 Modelos y Predicciones - Bondad del ajuste y efecto de las distintas medidas	36
4.6.3 Comparación del efecto de las distintas medidas en diferentes países	39

5	Resultados	41
5.1	Combinaciones de datos utilizados para el análisis.....	42
5.2	Agrupación y reducción de las medidas de intervención.....	43
5.2.1	Motivación	43
5.2.2	Criterios	45
5.3	Evaluación del modelo - Bondad del ajuste	48
5.3.1	Criterios para la evaluación de la bondad del ajuste.....	48
5.3.2	Bondad del ajuste de los distintos modelos	49
5.3.3	Evaluación de la bondad del ajuste - Conclusiones	50
5.4	Estimaciones para el número de reproducción básico (R_0)	51
5.5	Efecto de las distintas medidas de intervención	53
5.5.1	Efecto de las medidas de intervención - Definición	53
5.5.2	Datos utilizados para el análisis del efecto de las medidas de intervención.....	53
5.5.3	Resultados para el periodo del 22.01.2020 al 18.04.2021	55
5.5.4	Medidas más efectivas en el conjunto de países en distintos periodos	59
5.5.5	Efecto de las medidas de intervención - Conclusiones.....	62
5.6	Predicciones en distintos escenarios.....	63
5.6.1	Modelos con datos hasta el 18.02.2021 / Predicciones hasta el 18.04.2021	64
5.6.2	Modelos con datos hasta el 18.04.2021 / Predicciones hasta el 18.06.2021	67
5.6.3	Resultados de las predicciones - Conclusiones.....	72
6	Discusión	73
7	Conclusiones	74
7.1	Grado de cumplimiento de los objetivos del trabajo	74
7.2	Seguimiento de la planificación.....	75
7.3	Conclusiones obtenidas a partir de los resultados del trabajo.....	76
7.4	Limitaciones de este trabajo	77
7.5	Líneas de futuro	79
8	Glosario	81
9	Bibliografía	84

10	Anexos	88
10.1	Anexo I - Nombres de las medidas de intervención	88
10.2	Anexo II - Análisis de los datos a nivel regional para España y Alemania	90
10.3	Anexo III - Ejemplos de los resultados de predicciones hasta el 18.04.2021	92
10.4	Anexo IV - Ejemplos de resultados de ajuste poco satisfactorios	96
10.5	Anexo V - Enlaces para el acceso a todos los resultados	97
10.6	Anexo VI - Herramienta informática - aspectos técnicos	99

Lista de figuras

- Figura 1 - Definición y planificación de las tareas del Trabajo Fin de Máster
- Figura 2 - Componentes del modelo mecanístico
- Figura 3 - Serial Interval (SI), Infection-To-Death (ITD), Infection Fatality Ratio (IFR)
- Figura 4 - Herramienta informática - Componentes
- Figura 5 - Ejemplo - Predicciones para R_t y para el número de infectados y muertos por COVID-19
- Figura 6 - Ejemplo - Estimación del efecto de las medidas de intervención en un país (1)
- Figura 7 - Ejemplo - Estimación del efecto de las medidas de intervención en un país (2)
- Figura 8 - Ejemplo - Efectividad de las medidas de confinamiento en distintos países
- Figura 9 - Comparación del ajuste con distintos grupos de medidas de intervención
- Figura 10 - Efecto de las medidas de intervención - estimaciones con todas las medidas
- Figura 11 - Efecto de las medidas de intervención - estimaciones con las medidas optimizadas
- Figura 12 - Clasificación de modelos en función de la bondad del ajuste
- Figura 13 - Distribución de los modelos para 32 países europeos en función de la bondad del ajuste
- Figura 14 - Efecto de las medidas de intervención en España
- Figura 15 - Efecto de las medidas de intervención en la República Checa
- Figura 16 - Efecto de la limitación de reuniones y contactos sociales en distintos países
- Figura 17 - Efecto del periodo navideño en distintos países
- Figura 18 - Francia - Predicciones para el número de reproducción (R_t) hasta el 18.04.2021
- Figura 19 - Francia - Predicciones y datos reales hasta el 18.04.2021
- Figura 20 - Predicciones hasta el 18.06.2021 (escenario 1) - España, Grecia
- Figura 21 - Predicciones hasta el 18.06.2021 (escenario 2) - Portugal, Alemania
- Figura 22 - Predicciones adicionales hasta el 18.06.2021 (escenarios 1 y 2) - Alemania
- Figura 23 - Resultados del análisis para regiones españolas y alemanas
- Figura 24 - Predicciones hasta el 18.04.2021 - Austria, Bélgica
- Figura 25 - Predicciones hasta el 18.04.2021 - Croacia, República Checa
- Figura 26 - Predicciones hasta el 18.04.2021 - Francia, Italia
- Figura 27 - Predicciones hasta el 18.04.2021 - España, Alemania
- Figura 28 - Ejemplos de ajuste poco satisfactorio - España, Italia

Lista de tablas

Tabla 1 - Datos necesarios para el análisis

Tabla 2 - Estimación numérica del efecto de las medidas de intervención en un país

Tabla 3 - Combinaciones de datos utilizados para el análisis

Tabla 4 - Medidas de intervención - agrupación en categorías para la comparación entre países

Tabla 5 - Número de reproducción básico (R_0) - resultados de este trabajo y de otros modelos

Tabla 6 - Datos utilizados para el análisis del efecto de las medidas de intervención

Tabla 7 - Datos utilizados para la comparación del efecto de las medidas entre distintos países

Tabla 8 - Medidas más efectivas en distintos países (22.01.2020 - 31.07.2020)

Tabla 9 - Medidas más efectivas en distintos países (22.01.2020 - 31.07.2020) - resumen

Tabla 10 - Medidas más efectivas en distintos países (01.08.2020 - 18.04.2021)

Tabla 11 - Medidas más efectivas en distintos países (01.08.2020 - 18.04.2021) - resumen

Tabla 12 - Medidas más efectivas en distintos países (22.01.2020 - 18.04.2021)

Tabla 13 - Datos utilizados para el cálculo de predicciones

Tabla 14 - Francia - Medidas de intervención hasta el 18.04.2021

Tabla 15 - Glosario

Tabla 16 - Significado de los nombres de las distintas medidas de intervención

Tabla 17 - Enlaces al servidor web - modelos

Tabla 18 - Enlaces al servidor web - efecto de las medidas de intervención

Tabla 19 - Enlaces al servidor web - comparación del efecto de las medidas de intervención

Tabla 20 - Enlaces al servidor web - predicciones

Tabla 21 - Herramienta informática - aspectos técnicos

1 Resumen

Los coronavirus son una extensa familia de virus que pueden causar enfermedades tanto en animales como en humanos. En los humanos, se sabe que varios coronavirus causan infecciones respiratorias que pueden ir desde el resfriado común hasta enfermedades más graves [1].

El *COVID-19* (*coronavirus disease 2019*) es la enfermedad causada por un nuevo tipo de coronavirus llamado *SARS-CoV-2*. La Organización Mundial de la Salud (*OMS*) tuvo conocimiento de la existencia del mismo a finales del año 2019, a partir de un informe sobre un grupo de casos de *neumonía viral* en Wuhan (China) [1].

La rápida extensión del *COVID-19* en todo el mundo hizo que la *OMS* declarara en enero de 2020 la situación como emergencia de salud pública de interés internacional, para declararla oficialmente como pandemia el 11 de marzo de 2020. En ese momento, el número de casos confirmados de la enfermedad ascendía a más de 118.000, y más de 4.000 personas habían muerto a causa de la misma [2]. De acuerdo con datos oficiales de la *OMS*, el 1 de febrero de 2021 el número de casos confirmados superaba ya los 102 millones, el de muertos los 2,2 millones [3]. El 31 de mayo de 2021, se habían alcanzado los 169 millones de casos confirmados y los 3,5 millones de muertos [4].

En este contexto, entender los mecanismos de expansión del *COVID-19* y analizar el posible efecto de distintos factores en su evolución pueden ser de utilidad, por ejemplo, a la hora de tomar decisiones sobre futuras medidas o de hacer previsiones sobre la necesidad a medio plazo de recursos médicos y hospitalarios.

Con este *Trabajo Fin de Máster*, se ha extendido un modelo existente de transmisión del *COVID-19* [6]. Se han desarrollado funcionalidades con las que se mejora el análisis de múltiples periodos de diferente velocidad de transmisión y se posibilita la inclusión de un número potencialmente ilimitado de medidas de intervención, lo que permite una descripción más realista en rangos de tiempo más extensos que abarcan varias olas de la pandemia.

Los buenos resultados obtenidos para un total de 32 países europeos, con datos que siguen patrones muy distintos, permiten concluir, por un lado, que el modelo utilizado es un buen método para describir la evolución de la enfermedad.

Por otro lado, a partir de los resultados, se han extraído conclusiones sobre el efecto de las distintas medidas que son consistentes con la información conocida. Por ejemplo, los resultados para la primera ola de la pandemia indican que las medidas más efectivas fueron las relacionadas con el confinamiento, las medidas aplicadas conjuntamente en marzo de 2020 y el uso obligatorio de mascarillas.

Desde el punto de vista metodológico, lo que aporta este trabajo y lo diferencia de otras alternativas existentes es, principalmente:

1. La posibilidad de introducir de forma sencilla y flexible la información sobre un número potencialmente ilimitado de medidas de intervención, tanto para el análisis de datos como para la obtención de predicciones.
2. La posibilidad de incluir información sobre eventos y periodos adicionales, no necesariamente relacionados con medidas de intervención, pero que pueden influir (tanto positiva con negativamente) en la evolución de la enfermedad. Ejemplo: periodos festivos.
3. La opción de parametrizar la forma en que se estimará el efecto de las distintas medidas, permitiendo las siguientes variantes:
 - Ajustar los datos asumiendo que una misma medida tiene el mismo efecto en todos sus periodos de aplicación (opción por defecto) o que el efecto de una medida pueda ser diferente en distintos periodos
 - Distintos parámetros que se pueden utilizar para cada medida individual por separado: asumir que el efecto de una medida es positivo (opción por defecto), que puede ser negativo o que éste tiene un valor fijo predeterminado
4. La opción de modificar de forma sencilla algunos aspectos del comportamiento de la herramienta, bien de forma global, bien para países específicos, a través de la modificación de ciertos parámetros, sin necesidad de hacer cambios en el software.
5. La forma de estimar y comparar el efecto de las distintas medidas de intervención y de presentar los resultados de forma sencilla.

2 Introducción

Este capítulo resume el planteamiento del trabajo que se presenta con este documento, cuyos detalles se desarrollan en los capítulos siguientes.

2.1 Contexto y justificación del Trabajo

Los coronavirus son una extensa familia de virus que pueden causar enfermedades tanto en animales como en humanos. En los humanos, se sabe que varios coronavirus causan infecciones respiratorias que pueden ir desde el resfriado común hasta enfermedades más graves [1].

El *COVID-19* (*coronavirus disease 2019*) es la enfermedad causada por un nuevo tipo de coronavirus llamado *SARS-CoV-2*. La *Organización Mundial de la Salud (OMS)* tuvo conocimiento de la existencia del mismo a finales del año 2019, a partir de un informe sobre un grupo de casos de *neumonía viral* en Wuhan (China) [1].

La rápida extensión del *COVID-19* en todo el mundo hizo que la *OMS* declarara en enero de 2020 la situación como emergencia de salud pública de interés internacional, para declararla oficialmente como pandemia el 11 de marzo de 2020. En ese momento, el número de casos confirmados de la enfermedad ascendía a más de 118.000, y más de 4.000 personas habían muerto a causa de la misma [2]. De acuerdo con datos oficiales de la *OMS*, el 1 de febrero de 2021 el número de casos confirmados superaba ya los 102 millones, el de muertos los 2,2 millones [3]. El 31 de mayo de 2021, se habían alcanzado los 169 millones de casos confirmados y los 3,5 millones de muertos [4].

En este contexto, entender los mecanismos de expansión del *COVID-19* y analizar el posible efecto de distintos factores en su evolución pueden ser de utilidad, por ejemplo, a la hora de tomar decisiones sobre futuras medidas o de hacer previsiones sobre la necesidad a medio plazo de recursos médicos y hospitalarios.

Numerosos estudios han abordado el análisis de distintos aspectos relacionados con el *COVID-19*. Entre ellos destacan estudios médicos y epidemiológicos, con los que se busca una explicación a los mecanismos de la enfermedad y cuyos resultados han contribuido, por ejemplo, al desarrollo de vacunas efectivas. Otros estudios han centrado sus esfuerzos en analizar factores sociales que influyen en la propagación de la enfermedad y cuya modificación puede contribuir a frenarla.

Los estudios de carácter médico han constatado que el contagio de la enfermedad se produce por el contacto con individuos infectados. De acuerdo con la *OMS*, “los datos actuales sugieren que el virus se propaga principalmente por medio de gotículas respiratorias entre personas que estén en contacto cercano.” [5]. Las medidas no farmacológicas aplicadas para frenar la propagación de la enfermedad se centran en evitar o reducir el contacto entre personas. Éstas van desde medidas estrictas de confinamiento hasta el uso de mascarillas o la obligación de mantener una distancia mínima con otras personas.

Con este *Trabajo Fin de Máster* se ha extendido un modelo existente de transmisión del *COVID-19* [6]. Se han desarrollado funcionalidades con las que se mejora el análisis de múltiples periodos de diferente velocidad de transmisión y se posibilita la inclusión de un número potencialmente ilimitado de medidas de intervención, lo que permite una descripción más realista en rangos de tiempo más extensos que abarcan varias olas de la pandemia.

2.2 Objetivos del Trabajo

El modelo de transmisión del *COVID-19* desarrollado con este trabajo permite estimar el impacto de las medidas o intervenciones no farmacológicas en el desarrollo de la enfermedad en distintos países y en periodos extensos de tiempo.

Los objetivos del trabajo se dividen en dos categorías, metodología y análisis de datos:

1. En el aspecto **metodológico**, se ha definido como objetivo el desarrollo de una herramienta informática que implemente un modelo estadístico para el análisis de datos de *COVID-19*.
Ésta ha de permitir la inclusión en el análisis de múltiples periodos de diferente velocidad de transmisión, así como un número arbitrario de medidas de intervención.
La finalidad es obtener un modelo que permita describir la evolución de la enfermedad en rangos de tiempo extensos y estimar el efecto de las distintas medidas aplicadas.
2. La parte de **análisis de datos** consiste en la aplicación del modelo estadístico a los datos de países en los que la pandemia ha evolucionado de forma muy diferente con los siguientes objetivos:
 - Comprobar si el modelo es lo suficientemente robusto para ajustarse bien a datos que siguen patrones distintos
 - Evaluar y comparar el efecto de las distintas medidas en diferentes países
 - Hacer predicciones sobre la evolución de la enfermedad en varios países bajo distintos supuestos

2.3 Enfoque y método seguido

El modelo de transmisión del *COVID-19* que se desarrollado con este trabajo analiza la evolución de la enfermedad desde un punto de vista estadístico y de análisis de datos.

Los factores médicos y epidemiológicos no forman parte del trabajo, por lo que el enfoque y la metodología elegidos han sido los propios de un trabajo de análisis estadístico desarrollado con la ayuda de herramientas informáticas.

En el momento de iniciar este trabajo, en febrero de 2021, todavía había aspectos relevantes relacionados con el desarrollo de la enfermedad sobre los que no había información suficiente, entre otros:

- Efecto de nuevas variantes del virus
- Efecto de la vacuna, que empezó a suministrarse en diciembre de 2020

Por ello, dada la actualidad y el carácter dinámico del desarrollo de la enfermedad, el modelo resultante es necesariamente limitado. El planteamiento general ha sido pragmático y basado en algunos supuestos y simplificaciones que se asumen aceptables y que se documentan más adelante.

Esto no le resta utilidad al trabajo, ya que éste se ha enfocado de forma que el resultado sea un modelo modular, flexible y fácil de extender para incluir factores adicionales, que podrían tenerse en cuenta en futuras versiones del mismo.

De los distintos métodos para hacer análisis y predicciones sobre una serie temporal (*time series analysis / forecasting*), se ha elegido el de un modelo ya existente [6], creado a su vez a partir de un estudio previo [7]. Éste utiliza un modelo mecanístico basado en el análisis bayesiano [8], que define la evolución del número de casos y de muertes por *COVID-19* en función del efecto de las distintas medidas de intervención, y que describe muy bien la evolución de la enfermedad en distintos países europeos desde el inicio de la pandemia hasta la fecha de su publicación (mayo de 2020).

Este *Trabajo Fin de Máster* es una extensión del modelo mecanístico mencionado [6]. Entre los principales cambios se encuentra la posibilidad de introducir de forma sencilla un número potencialmente ilimitado de periodos de intervención. Además, se han añadido opciones que permiten un uso más flexible y que facilitan la extensión del modelo para considerar factores adicionales. Asimismo, se han desarrollado nuevas funciones para comparar el efecto de las diferentes medidas, tanto dentro de un país como entre países distintos.

Los detalles se describen en el capítulo “4 Metodología”. Un resumen de algunas diferencias de este modelo con trabajos previos se incluye en “3.3 Aportación de este Trabajo Fin de Máster”.

2.4 Planificación del Trabajo

El gráfico de la página siguiente muestra las distintas tareas en las que se ha desglosado este *Trabajo Fin de Máster* y su planificación. Éste refleja tanto las actividades concretas como el planteamiento general del proyecto, dividido en cuatro fases:

- Preparación Elección del tema del trabajo y definición de los objetivos, búsqueda bibliográfica, elección de las fuentes de datos
- Planificación Definición de los objetivos concretos, desglose en tareas y planificación, recopilación de datos
- Desarrollo Desarrollo de una herramienta informática para el análisis de datos de COVID-19, análisis de datos haciendo uso de la herramienta
- Finalización Documentación del trabajo realizado y de los resultados obtenidos

Para que el gráfico sea más legible, se prescinde de mostrar las fechas de inicio y final de cada una de las tareas, mostrando únicamente la fechas que se consideran necesarias para entender el planteamiento general, la planificación de los bloques principales y los hitos. Se definen como hitos algunos pasos y resultados intermedios, coincidentes con las fechas de entrega de las distintas *Pruebas de Evaluación Continua (PEC)*.

El trabajo se ha organizado de forma que las tareas se distribuyan de forma razonable teniendo en cuenta las dependencias entre ellas y las distintas fechas de entrega. El tiempo total estimado supone un promedio de trabajo de aproximadamente tres horas al día, si bien la carga de trabajo ha variado en distintas fases del proyecto.

El tiempo previsto para cada tarea se definió al inicio del trabajo a partir de estimaciones que se consideran realistas. Dada la experiencia previa, las tareas relacionadas con la programación informática se estimaron con valores más ajustados, mientras que para las tareas de análisis de datos y de documentación las estimaciones contienen un porcentaje adicional que puede considerarse de reserva.

Hay actividades recurrentes que se han ejecutado durante todo el proyecto y que no se mencionan de forma explícita, sino que se incluyen implícitamente como parte de las tareas definidas. Entre éstas se encuentra, por ejemplo, la consulta de bibliografía sobre el tema objeto del trabajo.

Además, durante el progreso del trabajo, se consideró adecuada la creación de un servidor *web* [9] para almacenar los distintos resultados y ofrecer un acceso a los mismos de forma sencilla e intuitiva. Esta tarea no se incluye de forma explícita en el plan de trabajo, sino que se considera parte de las tareas relacionadas con el análisis de datos y con la preparación de la documentación.

Definición y planificación de las distintas tareas

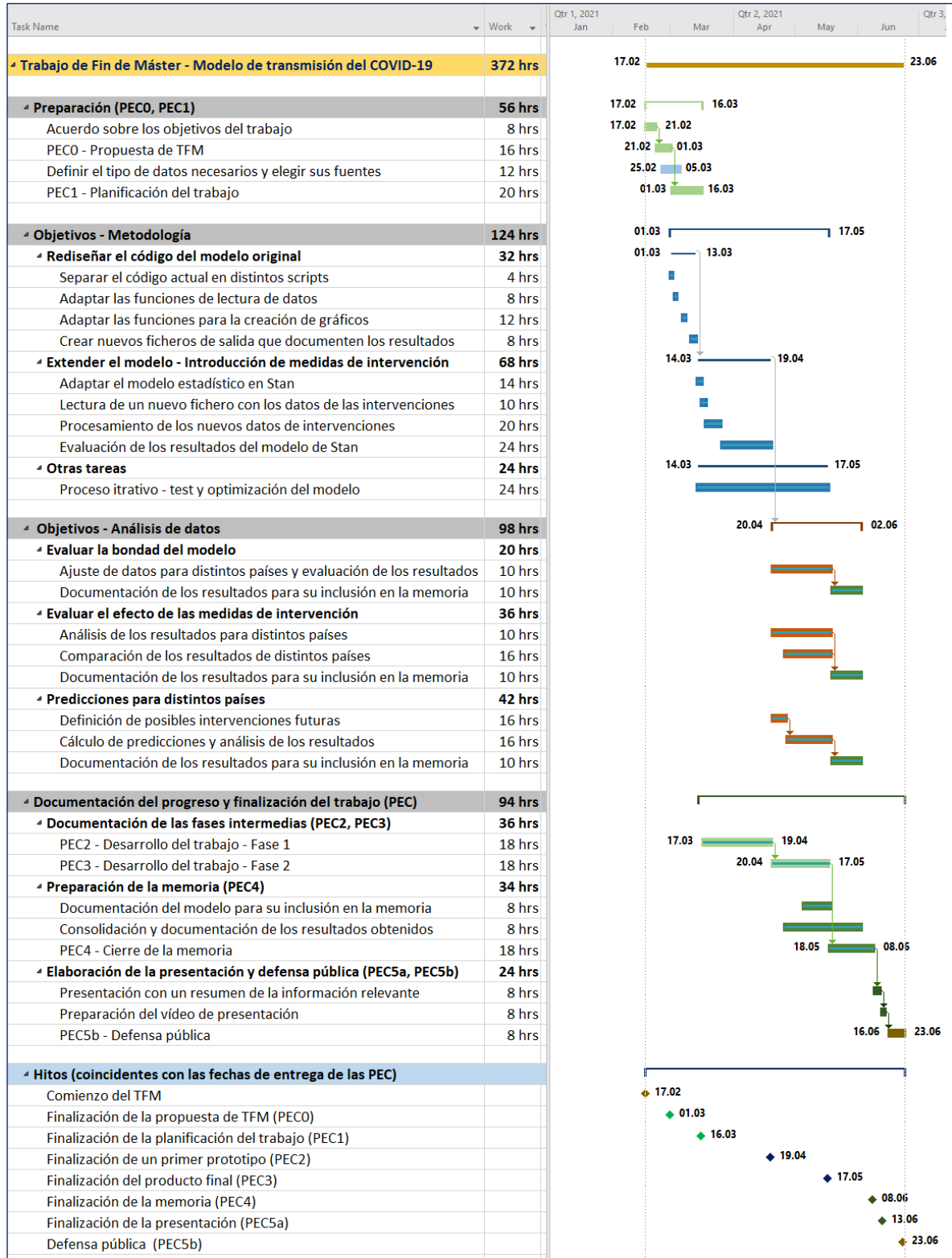


Figura 1 - Definición y planificación de las tareas del Trabajo Fin de Máster

2.5 Breve resumen de los productos obtenidos

El resultado de este trabajo es, por un lado, una herramienta informática de uso sencillo con la que se pueden obtener estimaciones sobre el efecto de las distintas medidas no farmacológicas en la evolución de la pandemia. Por otro lado, se ofrecen los resultados de aplicar la herramienta informática a datos de distintos países europeos, con lo que se obtienen conclusiones sobre el efecto de las medidas de intervención y predicciones sobre el desarrollo futuro de la enfermedad en distintos escenarios.

La información relevante se incluye en este documento y se completa con otros productos, de forma que el resultado de este *Trabajo Fin de Master* consta de las siguientes partes:

- Memoria final (este documento) ([🔗](#))
- Código y otros componentes de la herramienta informática ([🔗](#))
- Documentación de los detalles de las implementación y manual de uso de la herramienta ([🔗](#))
- Resultados completos obtenidos del uso de la herramienta informática para datos de diferentes países ([🔗](#))

Tanto este documento como los otros productos y resultados se encuentran en un **servidor web** [9]:

Enlace <http://jblecua.spdns.org:9095/>
Usuario *tfm*
Password *tfm*

2.6 Breve descripción de los otros capítulos

Los capítulos siguientes describen los detalles del proyecto desarrollado con este *Trabajo Fin de Máster*.

El capítulo “3 Estado del arte” contiene información sobre el estado actual del análisis estadístico de datos de *COVID-19*, documentando las características principales de trabajos previos y explicando lo que aporta este trabajo como complemento a estudios ya existentes.

En el capítulo “4 Metodología” se describen los detalles sobre los aspectos metodológicos del trabajo, consistentes en el desarrollo de una herramienta informática para el análisis de datos de *COVID-19*.

El capítulo “5 Resultados” documenta los resultados obtenidos del análisis de datos de *COVID-19* para distintos países haciendo uso de la herramienta informática desarrollada con este trabajo.

En “6 Discusión” se comentan los resultados de los capítulos anteriores bajo la perspectiva de su relación con los objetivos definidos.

El capítulo “7 Conclusiones” cierra la parte principal de la memoria. Éste incluye, por un lado, una evaluación sobre el grado de cumplimiento de los objetivos planteados. Además, se documentan las conclusiones principales de los resultados de los análisis y se comentan posibles limitaciones del producto obtenido que podrían subsanarse en trabajos futuros.

3 Estado del arte

Este capítulo describe el estado actual (*state of the art*) de investigaciones y estudios relacionados con el tema objeto de este *Trabajo Fin de Máster*, que consiste en el análisis de datos de *COVID-19* con la finalidad de estimar el efecto de las medidas no farmacológicas aplicadas en los distintos países.

Se comentan brevemente algunos trabajos de temática similar y se explica lo que aporta este trabajo y lo que lo diferencia de los estudios y modelos existentes.

3.1 Número de reproducción (R_t)

Un concepto central, común a la mayoría de los modelos que describen la evolución de la enfermedad, es el del parámetro R_t (**número de reproducción**), llamado *reproduction number* en inglés, a partir del que se estiman el número de infecciones por *COVID-19*.

Los estudios diferencian los conceptos de *número de reproducción básico* (R_0) y *número de reproducción efectivo* (R_t), siendo el primero el valor inicial y el segundo el valor que cambia a lo largo del tiempo. Una explicación detallada y más genérica se puede encontrar en [The Mathematics of Infectious Diseases](#) [10].

Su uso en el contexto de este trabajo se explica en el capítulo “4 Metodología” pero, de forma general, se puede decir que este parámetro es una indicación del ritmo de propagación de la enfermedad. El *número de reproducción* representa el promedio de casos *secundarios* que se producirán en un grupo de población no inmune a partir de un caso *primario* [11]. Esto se podría reformular diciendo que este valor representa el número de personas a las que contagia, en promedio, un individuo infectado. Valores mayores de 1 indican una tendencia creciente en el número de casos, mientras que un descenso por debajo de 1 se asocia con la contención en la expansión de la pandemia [10], [11].

3.2 Estudios previos y métodos existentes

A continuación, se comentan algunos trabajos en los que se analiza el efecto de las medidas de intervención en la evolución de la enfermedad, describiendo brevemente aspectos de su metodología y posibles limitaciones:

1. Entre los estudios que analizan el efecto de las medidas de intervención, se encuentran los dos, ya mencionados, en los que se basa este trabajo [6], [7]. Ambos utilizan para el análisis métodos de inferencia bayesiana [8].

El planteamiento de [Fernández-Recio](#) [6] es, en lo fundamental, el que se utiliza en este *Trabajo Fin de Máster*, y se describe en detalle en los siguientes capítulos.

El estudio de [Flaxman et col.](#) [7] es la base del anterior. Sin embargo, la metodología y algunos de los supuestos son ligeramente distintos. En concreto, en este estudio se analizan conjuntamente los datos de varios países, asumiendo que el efecto de las distintas medidas de intervención será el mismo en todos ellos. Este supuesto, como se verá más adelante, se contradice con los resultados obtenidos en este *Trabajo Fin de Máster*, en el que se analiza cada país por separado y se compara el efecto de las distintas medidas en los distintos países.

Ambos estudios obtienen buenos resultados pero, como se ha comentado antes, el número de periodos de intervención que pueden analizar está limitado a tres, algo que se solventa con este *Trabajo de Fin de Máster*.

2. Entre los trabajos consultados, el de [Haug, et col.](#) [12], que utiliza tanto métodos de regresión como de aprendizaje automático (*machine learning*), es el que presenta resultados cuantitativos más claros y detallados sobre el efecto de distintas medidas de intervención en distintos países.

Si bien la nomenclatura es distinta a la utilizada en este trabajo, los resultados entre ambos son consistentes al mostrar, entre las medidas más efectivas, las relacionadas con el confinamiento, con la limitación de reuniones y contactos sociales y con el cierre de instituciones educativas.

Aparte de los aspectos metodológicos, la principal diferencia a la hora de evaluar los resultados es que el trabajo mencionado [12] utiliza como criterio para cuantificar el efecto de las intervenciones la diferencia, en valor absoluto, del valor de R_t tras la aplicación de cada intervención. Sin embargo, en este *Trabajo Fin de Máster*, como en otros estudios, se considera más adecuado utilizar para este fin el factor de reducción sobre el valor de R_t , más apropiado para análisis más genéricos y que permite la comparación del efecto de distintas medidas también en periodos con una menor velocidad de transmisión.

El uso del número de casos documentados de *COVID-19* como punto de partida para el análisis, así como el hecho de cuantificar el efecto de las medidas a partir de diferencias absolutas en el valor de R_t , se consideran limitaciones de este trabajo.

3. Desde el punto de vista metodológico, el planteamiento de [Linka et col. \[13\]](#) es muy interesante y algo más elaborado que el utilizado en este *Trabajo Fin de Máster*.

Se utiliza un modelo *SEIR* [13], [14] y se define el número de reproducción efectivo (R_t) como función de una tangente hiperbólica, donde se incluyen variables que describen el efecto en el tiempo de distintas medidas de intervención. Los valores de R_t se estiman con métodos de inferencia bayesiana [8]. Se documentan los valores estimados de R_t para distintos países europeos al inicio de la pandemia y el 10 de mayo de 2020, siendo éstos consistentes con los de otros estudios, y también con los de este *Trabajo Fin de Máster*.

Los resultados muestran una correlación clara entre las medidas de intervención y la evolución de la enfermedad. Sin embargo, no está claro si este método funcionará bien en periodos de tiempo más extensos y en los que las medidas de intervención no tengan un efecto tan significativo como las aplicadas al inicio de la pandemia.

Una posible debilidad del modelo reside en que hace sus estimaciones a partir del número de casos documentados de *COVID-19*. La relación entre este valor y el número real de infectados depende de la capacidad de detección, que varía mucho entre los distintos países y que ha ido cambiando con el avance de la pandemia [15].

4. El estudio de [Giordano et col. \[16\]](#) analiza los datos de Italia entre el 20 de febrero y el 5 de abril de 2020, y tiene su interés por la base matemática del método utilizado.

Utiliza un método en el que la población se divide en ocho grupos distintos (entre ellos, por ejemplo, están el grupo de infectados y el de curados), denominado *SIDARTHE* por las iniciales de cada uno de ellos. Es un planteamiento similar, aunque más complejo, al del método *SEIR* [13], [14] mencionado antes.

La evolución del número de individuos de cada grupo se define a partir de ocho ecuaciones diferenciales, dependientes de una serie de parámetros, entre los que se encuentran, por ejemplo, la proporción de contagios o la probabilidad de detección, y que pueden variar con la aplicación de las distintas medidas de intervención.

El efecto de las medidas se estima a partir de las combinaciones de parámetros que hacen que las predicciones del modelo se adapten lo mejor posible a los datos conocidos, utilizando como criterio para la estimación de la bondad del ajuste la suma de los cuadrados de los errores.

Las distintas combinaciones de parámetros se obtienen por métodos de simulación numérica.

Este método ofrece un buen ajuste y resultados plausibles sobre el efecto de las medidas de intervención consideradas (confinamiento, distanciamiento social, aumento de pruebas para la detección de individuos infectados y rastreo de contactos de las personas infectadas) en el periodo de tiempo analizado. Sin embargo, por su complejidad, no es sencillo extenderlo para incluir un número alto de medidas de intervención, lo que limita mucho su utilidad para periodos extensos de tiempo en los que, como se verá, el número de medidas de intervención aplicadas en los distintos países ha sido muy alto.

5. El trabajo de [Brauner et col. \[17\]](#), publicado en diciembre de 2020, extiende el modelo del estudio de [Flaxman et col. \[7\]](#), mejorando algunos aspectos de éste.

Por un lado, se analizan los datos de cada país por separado, de forma que no se asume que el efecto de las intervenciones será el mismo en todos ellos. Por otro lado, se modifican algunos aspectos del modelo para hacerlo más flexible.

Por su metodología, éste es el estudio que más se aproxima a lo desarrollado con este *Trabajo Fin de Máster*, siendo el principal punto el común el hecho de permitir más de tres intervenciones para el análisis de los datos.

Sin embargo, en algunos aspectos este estudio es algo más limitado. Por un lado, se analizan datos únicamente entre enero y mayo de 2020, y no se consideran los periodos y la evolución de la enfermedad una vez se han desactivado las medidas. Por otro lado, se analiza el efecto de un máximo de siete medidas de intervención. Esta limitación es aceptable para el periodo analizado, pero es demasiado restrictiva para periodos más amplios en los que, como se verá, el número de medidas aplicadas es mucho mayor.

El estudio, sin embargo, contempla aspectos no considerados en el trabajo *Fin de Máster*, como es, por ejemplo, el análisis de datos de países no europeos, para lo que se utilizan varias fuentes de información sobre las medidas de intervención.

6. Como ejemplo de una metodología distinta a las mencionadas hasta ahora, pero habitual para el análisis de series temporales (*time series analysis*), se menciona el trabajo de [Elhassan et col. \[18\]](#).

De todos los trabajos mencionados, éste y el de [Bendavid et col. \[19\]](#) son los únicos que no utilizan el parámetro R_t .

El trabajo analiza la evolución de la enfermedad en Arabia Saudí ajustando los datos con dos modelos distintos: *ARIMA (Autoregressive integrated moving average)* y *Modelo de crecimiento logístico (logistic growth model)*.

En ambos casos, se trata de métodos de ajuste puramente matemático, en los que el efecto de las medidas de intervención queda reflejado en los valores óptimos de ciertos factores en las ecuaciones que definen el modelo.

Por su planteamiento, estos métodos no permiten cuantificar de forma tan clara el efecto de distintas medidas de intervención, especialmente en periodos extensos y con muchas medidas. Sin embargo, ofrecen un buen ajuste para periodos breves de tiempo y para predicciones a corto plazo.

7. El trabajo de [Bendavid et col. \[19\]](#) desarrolla un método interesante para *aislar* el efecto de las intervenciones individuales con un mayor impacto y luego cuantificarlo a través de un modelo de regresión.

Se asume que se puede extraer el efecto de medidas individuales a partir de la comparación del impacto de conjuntos de medidas entre distintos países.

Es una metodología novedosa con la que se intentan mitigar algunas de las limitaciones de otros estudios.

Sin embargo, la forma de extraer el efecto de las medidas con mayor impacto asume que el resto de las medidas tendrán efectos similares en los distintos países. Esto se considera cuestionable, como ya se ha dicho en relación con el estudio de [Flaxman et col. \[7\]](#).

8. El artículo de [Bryant et col. \[20\]](#) es una voz crítica con la metodología bayesiana utilizada en algunos de los estudios anteriores.

En particular, analiza de forma crítica los resultados del estudio de [Flaxman et col. \[7\]](#). Como se ha dicho, éste es la base del método de [Fernández-Recio \[6\]](#), a partir del que se ha desarrollado la metodología de este *Trabajo Fin de Máster*.

Los autores llegan a la conclusión de que el modelo asigna un efecto demasiado alto a la última medida aplicada, independientemente de cuál sea ésta. En la fecha de publicación de [Flaxman et col. \[7\]](#), la última intervención en la mayoría de los países había sido el confinamiento, y este estudio concluye que, por este motivo, el modelo asigna a esta medida un efecto desproporcionadamente alto que, además, no explica el descenso del valor de R_t en Suecia, similar al de otros países sin haber aplicado medidas de confinamiento.

Los argumentos son consistentes y deben tenerse en cuenta. Pero este *Trabajo Fin de Máster* muestra que la metodología bayesiana utilizada, con algunas limitaciones que se describen más adelante, es adecuada, y que las inconsistencias detectadas en el estudio original [7] tienen una causa que se ha eliminado en el trabajo presente y también en otros trabajos posteriores, por ejemplo en el de [Brauner et col. \[17\]](#). En concreto, las inconsistencias se deben, principalmente, al hecho de que el estudio original [7] analiza conjuntamente los datos de varios países, asumiendo que el efecto de las distintas medidas de intervención será el mismo en todos ellos. Como se ha dicho antes, este supuesto contradice los resultados del análisis de cada país por separado, como se verá más adelante.

3.3 Aportación de este Trabajo Fin de Máster

Dada la cantidad de trabajos dedicados al análisis de datos de *COVID-19*, la lista del apartado anterior no es completa. Aun así, es suficiente como muestra de la variedad de métodos de análisis. Un resumen más exhaustivo se puede encontrar en un estudio del *Robert Koch-Institut* [21].

Todos los métodos tienen su justificación y obtienen buenos resultados, pero también tienen ciertas limitaciones. Algunas están impuestas por el momento en el que se publicaron los estudios (por ejemplo, falta de datos), otras, por la propia definición del método de análisis (por ejemplo, límite del análisis a un periodo reducido de tiempo).

Con este *Trabajo Fin de Máster*, se continúa el trabajo de estudios anteriores [6], [7] para mitigar algunas de sus limitaciones. Se obtiene como resultado un método flexible, adecuado para el análisis de datos de *COVID-19* en periodos largos y con un número elevado de medidas de intervención.

Aunque no exento de limitaciones, que se comentarán más adelante, este *Trabajo Fin de Máster* extiende y complementa trabajos anteriores con funciones y opciones no disponibles hasta ahora en una única herramienta. Esto permite extender el alcance de algunos de los análisis.

Entre los aspectos metodológicos que diferencian a este trabajo de estudios anteriores se pueden destacar los siguientes:

1. La posibilidad de introducir de forma sencilla y flexible la información sobre un número potencialmente ilimitado de medidas de intervención, tanto para el análisis de datos como para la obtención de predicciones.
 2. La posibilidad de incluir información sobre eventos y periodos adicionales, no necesariamente relacionados con medidas de intervención, pero que pueden influir (tanto positiva como negativamente) en la evolución de la enfermedad. Ejemplo: periodos festivos.
 3. La opción de parametrizar la forma en que se estimará el efecto de las distintas medidas, permitiendo las siguientes variantes:
 - Ajustar los datos asumiendo que una misma medida tiene el mismo efecto en todos sus periodos de aplicación (opción por defecto) o que el efecto de una medida pueda ser diferente en distintos periodos
 - Distintos parámetros que se pueden utilizar para cada medida individual por separado: asumir que el efecto de una medida es positivo (opción por defecto), que puede ser negativo o que éste tiene un valor fijo predeterminado
 4. La opción de modificar de forma sencilla algunos aspectos del comportamiento de la herramienta, bien de forma global, bien para países específicos, a través de la modificación de ciertos parámetros, sin necesidad de hacer cambios en el software.
 5. La forma de estimar y comparar el efecto de las distintas medidas de intervención y de presentar los resultados de forma sencilla.
 6. Simplicidad del modelo matemático utilizado, en comparación con la mayoría de los trabajos mencionados en el apartado anterior. Esto no es mérito de este trabajo, sino de los estudios en los que se basa [6], [7].
-

4 Metodología

Distintos estudios han analizado la evolución del *COVID-19* con la finalidad de evaluar el efecto de las medidas de intervención y de hacer predicciones sobre la evolución futura de la enfermedad (ver “3.2 Estudios previos y métodos existentes”).

Un método novedoso, en cuanto a su fundamento estadístico, es el modelo de transmisión propuesto por [Fernández-Recio \[6\]](#), basado en el análisis bayesiano [8]. Para ello, el autor adaptó un modelo ya existente [7] para analizar la evolución de la enfermedad en algunos países europeos y el efecto de las medidas de intervención aplicadas.

Este modelo ofrece muy buenos resultados para los datos analizados pero, por el momento en el que se creó, tiene algunas limitaciones, debidas, principalmente, a la falta de información existente hasta entonces, por ejemplo:

- Datos sobre la evolución del *COVID-19* a más largo plazo
- Información sobre algunas de las intervenciones no farmacológicas introducidas en fechas posteriores

Desde el inicio de la pandemia, se han aplicado distintas medidas o intervenciones no farmacológicas para frenar la expansión de la enfermedad. Ejemplos de las mismas son el distanciamiento social, el uso de mascarillas, el teletrabajo o el confinamiento.

Una de las limitaciones del modelo de transmisión mencionado [6] es que el número de intervenciones que puede procesar como parte de su análisis está limitado a un máximo de tres. Esta restricción era aceptable en el modelo inicial, pero limita mucho la capacidad de análisis para datos posteriores, ya que durante los últimos meses se han introducido, en periodos distintos, un número bastante elevado y variado de medidas.

Por ejemplo, de acuerdo con la información disponible en el [European Centre for Disease Prevention and Control \[22\]](#), entre el 9 de marzo de 2020 y el 30 de marzo de 2021, en España se aplicaron un total de 35 medidas distintas, habiendo periodos en los que varias de estas medidas estaban activas simultáneamente, con medidas que se han desactivado y vuelto a activar más adelante, etc., resultando en un total de 43 periodos de intervención.

Para que el modelo sea más robusto y sus predicciones más fiables, se hace necesario extenderlo de forma que sea posible considerar para el análisis un número arbitrario de intervenciones y poder analizar el efecto de la activación y desactivación de cada una de ellas en el desarrollo posterior de la enfermedad.

Desde el punto de vista metodológico, este trabajo consiste, por tanto, en la extensión de un modelo mecanístico basado en el análisis bayesiano para que sea posible introducir un número potencialmente ilimitado de periodos de intervención. Los detalles se describen en los siguientes apartados.

4.1 Número de reproducción (R_t)

En el capítulo “3.1 Número de reproducción (R_t)” se ha introducido el concepto de **número de reproducción (R_t)** en el contexto del análisis de la evolución del COVID-19. Como se ha dicho, su valor representa el promedio de casos *secundarios* que se producirán en un grupo de población no inmune a partir de un caso *primario* [11] o, de forma simplificada, el número de personas a las que contagia, en promedio, un individuo infectado. Valores mayores de 1 indican una tendencia creciente en el número de casos, mientras que un descenso por debajo de 1 se asocia con la contención en la expansión de la pandemia [10], [11].

El valor al principio de la pandemia, denominado *número de reproducción básico* (R_0), está relacionado con el ritmo al que aumentaría el número de casos en un grupo de población no inmune si no se tomara ningún tipo de medida. Su valor es “esencialmente la combinación de tres factores: el número (medio) de contactos diarios que tiene una persona infectada, la probabilidad de transmisión de la enfermedad durante cada contacto y el número (medio) de días que una persona infectada tiene capacidad de contagiar. Mientras que el último factor depende de las características biológicas de la enfermedad, los dos primeros factores dependen en gran medida de los hábitos sociales” [11]. Es razonable suponer, por tanto, que el valor de R_0 será distinto dependiendo del país.

La aplicación de distintas medidas de intervención supone una reducción en la transmisión de la enfermedad, con lo que el *número de reproducción efectivo* (R_t) resultante es menor que el valor inicial.

En el estudio de *Brauner et col.* [17] se dice que “la desactivación de una medida de intervención no farmacológica no implica que el ritmo de transmisión volverá a su nivel original” [17]. Sin embargo, si bien puede considerarse una simplificación, en este trabajo se asume que la desactivación de una medida supone un aumento en el valor R_t en la misma proporción en que éste disminuyó con su aplicación. Esto coincide con el planteamiento de otros trabajos [6], [7] y, como se verá, permite obtener buenos resultados de ajuste.

Bajo el supuesto anterior, la vuelta a la *normalidad*, definida aquí informalmente como el momento en el que se desactiven todas las medidas no farmacológicas, supondrá la vuelta al *número de reproducción básico* (R_0). Aun así, el *número de reproducción efectivo* en ese momento será menor, ya que una parte de la población estará inmunizada, bien por haber pasado la enfermedad, bien por haber sido vacunada, y, por tanto, no será susceptible de ser contagiada.

El *número de reproducción efectivo* en el momento de la vuelta a la *normalidad* será $R_t = R_0 (1 - P)$, siendo P la proporción de población inmune [11]. Por tanto, cuando la proporción de población inmune supere el valor $1 - 1/R_0$, el *número de reproducción efectivo*, una vez desactivadas todas las medidas, será menor de 1 y la enfermedad dejará de extenderse. Esto muestra la importancia del valor de R_0 , a partir del cual se puede estimar el porcentaje de población que debe estar vacunada para poder volver a la *normalidad* sin riesgo de una nueva ola de la pandemia.

La estimación del efecto de las medidas de intervención, de acuerdo con el modelo mecanístico utilizado en este trabajo, requieren también estimaciones de los valores de R_0 . Los resultados obtenidos para R_0 se resumen en el apartado “5.4 Estimaciones para el número de reproducción básico (R_0)”, donde se comentan y se comparan con los de otros estudios.

4.2 Planteamiento

Con el método estadístico que se describe en el siguiente apartado, se analizan los datos existentes y se estima el efecto de las distintas medidas en la evolución del número de infecciones y muertes por COVID-19. Con esta información, se pueden hacer predicciones sobre el efecto de posibles intervenciones futuras en la evolución de la enfermedad.

El efecto de cada medida queda reflejado en su contribución parcial a la reducción del valor de R_t (*número de reproducción*), a partir del que se estiman el número de infecciones y de muertes en una fecha dada. Ver detalles en [Fernández-Recio \[6\]](#) y, de forma resumida, en el siguiente apartado.

Para que la herramienta sea más genérica, se contempla también la opción de definir medidas o eventos que puedan tener un efecto negativo en el desarrollo de la enfermedad, es decir, que puedan tener como efecto no una reducción, sino un aumento del valor de R_t (*número de reproducción*).

Con las estimaciones sobre las contribuciones parciales de cada medida a los cambios del valor de R_t , se pueden hacer predicciones sobre el efecto de la aplicación de medidas futuras en la evolución de la enfermedad, asumiendo que el efecto de medidas futuras se puede obtener a partir de las estimaciones obtenidas para medidas pasadas.

El análisis de los datos y la obtención de predicciones se hacen, en todos los casos, para cada país por separado. Para ello se requiere, como punto de partida, una serie de datos cuyos detalles se explican en “4.5 Datos necesarios para el análisis”.

El resultado del análisis de los datos existentes es un **modelo**, en el que se obtienen estimaciones que cuantifican el efecto de cada una de las medidas sobre la evolución del valor de R_t (*número de reproducción*). Este modelo puede utilizarse para hacer **predicciones** sobre la evolución futura de la enfermedad. Se puede definir el rango de tiempo sobre el que hacer predicciones teniendo en cuenta distintos escenarios en relación con las medidas de intervención, por ejemplo:

- Desactivación de medidas de intervención
- Activación de nuevas medidas
- Definición de eventos que puedan tener un efecto negativo sobre la evolución de la enfermedad (es decir, que puedan tener como consecuencia un aumento del valor de R_t).

Con la información de modelos generados para varios países, se puede **comparar el efecto de las distintas medidas**. Esto incluye tanto la evaluación del efecto de las medidas en cada país por separado como la comparación del efecto de medidas individuales entre distintos países.

Por tanto, las funciones principales de la herramienta informática con la que se implementa la funcionalidad descrita son:

- La creación de modelos estadísticos que se ajusten lo mejor posible a los datos existentes
- Utilizar modelos creados previamente para hacer predicciones en distintos escenarios
- Comparar el efecto de las distintas medidas de intervención

4.3 Fundamento estadístico - Modelo mecanístico

De los distintos métodos para hacer análisis y predicciones sobre una serie temporal (*time series analysis / forecasting*), se ha elegido el de un modelo ya existente [6], creado a su vez a partir de un estudio previo [7]. Éste consiste en un modelo mecanístico basado en el análisis bayesiano [8].

Se utilizan distintos componentes para obtener una estimación del número de muertes por la enfermedad en una fecha dada. La fiabilidad del ajuste puede evaluarse comparando las estimaciones con los datos reales de muertes, disponibles en distintas fuentes.

La estimación del número de nuevos contagios (infecciones) en una fecha dada (I_j) se obtiene en función del número de infecciones de cada uno de los días previos, a partir de una distribución *Serial Interval (SI)* y del llamado *número de reproducción (R_t)*, de acuerdo con la siguiente fórmula:

$$I_j = \sum_{k=1}^{j-1} I_k \cdot (R_t)_j \cdot SI_{j-k}$$

Por otra parte, el número de muertes en un día dado (D_i) se obtiene a partir del número de infecciones de días previos, de acuerdo con una distribución y un parámetro que describen, básicamente, la probabilidad de muerte de un infectado, *Infection-To-Death (ITD)* e *Infection Fatality Ratio (IFR)*:

$$D_i = \sum_{j=1}^{i-1} I_j \cdot IFR \cdot ITD_{i-j}$$

Los detalles sobre el significado de las distintas distribuciones (*SI*, *ITD*) y parámetros (*IFR*) pueden encontrarse en [Fernández-Recio \[6\]](#) y se comentan brevemente en “4.5 Datos necesarios para el análisis”. Se asume que los valores de estas distribuciones y parámetros, basados en datos empíricos, son conocidos, como también lo son los datos diarios de muertes por *COVID-19*.

La finalidad es encontrar un modelo que describa la evolución del *número de reproducción (R_t)* y que permita cuantificar el efecto de las distintas medidas (intervenciones) sobre el mismo, de forma que las predicciones sobre el número de muertes se adapten lo mejor posible a los datos reales.

Se define un modelo en el que el valor de R_t (*número de reproducción efectivo*) en un día dado se obtiene como el producto de un valor inicial R_0 (*número de reproducción básico*), que se estima durante el proceso de ajuste de los datos, y un factor que incluye el efecto de cada una de las medidas de intervención activas en el día correspondiente.

El efecto de una medida dada i queda reflejado en el valor de un parámetro α_i , que se asume constante durante todo el periodo en que una medida está activa y que se aplica para el cálculo de R_t en ese periodo. Los valores de los parámetros α_i se estiman durante el proceso de ajuste.

Para cada día j , el valor resultante de R_t se obtiene a partir de la siguiente fórmula, siendo α_i el valor correspondiente al efecto de cada una de las medidas activas en el día j :

$$(R_t)_j = R_0 \cdot \exp(-\sum_i \alpha_i)$$

El modelo creado con este trabajo sigue, en lo fundamental, el mismo planteamiento del estudio inicial [7], pero con la posibilidad de tener en cuenta un número potencialmente ilimitado de intervenciones, evaluando el efecto de cada una de ellas para cada país por separado.

El siguiente gráfico, basado en el publicado por *Flaxman et col.* [7], resume los distintos componentes del modelo.

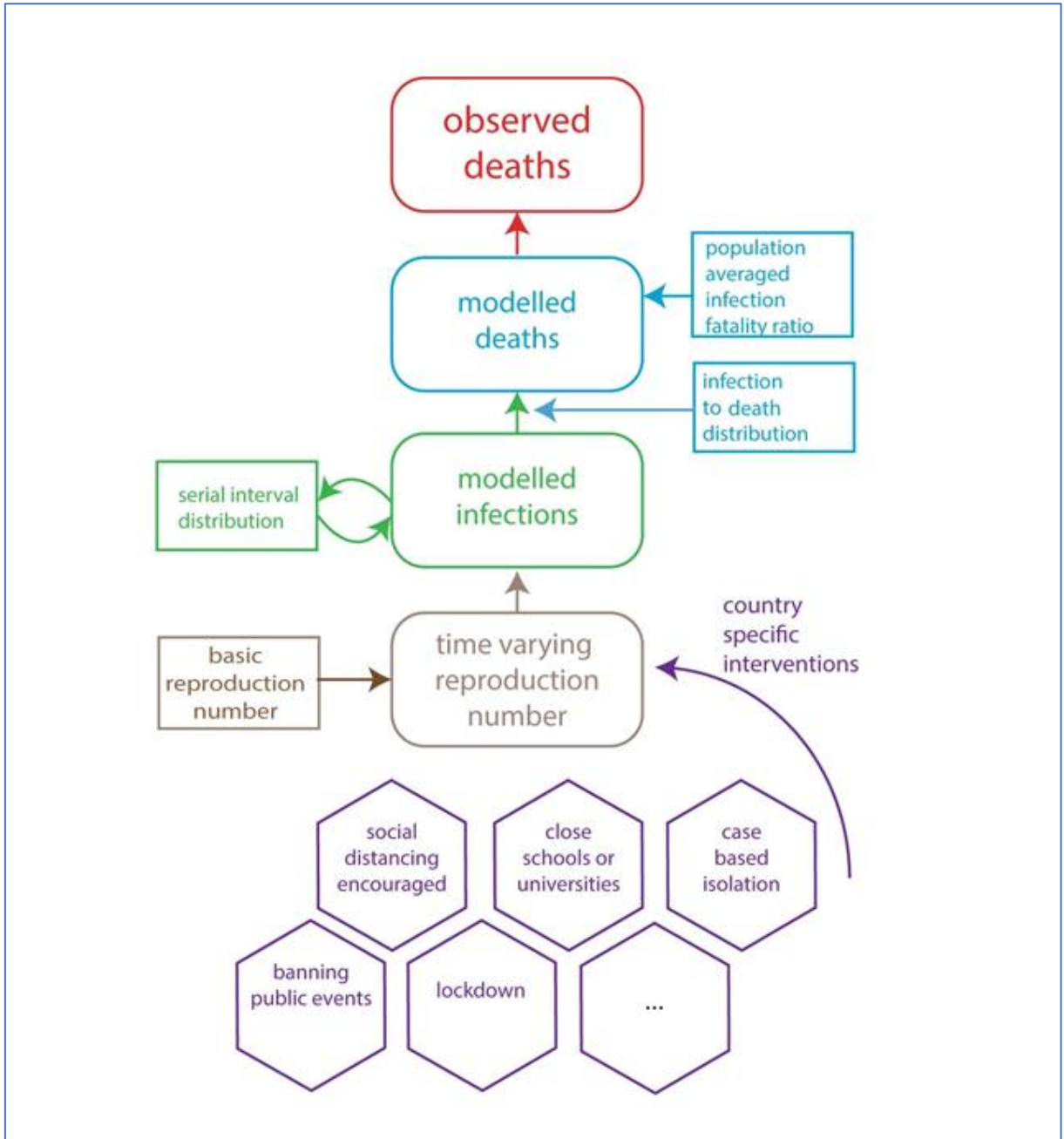


Figura 2 - Componentes del modelo mecanístico

4.4 Metodología - Análisis bayesiano

El apartado anterior describe la base matemática del modelo mecanístico que se utiliza en este trabajo para describir la evolución del número de infecciones y muertes por COVID-19.

Con este modelo como base, se utiliza un proceso de análisis de datos bayesiano para obtener estimaciones de los parámetros. Para ello, se siguen los pasos que se describen en el libro [Doing Bayesian Data Analysis \[8\]](#), y que aquí se formulan de forma simplificada:

1. *Identificación de los datos relevantes para el análisis, de las variables sobre las que han de hacerse predicciones y de las que se utilizan como predictores*

Ver “4.3 Fundamento estadístico - Modelo mecanístico” y “4.5 Datos necesarios para el análisis”

2. *Definición de un modelo descriptivo con los datos relevantes*

Ver “4.3 Fundamento estadístico - Modelo mecanístico”

3. *Definición de una distribución inicial (prior) de los parámetros*

Los parámetros que deben predecirse son, entre otros, los que definen el efecto de las diferentes medidas de intervención sobre el valor de R_t (*número de reproducción*) y el valor inicial de esta variable, definido como R_0 (*número de reproducción básico*) en el apartado anterior.

Como distribución inicial (*prior*) para los distintos parámetros se define la utilizada en el estudio que se toma como base para este trabajo [6].

Hay un parámetro adicional, no incluido en el modelo original, necesario para medidas cuyo efecto sobre el valor de R_t puede ser negativo (es decir, que pueden hacer que éste aumente). La distribución inicial (*prior*) de este parámetro se define como una normal con media 0 y desviación estándar 1.

4. *Utilización de la inferencia bayesiana para obtener una distribución posterior de los parámetros*

Las estimaciones de los parámetros (*distribución posterior*) se obtienen a través de un *muestreo de Monte Carlo Hamiltoniano* [8], [23]

5. *Comprobar la bondad del ajuste a los datos reales de las predicciones basadas en la distribución posterior de los parámetros*

Esto se hace únicamente de forma visual a través de los gráficos que se obtienen como resultado del análisis, comparando la curva de las predicciones del modelo para el número de muertes con los datos reales.

Como resultado de cada ajuste se obtiene, entre otros, un gráfico que representa la franja que incluye el 95% de los datos de todas las predicciones para el número de muertes, calculadas a partir de las estimaciones (*distribución posterior*) para los distintos parámetros. Cuanto más estrecha sea esta franja y mejor se adapte a los datos reales, mejor es el ajuste obtenido con el modelo.

Inferencia bayesiana

De forma muy simplificada, se puede decir que los métodos de inferencia bayesiana, como el que se utiliza en este trabajo, permiten sacar conclusiones, con la distribución *posterior* que se obtiene para ciertos parámetros, a partir de conocimientos o supuestos previos (*prior*) y de evidencias o datos conocidos.

Ésta es una explicación minimalista y no muy estricta, pero suficiente para explicar la idea detrás del modelo que se desarrolla con este trabajo. Para una explicación más fundamentada y rigurosa se pueden consultar, por ejemplo, los libros de [Kruschke \[8\]](#) y de [Gelman et col. \[23\]](#).

Trasladada al contexto de este trabajo, la inferencia bayesiana se aplica siguiendo los pasos descritos en la página anterior, para los que se dispone de la información necesaria:

- Se ha definido un modelo que describe la evolución de los casos y las muertes por *COVID-19* en función de ciertos parámetros
- Se asume una cierta distribución inicial (*prior*) de los valores de estos parámetros. En este caso, se utiliza el conocimiento previo disponible en estudios anteriores, como son los que sirven de referencia para este trabajo [6], [7]
- Se dispone de datos de las muertes por *COVID-19* [24] y de los periodos de aplicación de las medidas de intervención [22]
- Con la información anterior, el proceso de inferencia consiste en calcular o estimar las distribuciones *posteriores* de los parámetros a optimizar que hagan que las predicciones obtenidas con el modelo propuesto se ajusten lo mejor posible a los datos disponibles. Éste es el paso 4 de la página anterior.

El resultado de la inferencia es, por tanto, una distribución de valores para cada uno de los parámetros a optimizar. Estas distribuciones describen, para cada parámetro, la probabilidad de que éste tenga ciertos valores teniendo en cuenta la evidencia de los datos conocidos.

Lo que se hace es, por tanto, el cálculo o la estimación de una probabilidad condicionada a los datos conocidos. Matemáticamente, la forma de plantear el cálculo de este tipo de probabilidades se basa en la aplicación del teorema de *Bayes*, que se formula como sigue (ver detalles, por ejemplo, en las referencias bibliográficas ya mencionadas [8], [23]):

$$P(A|B) = \frac{P(B|A) P(A)}{P(B)}$$

En este trabajo, lo que se busca es la distribución de valores para cada parámetro condicionada a la evidencia de los datos conocidos ($P(A|B)$ en la fórmula). Ésta se obtendría, utilizando el teorema de *Bayes*, a partir de las probabilidades *a priori*, para las que se tiene una distribución que se asume válida. Éstas se corresponden con $P(A)$ y $P(B)$ en la fórmula anterior y, a partir de ellas, se asume que se podría obtener $P(B|A)$.

En el caso de distribuciones de valores discretos, la probabilidad del denominador de la fórmula se obtendría como la suma de las probabilidades marginales, mientras que para valores continuos sería necesaria la resolución de una integral.

En general, con datos y modelos medianamente complejos, la resolución exacta de las integrales puede no ser posible por su complejidad. Para estos casos, se utilizan métodos de simulación, que consisten en generar conjuntos de valores para los distintos parámetros, hasta encontrar combinaciones que permitan un buen ajuste de los datos conocidos. Uno de estos métodos es el **muestreo de Monte Carlo Hamiltoniano** [8], [23], que es el que se utiliza en este trabajo y que se aplicará haciendo uso de su implementación en el lenguaje probabilístico *Stan* [25].

4.5 Datos necesarios para el análisis

La siguiente tabla documenta los datos necesarios para el análisis y las fuentes de los mismos:

Datos	Fuente
<p><i>Time series summary</i></p> <p>Número de infecciones confirmadas y de muertes diarias por COVID-19.</p>	<p><i>Johns Hopkins University [24]</i></p>
<p><i>Data on country response measures to COVID-19</i></p> <p>Información sobre las medidas de intervención aplicadas en los distintos países europeos.</p>	<p><i>European Centre for Disease Prevention and Control [22]</i></p>
<p><i>Serial Interval (SI)</i></p> <p>Distribución que representa la probabilidad del tiempo transcurrido desde que una persona se infecta hasta que contagia la enfermedad a otra persona.</p>	<p><i>Fernández-Recio [6]</i></p>
<p><i>Infection-To-Death (ITD)</i></p> <p>Distribución que representa la probabilidad del tiempo transcurrido desde la infección hasta el fallecimiento.</p>	<p><i>Fernández-Recio [6]</i></p>
<p><i>Infection Fatality Ratio (IFR)</i></p> <p>Probabilidad media de fallecimiento para un individuo infectado.</p> <p>Se trata de datos ajustados a la estructura de edades de cada país. Para países, para los que esta información no está disponible, se utiliza la media de los valores conocidos.</p>	<p><i>Fernández-Recio [6]</i></p>

Tabla 1 - Datos necesarios para el análisis

Casos confirmados y número de muertes por COVID-19

En la [Johns Hopkins University \[24\]](#) están disponibles los datos diarios acumulados de casos confirmados y de muertes por COVID-19 para cada país. La información proviene de distintas fuentes, desde instituciones internacionales a fuentes específicas de diferentes países. Los detalles se pueden encontrar en la documentación del repositorio [24].

Una dificultad del uso de estos datos es el procedimiento que se utiliza para hacer correcciones en caso de que se detecten errores en fechas pasadas. Por ejemplo, en el caso de España, hay un día en junio de 2021 con un número de muertes anómalamente alto, resultado de una corrección. Con excepción de algún dato llamativo, que se ha observado que cambia significativamente el sentido de los resultados, en este trabajo no se ha invertido esfuerzo en la identificación y tratamiento de datos anómalos (*outliers*).

Los valores para el número de muertes son los que se utilizan para evaluar la bondad del ajuste de los modelos generados con la herramienta informática. Para ello, se comparan los datos reales con las predicciones de los modelos.

Por la definición del modelo estadístico (“4.3 Fundamento estadístico - Modelo mecanístico”), para obtener el número de muertes es necesario estimar el número de infectados. Este valor no es comparable con el número de casos confirmados, pero la diferencia entre ambos sirve como indicación de la capacidad de detección de la enfermedad.

Medidas de intervención

La información sobre las medidas de intervención aplicadas en distintos países europeos está disponible en el [European Centre for Disease Prevention and Control \[22\]](#) y proviene de fuentes oficiales. Hay información para 32 países en un fichero que se actualiza regularmente.

Para cada país, se documentan los periodos de aplicación de las distintas medidas. Cada periodo está identificado por el nombre de la medida y las fechas inicial y final de su aplicación.

Ejemplo: “Spain”, “StayHomeOrder”, 2020-03-14, 2020-05-03

Es decir, en España hubo una medida llamada *StayHomeOrder* que estuvo activa desde el 14 de marzo al 3 de mayo de 2020.

Los nombres de las medidas individuales están documentados en fuente de los datos ([Data dictionary non-pharmaceutical measures \[22\]](#)), y se resumen en uno de los anexos de este documento (“10.1 Anexo I - Nombres de las medidas de intervención”).

El formato original puede extenderse para activar opciones adicionales para intervenciones individuales con las que se modifica la forma en la que se procesarán durante el análisis. Los detalles pueden verse en la [documentación de la herramienta informática \[9\]](#).

Distribuciones (SI, ITD) y probabilidad de fallecimiento (IFR)

Los datos de las distribuciones *Serial Interval (SI)* e *Infection-To-Death (ITD)* y los valores para la probabilidad media de fallecimiento de un infectado (*Infection Fatality Ratio (IFR)*) se han obtenido del estudio de *Fernández-Recio [6]*, y se representan en los siguientes gráficos:

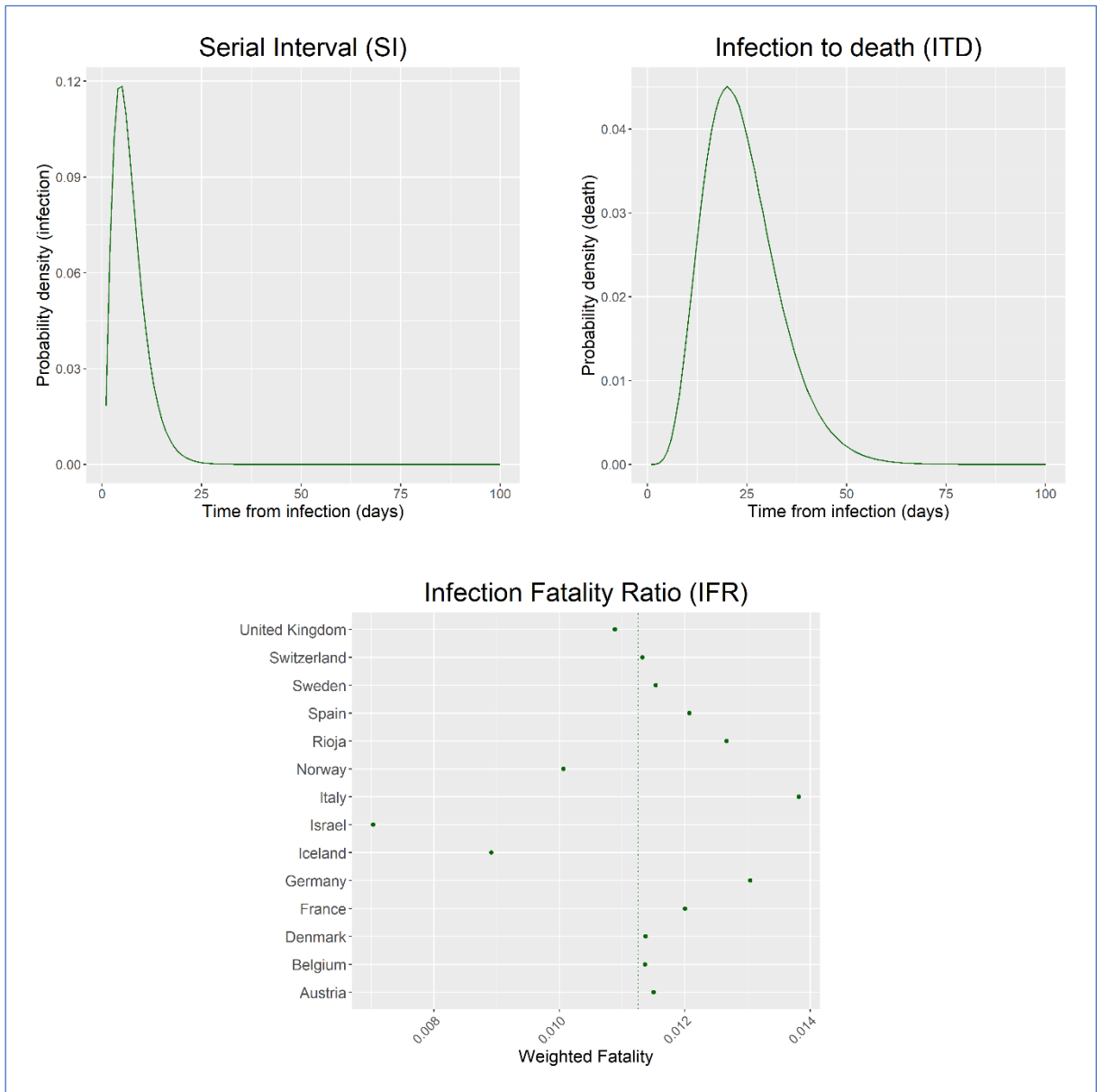


Figura 3 - Serial Interval (SI), Infection-To-Death (ITD), Infection Fatality Ratio (IFR)

4.6 Funcionalidad

Una descripción detallada de la herramienta informática está disponible en el servidor *web* [9] mencionado anteriormente. Aquí se describen de forma resumida las funciones principales, y el capítulo “10.6 Anexo VI - Herramienta informática - aspectos técnicos” contiene algunos detalles técnicos.

El gráfico de la página siguiente representa gráficamente los distintos componentes de la herramienta y las interacciones entre ellos para, a continuación, mostrar ejemplos de resultados obtenidos y describir brevemente cómo interpretarlos. Las funciones principales son:

1. Creación de modelos estadísticos que se ajusten lo mejor posible a los datos existentes

Para la creación de los modelos se utiliza el *script covid19_base.r* con el que, a partir de la información sobre casos y muertes por COVID-19 [24] y sobre las medidas de intervención [22], se estima el efecto de las distintas medidas, de forma que las predicciones del modelo para el número de muertes se ajusten lo mejor posible a los datos reales.

Las estimaciones se obtienen como resultado de un *muestreo de Monte Carlo Hamiltoniano* [8], [23] utilizando el lenguaje probabilístico *Stan* [25].

Como resultado se obtienen

- Un fichero *Rdata* con los resultados obtenidos, que pueden utilizarse para el cálculo de predicciones
- El resumen de los pasos ejecutados y de resultados intermedios (*log*)
- Gráficos (*pdf, png*) que representan las predicciones del modelo para el número de infecciones y muertes por COVID-19 y para el valor de R_t (*número de reproducción*)
- Ficheros de texto (*txt*) con contenido en formato *csv* con las estimaciones de los parámetros que cuantifican el efecto de las distintas medidas de intervención

2. Predicciones para periodos futuros bajo distintos supuestos

Las predicciones sobre el número de casos y de muertes para periodos futuros se obtienen con el *script covid19_predict.r*, que utiliza los ficheros *Rdata* de modelos creados previamente e información sobre medidas de intervención futuras.

Como resultado se obtienen

- Un resumen de los pasos ejecutados y de resultados intermedios (*log*)
- Gráficos que extienden los del modelo inicial para incluir las predicciones sobre periodos futuros

3. Comparación del efecto de las distintas medidas

El *script covid19_compare_measure_effect.r* procesa los resultados numéricos sobre el efecto de las medidas de intervención para los distintos países obtenidos de la creación de modelos.

El resultado son gráficos que permiten comparar el efecto de las distintas medidas en cada uno de los países analizados así como hacer comparaciones entre los diferentes países.

4.6.1 Herramienta informática - Componentes

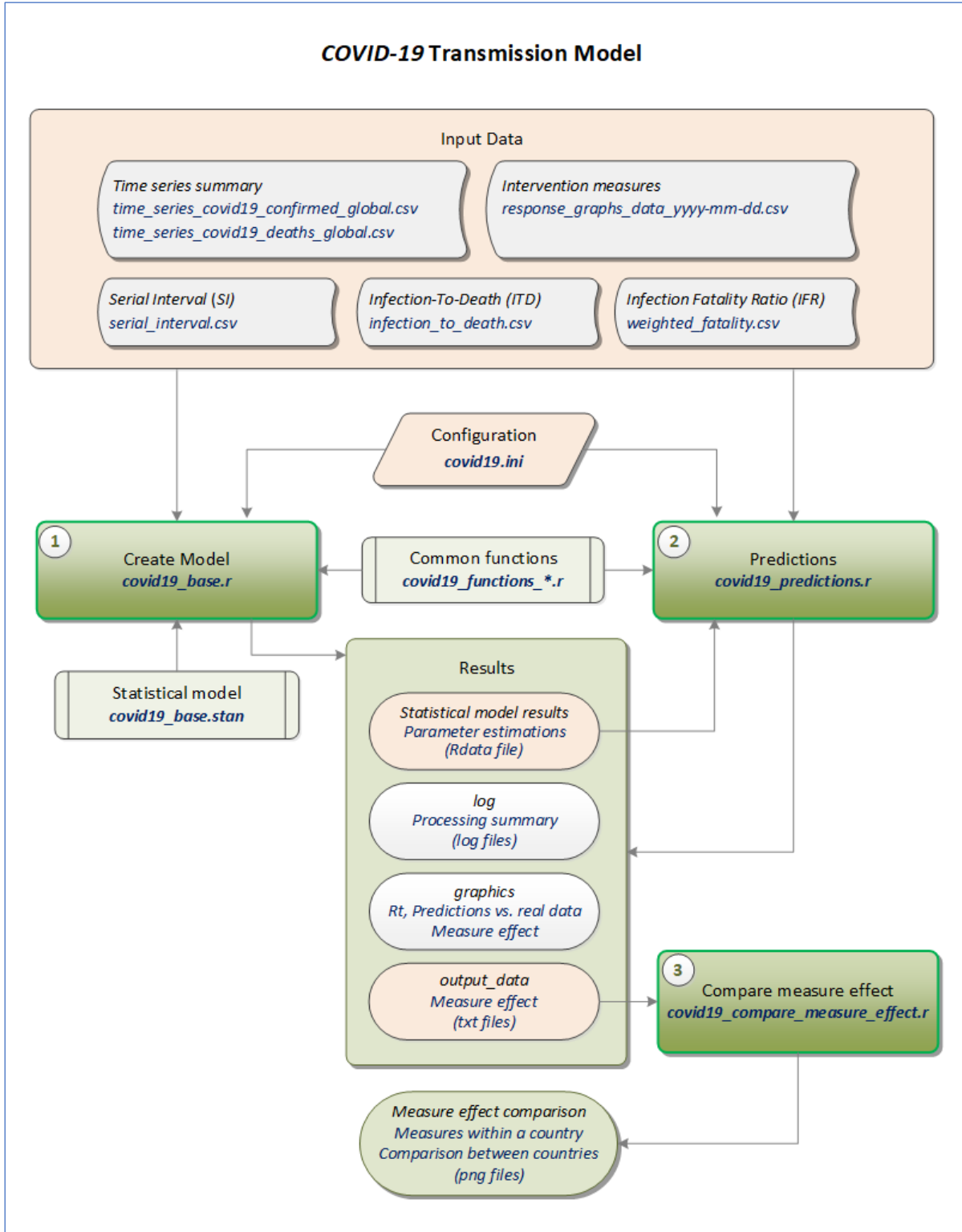


Figura 4 - Herramienta informática - Componentes

4.6.2 Modelos y Predicciones - Bondad del ajuste y efecto de las distintas medidas

Los siguientes gráficos son ejemplos de los resultados que se obtienen con la ejecución de la herramienta informática. Su contenido se explica en las siguientes páginas.

Se trata de gráficos para predicciones obtenidas a partir de un modelo creado previamente para el Reino Unido con datos desde el 22.01.2020 al 18.04.2021.

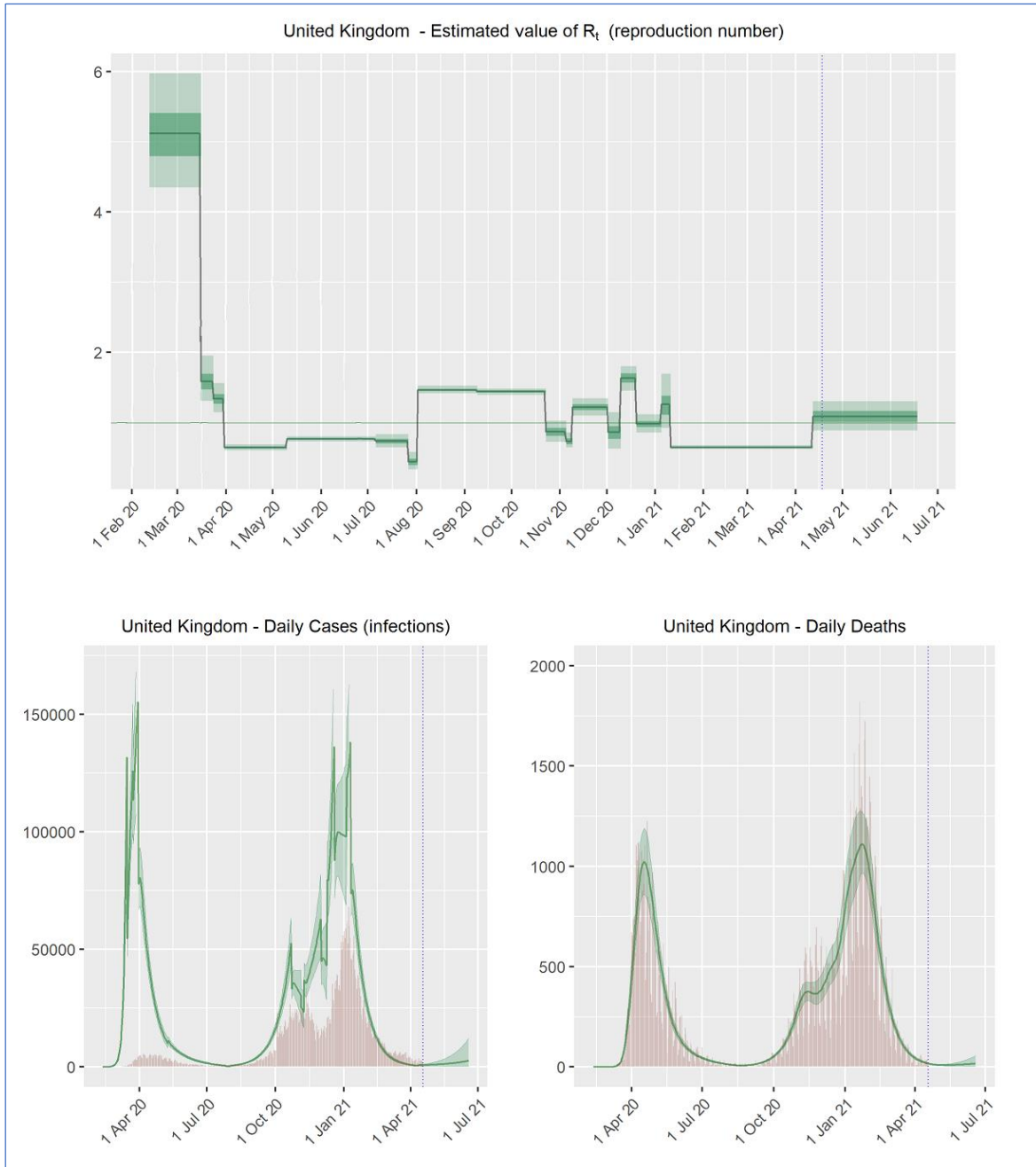


Figura 5 - Ejemplo - Predicciones para R_t y para el número de infectados y muertos por COVID-19

En los gráficos de la página anterior, la línea vertical azul marca la fecha hasta la que hay datos. Hasta esta fecha, los valores son los obtenidos durante la creación del modelo, mientras que a partir de esa fecha se trata de predicciones futuras (*forecast*).

El espacio delimitado por una franja verde contiene el rango que incluye el 95% de los datos de las estimaciones. Cuanto más estrecha sea esta franja, se puede asumir que hay menos incertidumbre en las estimaciones/predicciones.

Estimated value of R_t (reproduction number)

Estimación del valor de R_t .

Cada cambio en el valor de este parámetro está asociado a la activación o desactivación de alguna medida de intervención. El efecto de la activación de cada medida queda reflejado en su contribución a la reducción de R_t a través del factor $\exp(-\alpha_i)$, tal y como se define en el modelo mecanístico (ver “4.3 Fundamento estadístico - Modelo mecanístico”).

Daily Cases (infections)

La franja verde representa las predicciones del modelo para el número de infecciones por COVID-19, mientras que el histograma representa los casos documentados.

Asumiendo que las predicciones son buenas, lo que se puede corroborar con el siguiente gráfico, una mayor distancia entre la franja verde y las barras verticales indica un menor porcentaje de detección de los casos de COVID-19.

Daily Deaths

La franja verde representa las predicciones del modelo para el número de muertes, mientras que el histograma representa las muertes reales.

Cuanto más se aproxime la franja verde a la silueta definida por las barras verticales, mejor es el ajuste.

Los gráficos de la página anterior permiten extraer las siguientes conclusiones:

- El ajuste para el número de muertes es muy bueno: la franja verde se adapta muy bien a los datos reales y es bastante estrecha, lo que indica que no hay mucha incertidumbre en las predicciones.
- El buen ajuste del número de muertes permite asumir que la estimación del número de infecciones también será buena.
Comparando las predicciones con los casos documentados, se puede concluir que, durante la primera ola de la pandemia, la proporción de casos detectados era muy limitada, mientras que, con el paso del tiempo, ésta ha mejorado significativamente.

Effect of measures on R_t

Con la creación de los modelos se obtienen también las estimaciones del efecto de las distintas medidas (factor $\exp(-\alpha_i)$), tanto de forma gráfica como de forma numérica:

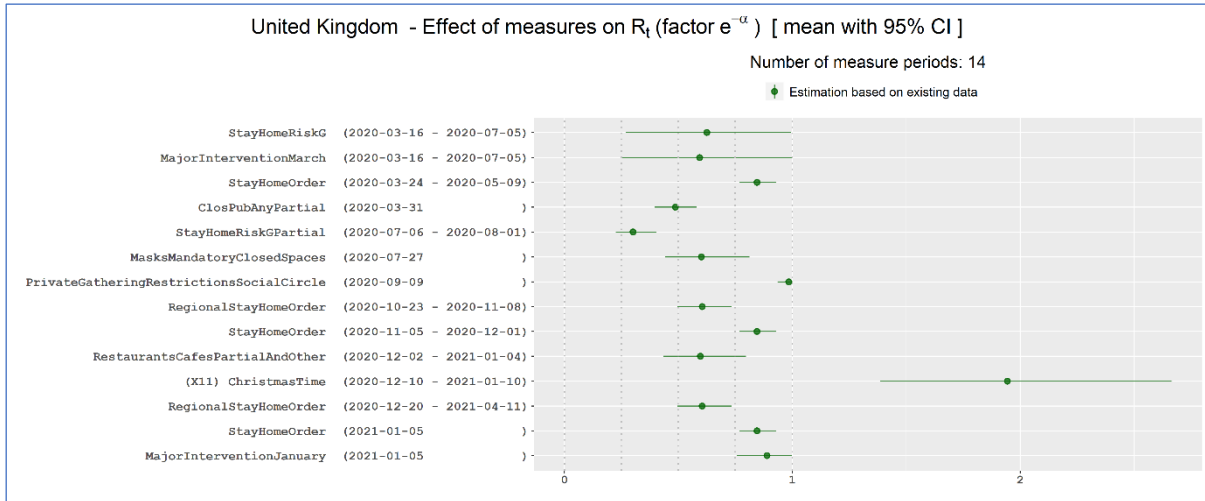


Figura 6 - Ejemplo - Estimación del efecto de las medidas de intervención en un país (1)

Lo que se representa es el valor medio de $\exp(-\alpha_i)$ con el rango de valores que incluye el 95% de todas las predicciones. Por simplicidad, este rango se describe en el título del gráfico como "95% CI", si bien, estrictamente, no se trata de un intervalo de confianza.

Valores menores de 1 suponen una reducción de R_t y, por tanto, representan medidas efectivas contra la pandemia, mientras que valores mayores de 1 indican eventos que han contribuido a un aumento del valor de R_t y han supuesto, por tanto, un incremento en el número de contagios (en este ejemplo, el periodo navideño: *ChristmasTime*).

Los valores numéricos de las estimaciones del efecto de las distintas medidas también se documentan en un fichero que se genera junto con el gráfico, y del que aquí se muestra un extracto:

Country	Response measure	date_start	date_end	inverse_exp_alpha_mean
United Kingdom	MajorInterventionMarch	16/03/2020	05/07/2020	0.593709928
United Kingdom	StayHomeRiskG	16/03/2020	05/07/2020	0.625884142
United Kingdom	StayHomeOrder	24/03/2020	09/05/2020	0.844772191
United Kingdom	ClosPubAnyPartial	31/03/2020	NA	0.486889241
United Kingdom	StayHomeRiskGPartial	06/07/2020	01/08/2020	0.301111488
United Kingdom	MasksMandatoryClosedSpaces	27/07/2020	NA	0.600864135
United Kingdom	PrivateGatheringRestrictionsSocialCircle	09/09/2020	NA	0.984846793
United Kingdom	RegionalStayHomeOrder	23/10/2020	08/11/2020	0.604421782
United Kingdom	StayHomeOrder	05/11/2020	01/12/2020	0.844772191
United Kingdom	RestaurantsCafesPartialAndOther	02/12/2020	04/01/2021	0.59735299
United Kingdom	(X11) ChristmasTime	10/12/2020	10/01/2021	1.943573129
United Kingdom	RegionalStayHomeOrder	20/12/2020	11/04/2021	0.604421782
United Kingdom	MajorInterventionJanuary	05/01/2021	NA	0.889117633
United Kingdom	StayHomeOrder	05/01/2021	NA	0.844772191

Tabla 2 - Estimación numérica del efecto de las medidas de intervención en un país

4.6.3 Comparación del efecto de las distintas medidas en diferentes países

Con modelos generados para varios países, se dispone de información que permite la comparación del efecto de las distintas medidas de intervención.

Por un lado, se obtienen gráficos que representan de otra forma la información de los diagramas como el de la página anterior, es decir, el efecto de las distintas medidas en la evolución de la enfermedad en **cada país por separado**.

En estos gráficos, se representa el valor de $\exp(-\alpha_i)$, al igual que en el gráfico del apartado anterior. El formato es distinto, de forma que se pueda identificar de forma visual más fácilmente la efectividad de las distintas medidas. Las medidas más efectivas son las que tienen valores menores de $\exp(-\alpha_i)$, y se representan con barras hacia la izquierda, cuya mayor longitud indica una mayor contribución a la disminución de R_t y, por tanto, una mayor efectividad. Eventos o periodos con un impacto negativo en la evolución de la enfermedad tienen valores mayores de 1 y se representan con una barra roja hacia la derecha de este valor, indicando su contribución al aumento de R_t .

Medidas similares se agrupan para hacer los resultados más compactos. Por ejemplo, todas las medidas relacionadas con el confinamiento (*StayHomeOrder*, *StayHomeGen*, *StayHomeRiskG*, ...) se agrupan bajo el nombre genérico *StayHome*, al que se le asigna el mejor valor (el menor) entre todas las medidas que incluye.

Ejemplo:

Las medidas más efectivas en el Reino Unido fueron las relacionadas con el confinamiento (*StayHome*) y con el cierre de espacios públicos (*ClosPubAny*), mientras que el periodo navideño (*ChristmasTime*) supuso un aumento en el valor de R_t que tendría como consecuencia un aumento del número de casos.

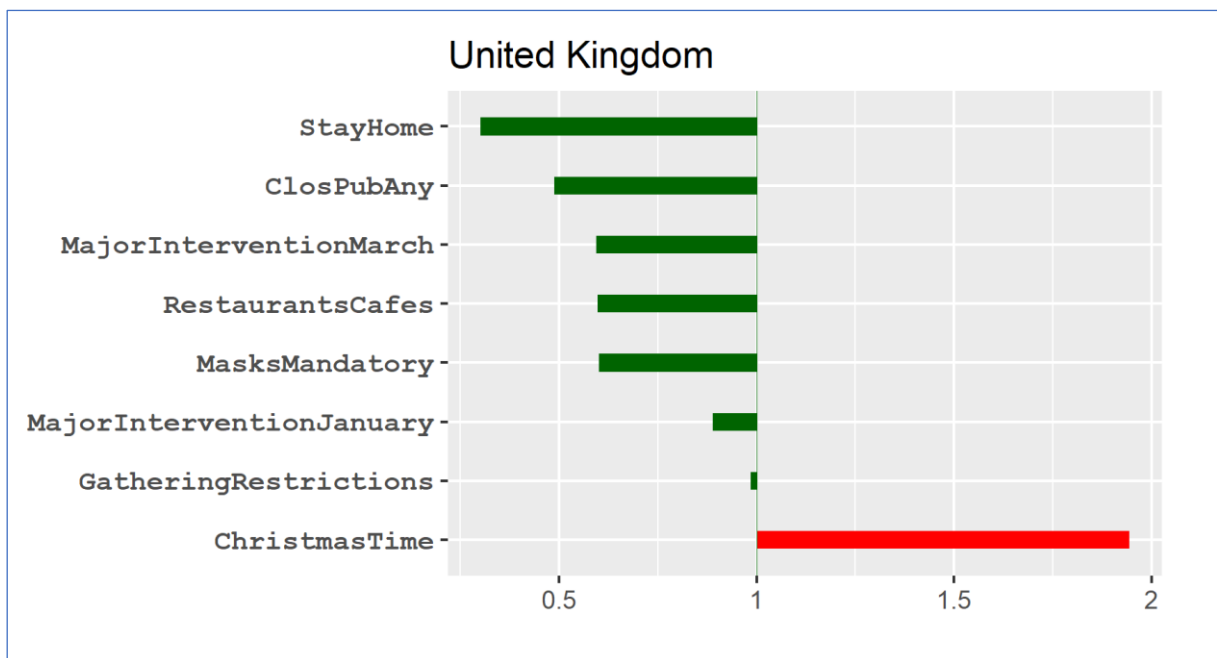


Figura 7 - Ejemplo - Estimación del efecto de las medidas de intervención en un país (2)

Por otro lado, se generan gráficos para **cada una de las medidas de intervención por separado** que documentan su efecto en cada país. Esto permite comparar la efectividad de medidas individuales entre distintos países. Se representa la media de las estimaciones para $\exp(-\alpha_i)$ junto con el rango que incluye el 95% de las predicciones.

Ejemplo:

En el periodo analizado (22.01.2020 - 18.04.2021), se ha estimado que en Polonia, Bélgica, y España fue donde más efectividad tuvieron las medidas relacionadas con el confinamiento (*StayHome*).

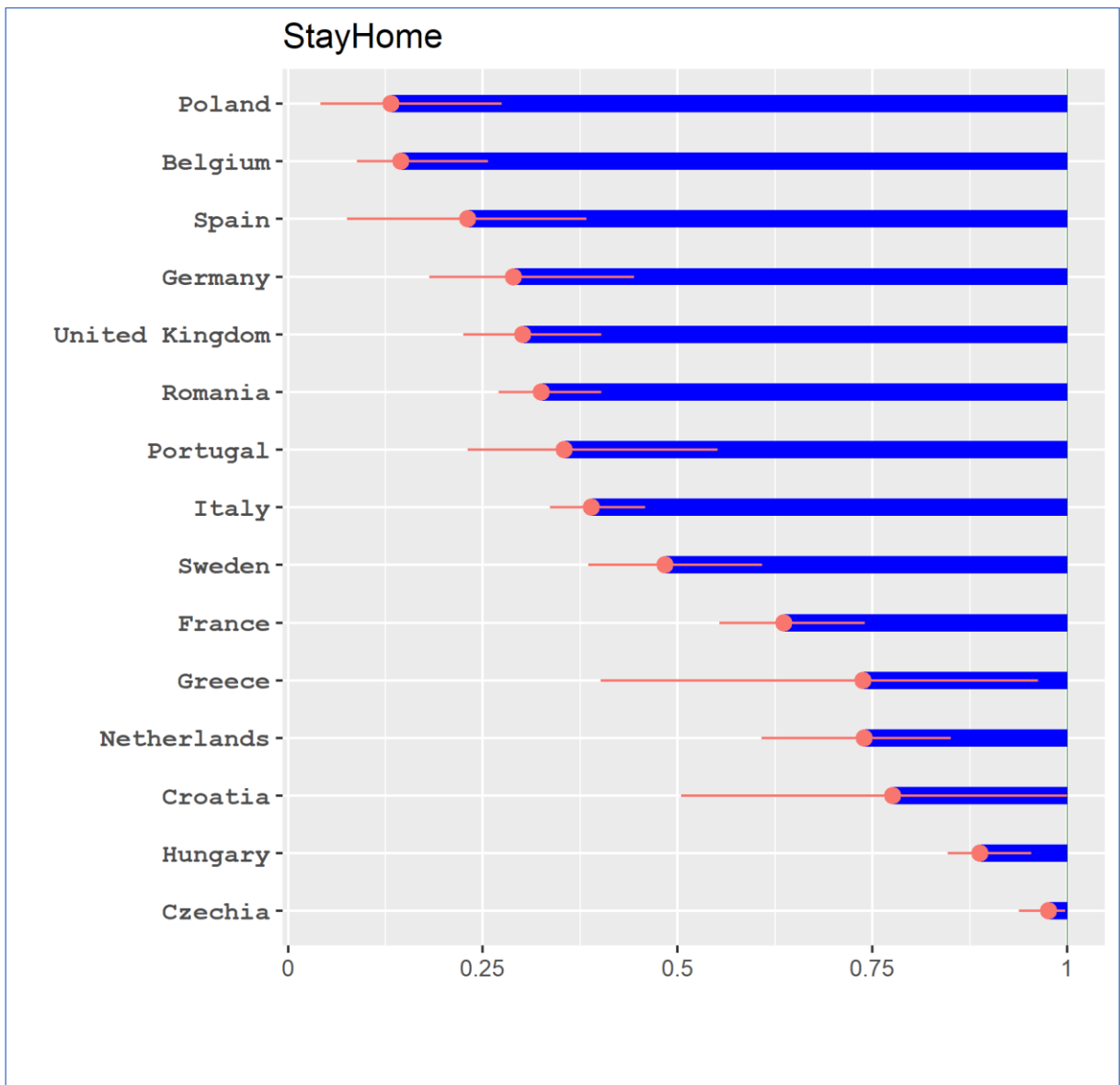


Figura 8 - Ejemplo - Efectividad de las medidas de confinamiento en distintos países

5 Resultados

Este capítulo documenta los resultados del análisis de datos de *COVID-19* de distintos países utilizando la herramienta informática con la que se han desarrollado los aspectos metodológicos de este trabajo, descritos en el capítulo anterior.

Los siguientes apartados describen los resultados de los distintos análisis, ejecutados con los siguientes objetivos:

- Evaluar la bondad del modelo estadístico
- Obtener información sobre el efecto de las distintas medidas en diferentes países
- Hacer predicciones sobre la evolución del *COVID-19* en varios países bajo distintos supuestos

Como se ha explicado en “4.6.2 Modelos y Predicciones - Bondad del ajuste y efecto de las distintas medidas”, la **bondad del ajuste** se evalúa comparando las predicciones para el número de muertes con el número de muertes reales, información disponible en uno de los gráficos que se generan con la creación de los modelos (*Daily Deaths*).

La evaluación de la bondad del ajuste de los modelos resultantes se hace exclusivamente de forma gráfica, comparando visualmente los resultados del ajuste de cada modelo con los datos reales. Si bien se generan datos que permitirían una evaluación numérica de la bondad de los ajustes, por ejemplo, utilizando el *RMSE (root mean square error)*, se ha prescindido de ello.

Se documentan también las predicciones para el *número de reproducción básico (R_0)* y se comparan los valores obtenidos con los de otros estudios.

La información sobre el **efecto de las medidas de intervención** se presenta principalmente de forma gráfica, aunque también están disponibles los valores numéricos, como se ha descrito en “4.6.2 Modelos y Predicciones - Bondad del ajuste y efecto de las distintas medidas”.

Dada la gran cantidad de datos producidos como resultado de los distintos análisis, este documento sólo contiene algunos ejemplos de los mismos, que se incluyen donde se ha considerado que son necesarios para hacer las explicaciones más comprensibles.

Se puede acceder a todos los resultados a través del servidor *web* [9] mencionado en apartados anteriores: <http://jblecua.spdns.org:9095/> (*usuario: tfm / password: tfm*).

5.1 Combinaciones de datos utilizados para el análisis

La parte central del proceso de análisis ha consistido en la evaluación del modelo implementado con la herramienta informática en cuanto a su capacidad de ajuste para datos que siguen patrones distintos. Para ello, se han utilizado distintas combinaciones de los datos disponibles:










	Datos de COVID-19		Intervenciones	Número de modelos resultantes	Resultados
	Países / Regiones	Rango de fechas			
1	Países europeos	22.01.2020 - 18.04.2021	Reducidas	32	
2	Países europeos	22.01.2020 - 18.02.2021	Reducidas	32	
3	Países europeos	22.01.2020 - 31.07.2020	Reducidas	32	
4	Países europeos	01.08.2020 - 18.04.2021	Reducidas	32	
5	Países europeos	22.01.2020 - 18.02.2021	Todas	32	
6	Países europeos	22.01.2020 - 31.07.2020	Todas	32	
7	Países europeos	01.08.2020 - 18.04.2021	Todas	32	
8	Regiones españolas	14.05.2020 - 09.04.2021	Reducidas	19	
9	Regiones alemanas	14.05.2020 - 09.04.2021	Reducidas	16	

Tabla 3 - Combinaciones de datos utilizados para el análisis

Se han utilizado datos de todos los países europeos para los que hay información sobre intervenciones no farmacológicas en el [European Centre for Disease Prevention and Control \[22\]](#). Se trata de un total de 32 países, entre los que se incluyen los 27 países de la Unión Europea y, además, Islandia, Liechtenstein, Noruega, Suiza y Reino Unido.

Como complemento, se han analizado también los datos de cada una de las regiones españolas y alemanas por separado. Los resultados se comentan brevemente en “10.2 Anexo II - Análisis de los datos a nivel regional para España y Alemania”.

Por el momento en el que se crearon, los estudios en los que se basa este trabajo [6], [7] utilizaban datos de un periodo relativamente reducido de tiempo. Con este trabajo, se ha creado una herramienta que funciona bien también para periodos más extensos, y algunos de los análisis incluyen todo el rango de datos disponibles, desde el inicio de la documentación de información sobre la pandemia (22.01.2020) hasta fechas recientes (18.04.2021), tal y como se describe en la tabla anterior.

El uso de intervenciones *Reducidas* en algunos de los análisis (en lugar de *Todas*) tiene que ver con la optimización del número de medidas documentadas en el [European Centre for Disease Prevention and Control \[22\]](#). La finalidad es *aislar* el efecto de medidas individuales, cuando esto es posible. Para ello, se eliminan o se agrupan algunas de las medidas originales siguiendo distintos criterios que se describen en “5.2 Agrupación y reducción de las medidas de intervención”.

5.2 Agrupación y reducción de las medidas de intervención

5.2.1 Motivación

En todos los países ha habido periodos en los que se han aplicado simultáneamente varias medidas de intervención, lo que dificulta la evaluación del efecto de medidas individuales.

Esto puede verse en el siguiente ejemplo, en el que se muestran los resultados del ajuste de los datos de Italia en el periodo del 22.01.2020 al 18.02.2021 en dos variantes: por un lado, utilizando todas las medidas de intervención documentadas y, por otro lado, utilizando sólo un conjunto reducido (optimizado) de medidas, resultantes de aplicar criterios que se detallan en el siguiente apartado.

De acuerdo con la información del [European Centre for Disease Prevention and Control \[22\]](#), entre esas dos fechas hubo en Italia un total de 60 periodos de intervención. La reducción y agrupación de medidas (ver "5.2.2 Criterios") permite hacer un buen ajuste de los datos con únicamente 8 intervenciones.

Los siguientes gráficos muestran el resultado del ajuste para las dos variantes. Se puede apreciar que los resultados son muy similares y que, por tanto, es posible describir la evolución de la enfermedad en el rango de fechas seleccionado en función de únicamente 8 periodos de intervención:

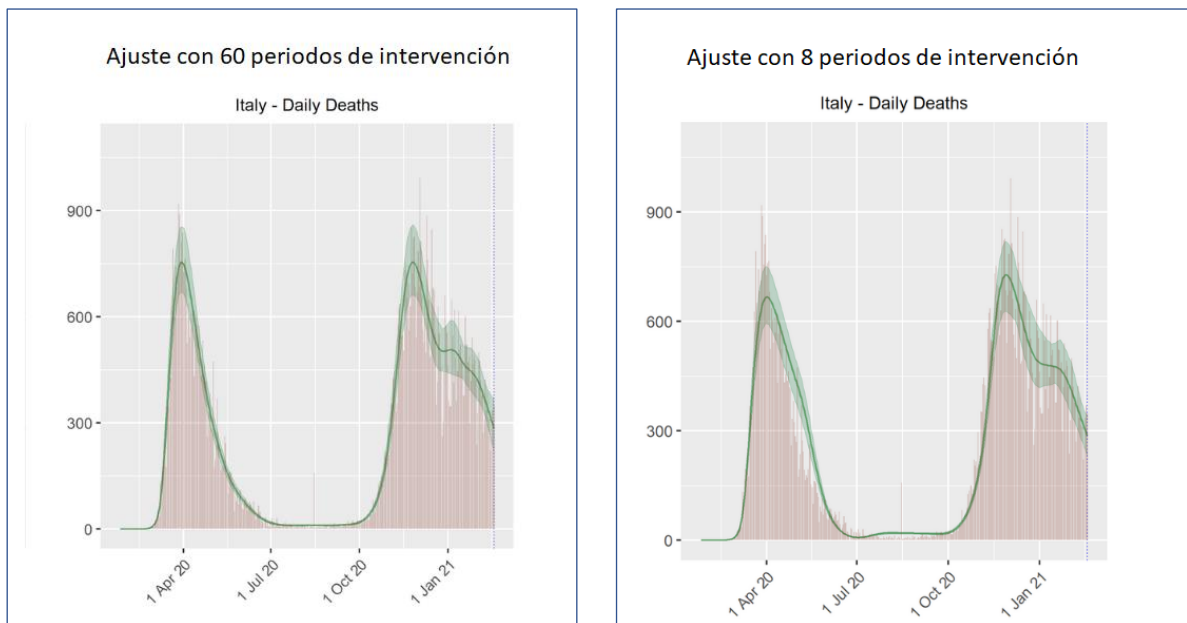


Figura 9 - Comparación del ajuste con distintos grupos de medidas de intervención

El **ajuste con las 60 intervenciones originales** muestra un alto grado de incertidumbre en las estimaciones del efecto de las distintas medidas, reflejado en la amplitud de los rangos de valores que representan el 95% de todas las predicciones:

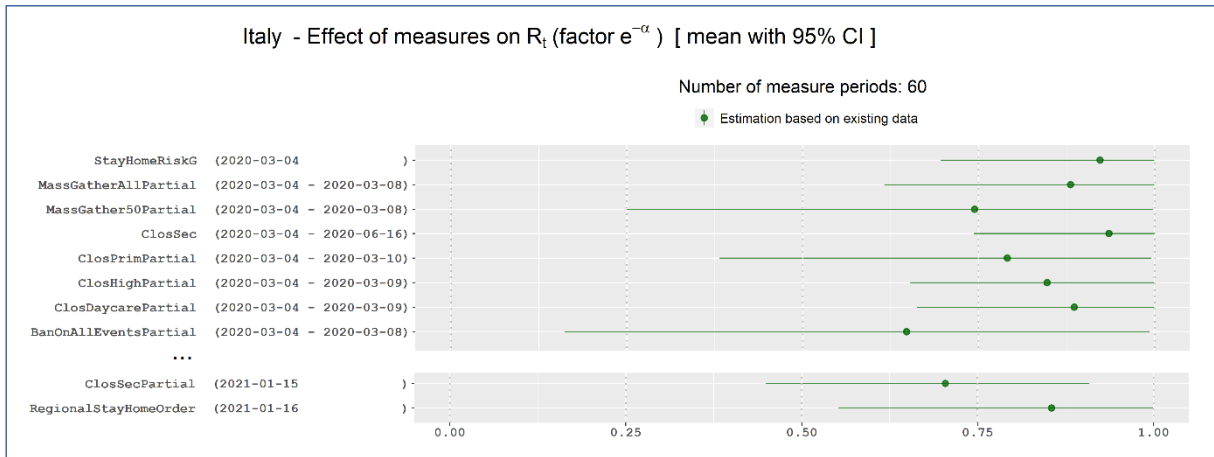


Figura 10 - Efecto de las medidas de intervención - estimaciones con todas las medidas

La **versión con únicamente 8 periodos de intervención** reduce la incertidumbre en los valores de las estimaciones:

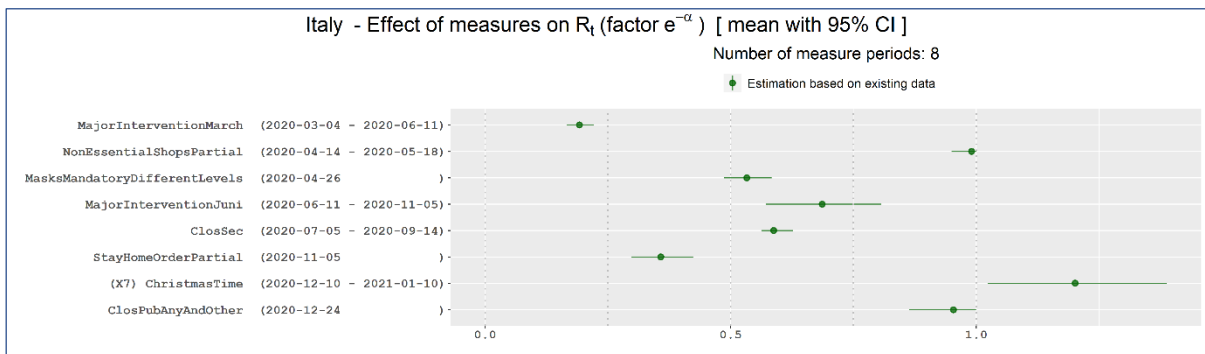


Figura 11 - Efecto de las medidas de intervención - estimaciones con las medidas optimizadas

El rango del 95% de las predicciones es, para todos los periodos, bastante reducido, lo que permite sacar conclusiones sobre el efecto de cada una de las medidas de intervención en el valor de R_t (número de reproducción) y, por tanto, en la evolución de la enfermedad:

- Las medidas que más contribuyeron a la reducción de R_t (número de reproducción) fueron la intervención de marzo de 2020 (*MajorInterventionMarch*) y el periodo de confinamiento parcial en mayo (*StayHomeOrderPartial*).
- Por otro lado, el periodo navideño supuso un aumento del valor de R_t (número de reproducción).

5.2.2 Criterios

Para obtener las medidas que reflejen el conjunto de intervenciones aplicadas en cada país, pero que al mismo tiempo permitan obtener estimaciones de su efecto con un grado aceptable de incertidumbre, se han reducido las medidas de intervención documentadas en el [European Centre for Disease Prevention and Control \[22\]](#) siguiendo distintos criterios.

El procedimiento ha sido prescindir de algunas medidas y agrupar otras en medidas genéricas, lo que hace que, en algunos casos, no sea posible obtener información sobre el efecto de cada una de las medidas individuales. Esto es razonable: si hay muchas medidas activas en un mismo periodo, es posible estimar el efecto asociado a todas las medidas en conjunto pero, sin información adicional, no es posible saber cuál es la contribución de cada una de las medidas individuales.

Criterios para la agrupación de medidas para cada uno de los países

Para la creación de los modelos con los conjuntos de medidas *Reducidas*, se han analizado las medidas de intervención definidas en cada país [22] y se han reorganizado de forma manual. Para ello, se han definido una serie de criterios concretos (ver abajo) que se han aplicado de forma idéntica a todos los países, para que así los resultados del análisis sean comparables en cuanto a las conclusiones sobre el procedimiento.

Los conjuntos de medidas resultantes son los que se han utilizado como parte del análisis para evaluar la bondad del modelo estadístico (“5.3 Evaluación del modelo - Bondad del ajuste”) y se corresponden a las combinaciones de datos que utilizan intervenciones *Reducidas*, tal y como se describen en “5.1 Combinaciones de datos utilizados para el análisis” .

La siguiente lista resume el proceso utilizado para la reducción (optimización) de las medidas:

1. Grupos de medidas activas simultáneamente se han agrupado en medidas genéricas denominadas *major interventions*.

Un ejemplo son los conjuntos de medidas aplicadas en marzo de 2020 en la mayoría de los países europeos, que se han agrupado bajo el nombre de *MajorInterventionMarch*.

2. Se han eliminado medidas que han estado activas un periodo muy corto de tiempo.

Hay casos en los que algunas medidas están activas un periodo demasiado breve como para que sean efectivas. Por ejemplo, en el caso de Irlanda se documenta una medida de confinamiento (*StayHomeGen*) que estuvo activa únicamente entre el 24.03.2020 y el 26.03.2020.

Se ha considerado que se puede prescindir de medidas activas menos de dos semanas si ha habido otras medidas activas en ese mismo periodo. La excepción son las medidas de confinamiento, a las que se ha dado *prioridad* por estimarse que su efecto será mayor que el de medidas menos restrictivas (ver siguiente punto).

3. En periodos con medidas de confinamiento, y con pocas medidas activas simultáneamente como para considerarlas *major interventions*, se han considerado para el análisis únicamente las medidas de confinamiento, prescindiendo de las demás. Este criterio se elige haciendo uso de resultados de otros estudios [6], [7], [12], [17], en los que se concluye el mayor efecto de las medidas de confinamiento sobre otras intervenciones.

4. Se han eliminado medidas redundantes activas en un mismo periodo.

De acuerdo con la documentación del [European Centre for Disease Prevention and Control \[22\]](#), la información sobre las medidas de intervención proviene de fuentes oficiales. A pesar de eso, hay datos que no son del todo consistentes y, en algunos casos, se documentan medidas redundantes.

Por ejemplo, para la República Checa se documenta la activación de las medidas *MassGatherAll* y *MassGather50* el 22.01.2020 por un periodo indefinido

Como se puede ver en “10.1 Anexo I - Nombres de las medidas de intervención”, la medida *MassGatherAll* consiste en la limitación de concentraciones de cualquier tamaño de hasta 1000 participantes.

Por otro lado, *MassGather50*, si bien no documentada explícitamente, significaría la limitación de concentraciones de hasta 50 personas, lo que ya estaría incluido en la medida *MassGatherAll*.

En este caso, se elimina la medida *MassGather50*.

5. Se han agrupado medidas similares en medidas conjuntas.

Por ejemplo, en Grecia se cerraron parte de las instituciones educativas entre el 11.03.2020 y el 01.06.2020. Este cierre se documentó con tres medidas distintas, una para cada tipo de institución: *ClosSec*, *ClosPrim*, *ClosDayCare*. Estas tres medidas se han agrupado para el análisis en una única, denominada *ClosSecPrimDayCare*.

Agrupación de medidas para la comparación de su efecto entre distintos países

Tras la creación de los modelos para cada país por separado, se han definido *categorías* que agrupan medidas similares, lo que posibilita las comparaciones entre los distintos países. Por ejemplo:

Mientras que en España ha habido periodos de confinamiento obligatorio (*StayHomeOrder*), en Alemania esta medida se ha aplicado sólo de forma parcial (*StayHomeOrderPartial*). Asumiendo que la comparación del efecto de estas dos medidas (y de medidas similares) tiene sentido, se agrupan todas las medidas relacionadas con el confinamiento bajo una categoría genérica denominada *StayHome*.

La siguiente tabla contiene la información sobre las categorías definidas. Las medidas que no aparecen en la tercera columna se utilizan con su nombre original para las comparaciones entre los distintos países:

Categoría	Descripción	Medidas incluidas en este grupo
<i>StayHome</i>	Medidas relacionadas con el confinamiento	<i>StayHomeOrder</i> <i>StayHomeGen</i> <i>StayHomeRiskG</i> <i>RegionalStayHomeOrder</i> <i>RegionalLockdown20</i>
<i>GatheringRestrictions</i>	Limitación de reuniones y contactos sociales	<i>BanOnAllEvents</i> <i>MassGatherAll</i> <i>PrivateGatheringRestrictions</i> <i>OutdoorOverNNN</i> <i>OutdoorOverNNN</i> <i>SocialCircle</i> <i>EntertainmentVenues</i>
<i>CloseEducation</i>	Cierre de centros educativos	<i>ClosDaycare</i> <i>ClosPrim</i> <i>ClosSec</i> <i>ClosHigh</i>
<i>MasksMandatory</i>	Uso obligatorio de mascarillas	<i>MasksMandatoryAllSpaces</i> <i>MasksMandatoryClosedSpaces</i>
<i>MasksVoluntary</i>	Uso voluntario de mascarillas	<i>MasksVoluntaryAllSpaces</i> <i>MasksVoluntaryClosedSpaces</i>
<i>XXX</i> <i>ClosPubAny</i> ...	Las medidas aplicadas de forma parcial se tratan como las medidas originales	<i>XXXPartial</i> <i>ClosPubAnyPartial</i> ...

Tabla 4 - Medidas de intervención - agrupación en categorías para la comparación entre países

5.3 Evaluación del modelo - Bondad del ajuste

El primero de los objetivos relacionados con el análisis de datos es la evaluación del ajuste de los modelos creados con distintos datos. Este capítulo resume los resultados obtenidos.

5.3.1 Criterios para la evaluación de la bondad del ajuste

Como se ha descrito en “5.1 Combinaciones de datos utilizados para el análisis”, se han creado un total de 259 modelos distintos para evaluar la bondad del ajuste de datos muy variados.

La parte principal del análisis ha consistido en la creación de modelos para un total de 32 países europeos en 7 variantes distintas (número total de modelos: 224).

Cada uno de los 224 modelos se ha clasificado en una de las categorías *Muy bueno*, *Bueno*, *Regular*, *Malo* en función de la bondad del ajuste para el número de muertes por COVID-19 (gráfico *Daily Deaths*).

La clasificación es subjetiva y basada exclusivamente en la evaluación visual de los gráficos, pero se ha hecho de forma conservadora para evitar resultados artificialmente optimistas.

Se han definido tres criterios para asignar cada modelo a uno u otro grupo:

- Tendencia de la franja de predicciones en comparación con la tendencia de los datos reales
- Proximidad de la curva del valor medio de las predicciones a los valores reales
- Nivel de incertidumbre, reflejado en la anchura de la franja de los valores de las predicciones

Estos dos gráficos son ejemplos de un ajuste *Muy bueno* y de otro *Malo* teniendo en cuenta los criterios mencionados:

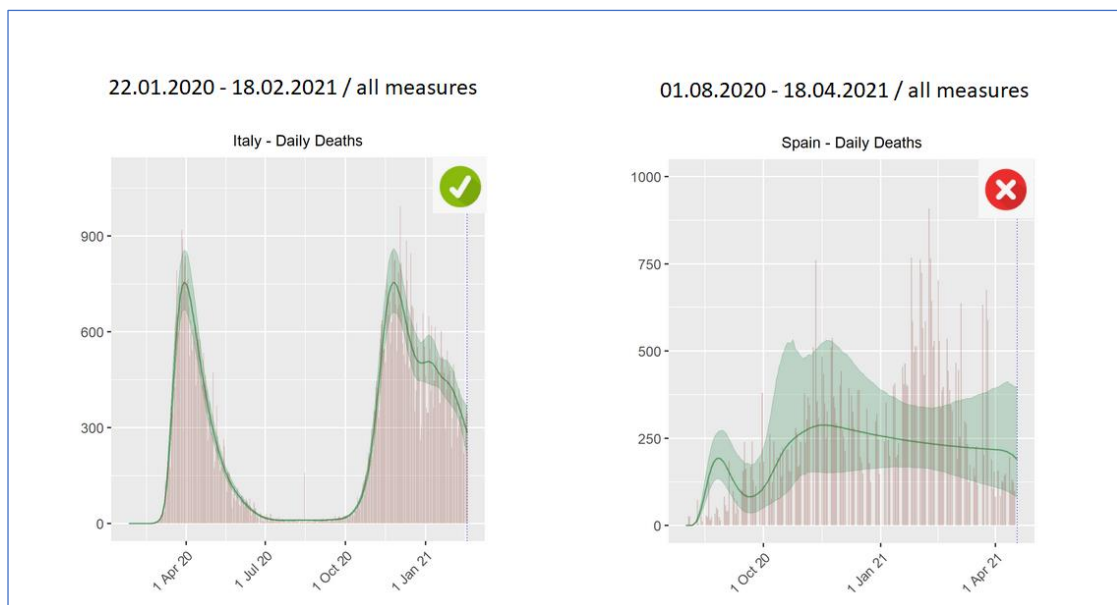


Figura 12 - Clasificación de modelos en función de la bondad del ajuste

5.3.2 Bondad del ajuste de los distintos modelos

Este gráfico resume la distribución de los modelos de los 32 países europeos analizados teniendo en cuenta su clasificación de acuerdo con los criterios descritos en el apartado anterior.

Los resultados detallados sobre la clasificación de cada modelo están disponibles en el servidor *web* [9]: [Valoracion_de_los_resultados_de_ajuste.xlsx](#)

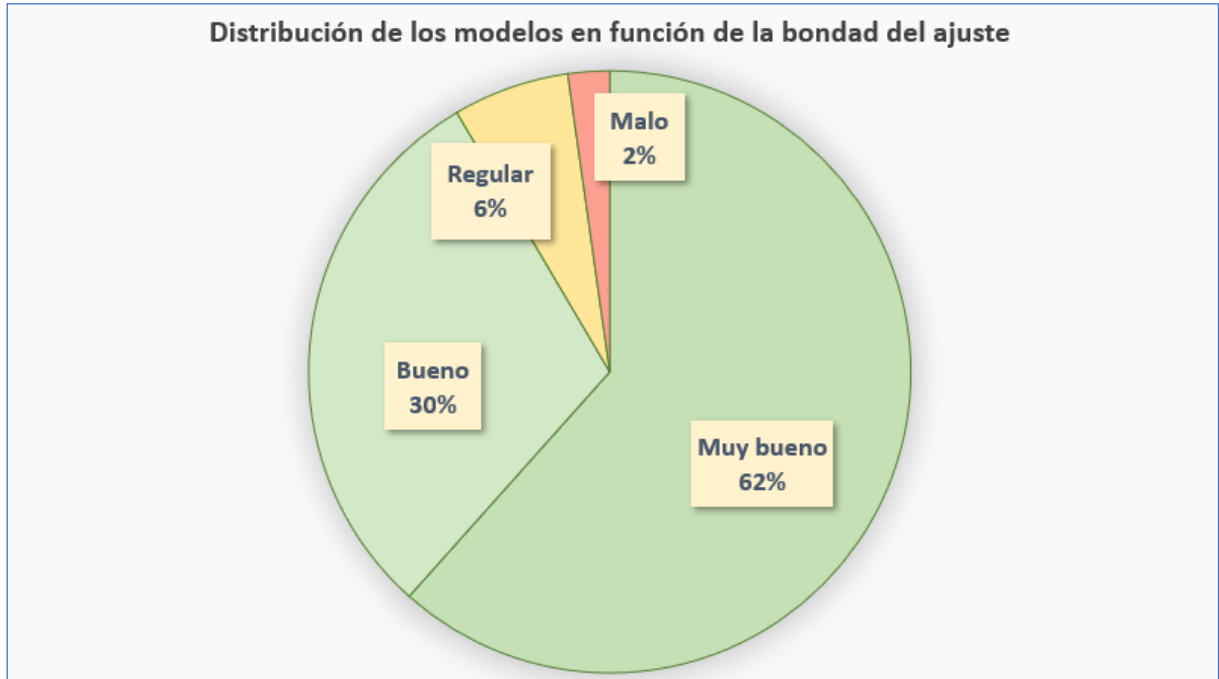


Figura 13 - Distribución de los modelos para 32 países europeos en función de la bondad del ajuste

Para 6 de los 32 países, todos sus modelos se han clasificado como *Muy buenos*, entre ellos algunos con un desarrollo muy distinto de la enfermedad, como son Italia y Portugal, por ejemplo.

Por el contrario, en el caso de *Malta*, sólo uno de sus 7 modelos se ha clasificado como *Muy bueno*, siendo la clasificación de los otros 6 sólo *Regular*. Esto es aceptable, al tratarse de un país con pocos casos de *COVID-19*, lo que dificulta el ajuste.

Un caso un tanto decepcionante es el de España: 4 modelos *Buenos*, 2 *Regulares* y uno *Malo*. Las discrepancias entre las predicciones y los datos reales se observan a partir de la segunda ola de la pandemia. Al igual que para otros países, esto parece indicar que ha habido otros factores, además de las medidas de intervención, que han influido en el desarrollo de la enfermedad y que no se han tenido en cuenta en el análisis, como pueden ser la reducción de casos a causa de la mayor capacidad de detección de la enfermedad o el efecto de la vacunación.

5.3.3 Evaluación de la bondad del ajuste - Conclusiones

De los 224 modelos generados para 32 países europeos, únicamente con 5 (el 2%) se han obtenido resultados de ajuste que pueden considerarse malos y que no permiten su uso para extraer conclusiones.

Además, 14 de los modelos (6%) se han clasificado como *Regular*. Estos modelos ofrecen, en general, un buen ajuste, pero no se han clasificado como *Buenos*, bien porque en algunos periodos las predicciones se desvían de los datos reales, bien porque el rango que incluye el 95% de las predicciones es demasiado amplio.

Con las distintas opciones de configuración que ofrece la herramienta informática, sería posible hacer pruebas con combinaciones de parámetros distintas a las usadas por defecto para encontrar variantes que mejoraran los modelos menos buenos. Sin embargo, se ha considerado adecuado mostrar los resultados originales, para hacer visibles no sólo las bondades, sino también las posibles limitaciones del modelo estadístico.

En cualquier caso, la gran mayoría de los resultados han sido *Buenos* (30%) o *Muy buenos* (62%), como se ha documentado en las páginas anteriores.

Esto permite concluir que el modelo estadístico utilizado en este trabajo es un buen método para describir la evolución del *COVID-19*, y que puede utilizarse tanto para obtener conclusiones sobre el efecto de las distintas medidas de intervención como para hacer predicciones sobre la evolución futura de la enfermedad.

5.4 Estimaciones para el número de reproducción básico (R_0)

Las estimaciones del efecto de las medidas de intervención se obtienen, de acuerdo con el modelo mecanístico (“4.3 Fundamento estadístico - Modelo mecanístico”), a partir de su contribución a la reducción del número de reproducción (R_t). Por tanto, una estimación realista del valor inicial R_0 (número de reproducción básico) es fundamental para obtener resultados plausibles.

Además, como se ha dicho (“4.1 Número de reproducción (R_t)”), este valor permite calcular el porcentaje de población que debe estar inmunizada poder volver a la normalidad sin riesgo de una nueva ola de la pandemia, alcanzando la denominada inmunidad de rebaño o colectiva (*herd immunity*) [28]. Esta información es relevante para definir las estrategias de vacunación.

En este trabajo, se obtienen resultados muy dispares para R_0 entre los distintos países europeos, lo cual es plausible y consistente con el hecho de que su valor depende, en buena medida, de factores relacionados con los hábitos sociales [11].

Para los 32 países europeos analizados, se obtiene como media ponderada en función de la población un valor de 4.69 (IC 95%: 3.81-5.77). Este valor es consistente con los resultados de Linka et col. [13], que estima un valor medio de 4.22 (± 1.69) para los 27 países de la Unión Europea.

Sin embargo, algunos estudios estiman valores de R_0 significativamente menores. Por ejemplo, el trabajo de Locatelli et col. [11] llega a la conclusión de que “la propagación de la enfermedad parece ser notablemente similar en la mayoría de los países europeos, lo que nos permite estimar un R_0 medio en Europa Occidental de 2.2 (IC 95%: 1.9-2.6)” [11]. Otros estudios han obtenido resultados similares, y son los que se utilizan como valores iniciales (*prior*) en algunos trabajos (por ejemplo [6], [7]) y también en este Trabajo Fin de Máster.

Por el contrario, otros estudios ofrecen resultados distintos y más consistentes con los obtenidos en este trabajo. Entre ellos, se encuentra el trabajo de Ke et col. [27], publicado recientemente (21 de mayo de 2021) y otros trabajos anteriores como, por ejemplo, el ya mencionado de Linka et col. [13].

La tabla de la página siguiente, también disponible en [este enlace](#) [9], resume las estimaciones de R_0 obtenidas con este Trabajo Fin de Máster (resultado del ajuste de los datos de la primera ola de la pandemia) junto con los resultados de otros estudios [13], [27]. Se incluyen también los porcentajes de inmunidad necesarios para detener la extensión de la pandemia. Éstos se calculan como $1 - 1/R_0$ (ver “4.1 Número de reproducción (R_t)”) y definen el límite para alcanzar la inmunidad de rebaño o colectiva (*herd immunity*) [28].

En general, si bien hay algunas diferencias, se puede concluir que las estimaciones de este trabajo son consistentes con los resultados de los otros estudios [13], [27] en dos aspectos fundamentales:

- Hay diferencias significativas en los valores de R_0 entre los distintos países
- El promedio es mayor que los valores de algunos estudios previos, que obtienen estimaciones en rangos en torno a 2.5

Número de reproducción básico (R_0)

La tabla contiene las estimaciones del número de reproducción básico (R_0) obtenidas con este Trabajo Fin de Máster junto con las de otros dos estudios [13], [27]. Se muestran también las estimaciones para el porcentaje de población inmunizada necesario para alcanzar el límite que asegura una inmunidad de rebaño o colectiva (herd immunity) [28].

Los datos sobre la población se han copiado del estudio de Linka et col. [13] para los 27 países de la Unión europea y de worldometer [29] para los 5 países restantes.

Country	Population	Blecua, J. (2021)		Linka, K. et al. (2020)		Ke, R. et al. (2021)	
		R_0 (95% CI)	% immunity	R_0 (95% CI)	% immunity	R_0 (95% CI)	% immunity
Austria	8,840,521	4.22 (3.06 - 5.69)	76%	4.38 (4.02 - 4.74)	77%		
Belgium	11,433,256	6.91 (5.76 - 8.30)	86%	5.00 (4.27 - 5.73)	80%	3.6 (2.9 - 4.6)	72%
Bulgaria	7,025,037	2.30 (1.52 - 3.00)	57%	1.29 (1.25 - 1.33)	22%		
Croatia	4,087,843	2.32 (1.56 - 3.25)	57%	0.93 (0.71 - 1.15)			
Cyprus	1,189,265	2.32 (1.60 - 3.04)	57%	3.35 (2.21 - 4.49)	70%		
Czechia	10,629,928	4.81 (3.47 - 6.46)	79%	2.92 (2.45 - 3.39)	66%		
Denmark	5,793,636	3.68 (2.78 - 5.07)	73%	2.00 (1.95 - 2.05)	50%		
Estonia	1,321,977	2.43 (0.99 - 3.41)	59%	3.12 (2.34 - 3.90)	68%		
Finland	5,515,525	4.36 (2.91 - 6.61)	77%	1.62 (1.57 - 1.67)	38%		
France	66,977,107	4.63 (3.93 - 5.41)	78%	3.46 (3.17 - 3.75)	71%	4.4 (3.6 - 5.4)	77%
Germany	82,905,782	4.74 (3.89 - 6.03)	79%	6.33 (5.69 - 6.97)	84%	4.7 (3.8 - 5.8)	79%
Greece	10,731,726	2.67 (2.14 - 3.70)	63%	1.66 (1.54 - 1.78)	40%		
Hungary	9,775,564	2.51 (2.12 - 3.14)	60%	1.97 (1.42 - 2.52)	49%		
Iceland	343,232	1.91 (1.13 - 2.58)	48%				
Ireland	4,867,309	6.45 (4.14 - 9.22)	84%	1.94 (1.88 - 2.00)	48%		
Italy	60,421,760	5.06 (4.67 - 5.44)	80%	4.25 (3.83 - 4.67)	76%	4.6 (3.7 - 5.8)	78%
Latvia	1,927,174	1.35 (0.82 - 2.34)	26%	2.50 (1.61 - 3.39)	60%		
Liechtenstein	38,226	2.25 (1.20 - 3.05)	56%				
Lithuania	2,801,543	2.30 (1.35 - 3.10)	57%	0.91 (0.03 - 1.79)			
Luxembourg	607,950	2.76 (2.22 - 3.96)	64%	2.42 (1.21 - 3.63)	59%		
Malta	484,630	2.37 (1.48 - 3.25)	58%	2.08 (1.94 - 2.22)	52%		
Netherlands	17,231,624	7.06 (5.28 - 8.41)	86%	5.88 (5.00 - 6.76)	83%	3.5 (3.0 - 4.2)	72%
Norway	5,459,408	2.95 (2.27 - 4.35)	66%				
Poland	37,974,750	3.79 (2.57 - 5.37)	74%	2.62 (2.36 - 2.88)	62%		
Portugal	10,283,822	4.32 (3.58 - 5.24)	77%	5.10 (4.24 - 5.96)	80%		
Romania	19,466,145	3.76 (2.89 - 4.98)	73%	6.06 (5.22 - 6.90)	83%		
Slovakia	5,446,771	2.57 (1.91 - 3.70)	61%	1.46 (1.42 - 1.50)	32%		
Slovenia	2,073,894	3.05 (2.17 - 4.83)	67%	3.83 (2.87 - 4.79)	74%		
Spain	46,796,540	5.85 (4.55 - 7.05)	83%	5.19 (4.69 - 5.69)	81%	6.4 (5.2 - 8.0)	84%
Sweden	10,175,214	3.97 (2.66 - 5.49)	75%	1.89 (1.80 - 1.98)	47%		
Switzerland	8,711,558	4.04 (3.30 - 5.06)	75%			3.5 (2.8 - 4.3)	71%
United_Kingdom	67,886,011	5.38 (4.64 - 6.24)	81%			3.9 (3.3 - 4.6)	74%

Tabla 5 - Número de reproducción básico (R_0) - resultados de este trabajo y de otros modelos

5.5 Efecto de las distintas medidas de intervención

5.5.1 Efecto de las medidas de intervención - Definición

El efecto de las medidas de intervención sobre la evolución de la enfermedad se cuantifica a partir de su contribución a los cambios en el valor de R_t (*número de reproducción*), de acuerdo con las fórmulas que describen el modelo mecanístico utilizado en este trabajo (ver “4.3 Fundamento estadístico - Modelo mecanístico”).

Cada medida o evento contribuye a un cambio en el valor de R_t de acuerdo con la siguiente fórmula.

$$(R_t)_d = R_0 \cdot \exp(-\sum_i \alpha_i)$$

Las estimaciones para los valores de α_i (más exactamente, las distribuciones *posteriores* de éstos) se obtienen como resultado del ajuste.

El factor $\exp(-\sum_i \alpha_i)$ es, por tanto, una indicación del efecto de cada medida sobre R_t . Valores menores que 1 suponen una reducción del valor de R_t , mientras que valores mayores hacen que éste aumente. En este trabajo, se utiliza este factor como medida cuantitativa del efecto de cada medida.

El estudio de [Haug et col. \[12\]](#) utiliza como criterio para medir y comparar la efectividad de las medidas las diferencias absolutas en el valor de R_t pero, como se ha comentado en “3 Estado del arte”, esto no se considera adecuado como solución genérica.

5.5.2 Datos utilizados para el análisis del efecto de las medidas de intervención

El efecto de las distintas medidas se puede analizar de forma individual para cada país, a partir de los datos, tanto numéricos como gráficos, que se generan con cada uno de los modelos (ver “4.6.2 Modelos y Predicciones - Bondad del ajuste y efecto de las distintas medidas”).

En este apartado, se utilizan los resultados numéricos de cada uno de los modelos para hacer un análisis conjunto del efecto de las distintas medidas en los distintos países. Además, se compara el efecto de medidas individuales entre los diferentes países y se extraen conclusiones sobre las medidas más efectivas en el conjunto de países analizados.

Los resultados se dividen en dos bloques y se documentan en las siguientes páginas:

- Análisis del efecto de las medidas de intervención en distintos países y regiones
- Comparación del efecto de las principales medidas de intervención entre distintos países

En ambos casos, se consideran únicamente los modelos generados con las intervenciones *Reducidas* (ver “5.2 Agrupación y reducción de las medidas de intervención”), que son los que permiten extraer conclusiones sobre el efecto de las medidas con un grado de incertidumbre aceptable.

Para las comparaciones entre distintos países, se han seleccionado 16 de los 32 países europeos que se han considerado más representativos, tanto por su población como por la incidencia del COVID-19. Además, se han considerado únicamente medidas de intervención aplicadas en al menos 6 de los 16 países seleccionados, si bien los periodos con un efecto negativo sobre R_t se han tenido en cuenta independientemente del número de países afectados.

Las siguientes tablas resumen los datos utilizados para el análisis. Los enlaces de la columna *Resultados* permiten acceder directamente a todos los resultados gráficos.

En los siguientes apartados se comentan algunos ejemplos y se resumen las conclusiones obtenidas a partir de los resultados.

Análisis del efecto de las medidas de intervención en distintos países y regiones













	Datos de <i>COVID-19</i>		Intervenciones	Número de modelos resultantes	Resultados	
	Países / Regiones	Rango de fechas				
1	Países europeos	22.01.2020 - 18.04.2021	Reducidas	32		
2	Países europeos	22.01.2020 - 18.02.2021	Reducidas	32		
3	Países europeos	22.01.2020 - 31.07.2020	Reducidas	32		
4	Países europeos	01.08.2020 - 18.04.2021	Reducidas	32		
8	Regiones españolas	14.05.2020 - 09.04.2021	Reducidas	19		
9	Regiones alemanas	14.05.2020 - 09.04.2021	Reducidas	16		

Tabla 6 - Datos utilizados para el análisis del efecto de las medidas de intervención

Comparación del efecto de las principales medidas de intervención entre distintos países









	Datos de <i>COVID-19</i>		Intervenciones	Número de países comparados	Resultados	
	Países / Regiones	Rango de fechas				
1	Países europeos	22.01.2020 - 18.04.2021	Reducidas	16		
2	Países europeos	22.01.2020 - 18.02.2021	Reducidas	16		
3	Países europeos	22.01.2020 - 31.07.2020	Reducidas	16		
4	Países europeos	01.08.2020 - 18.04.2021	Reducidas	16		

Tabla 7 - Datos utilizados para la comparación del efecto de las medidas entre distintos países

5.5.3 Resultados para el periodo del 22.01.2020 al 18.04.2021

Como ejemplo del tipo de análisis realizado, se eligen algunos resultados del periodo más extenso (22.01.2020 - 18.04.2021) para mostrar la forma de interpretar los datos y documentar algunas de las conclusiones obtenidas.

Resultados para cada país por separado

El siguiente ejemplo representa el efecto de las distintas medidas en España. Se puede acceder a los gráficos de los 32 países analizados a través de [este enlace \[9\]](#) o pulsando sobre el gráfico:

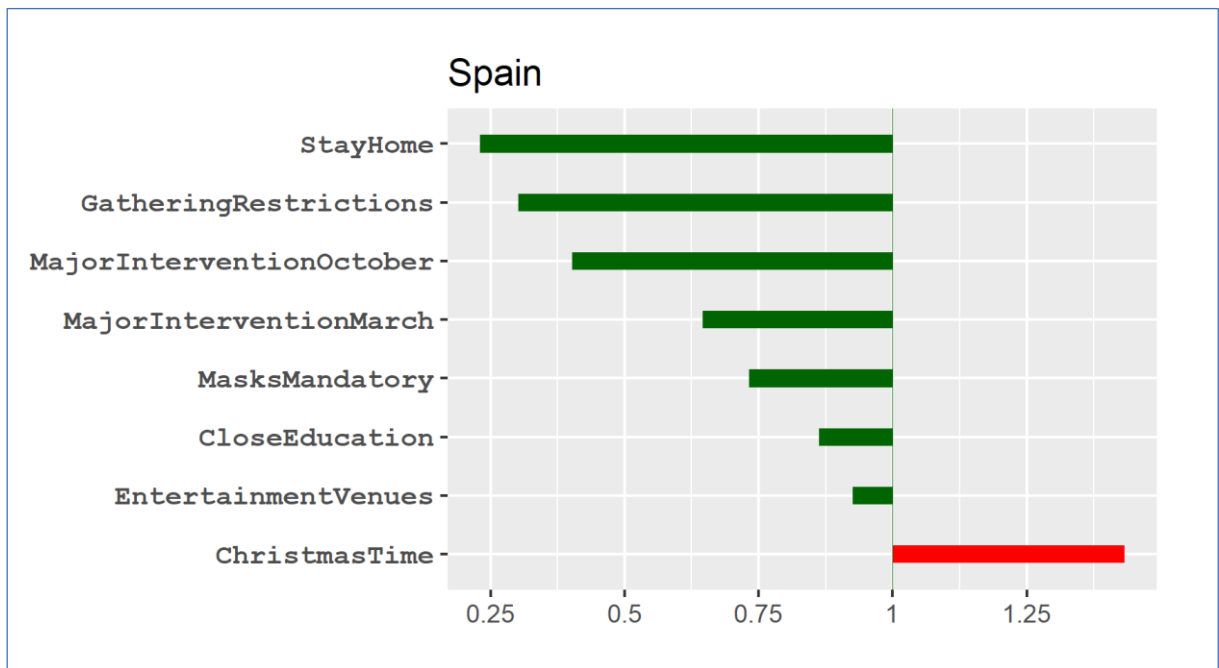


Figura 14 - Efecto de las medidas de intervención en España

Lo que se representa es la media del valor de cada factor $e^{-\alpha}$ (ver detalles en la descripción del modelo mecanístico: “4.3 Fundamento estadístico - Modelo mecanístico”), en un formato en el que una mayor longitud de las barras verdes significa una mayor contribución al descenso de R_t y, por tanto, una mayor efectividad de las medidas.

A partir de gráficos como éste se pueden extraer conclusiones sobre el efecto de las distintas medidas en el país que se analiza. En este caso, algunas conclusiones para España son:

- En el periodo analizado, las medidas más efectivas fueron las relacionadas con el confinamiento (*StayHome*) y con la limitación de reuniones y contactos sociales (*GatheringRestrictions*)
- El periodo navideño supuso un aumento de casos que no se explica con la desactivación de ninguna de las medidas activas. Esto se interpreta como una consecuencia de las reuniones familiares en el periodo festivo

En el gráfico de la página anterior, es llamativo el hecho de que la *MajorInterventionMarch* (amplio conjunto de medidas aplicadas en marzo de 2020) no aparezca entre las intervenciones con un mayor efecto. Esto se debe a que, de forma simultánea a la aplicación de estas medidas, se aplicó la medida de confinamiento obligatorio. Esta última, si bien más limitada en el tiempo que otras, muestra estadísticamente un mayor impacto en la reducción de R_t , lo que hace que el efecto de otras medidas quede parcialmente *oculto*.

En otros países, los resultados son opuestos por motivos similares. Por ejemplo, en la República Checa, la medida con mayor efecto es la documentada como *MajorInterventionMarch*, mientras que el valor estimado para la medida de confinamiento *StayHome* es prácticamente irrelevante:

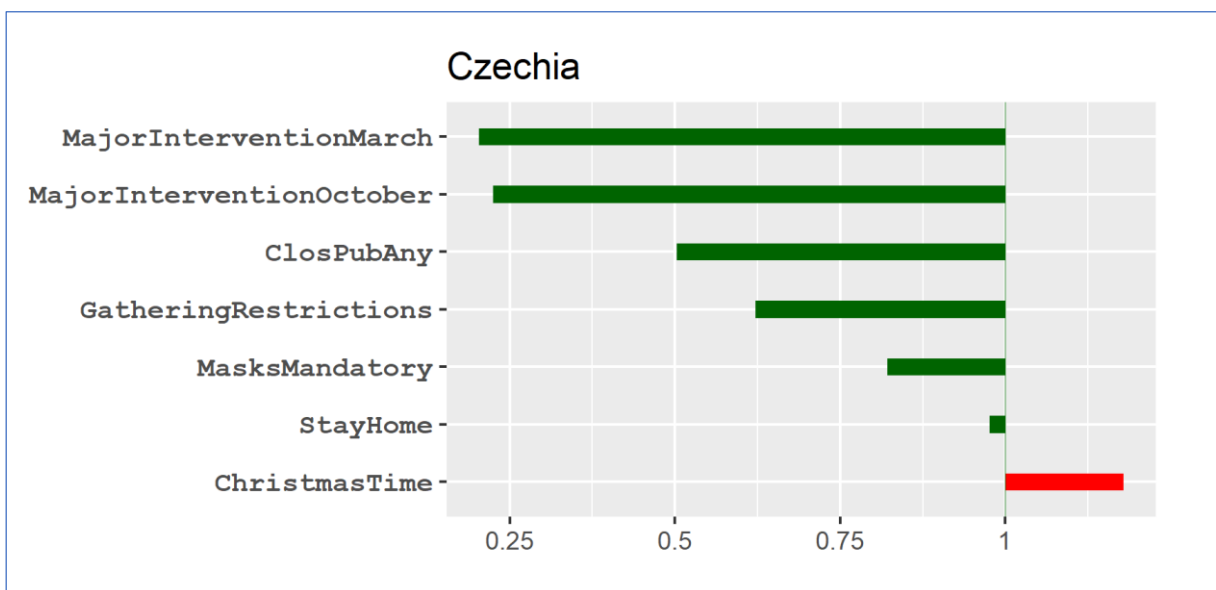


Figura 15 - Efecto de las medidas de intervención en la República Checa

Este tipo de resultados, aparentemente inconsistentes o poco plausibles, son muestras de las limitaciones del modelo, en este caso causadas por la naturaleza de los datos.

Esto ha de servir como indicación de que, más allá de los resultados puramente numéricos, los resultados del análisis requieren una interpretación individual y razonada antes de ser utilizados para sacar conclusiones a partir de ellos.

Comparación del efecto de cada medida en los distintos países

Otro resultado del análisis es la comparación del efecto de algunas de las medidas entre distintos países.

Como se ha dicho, para este análisis se han tenido en cuenta sólo datos de algunos países que se consideran más representativos. De los 32 países analizados, se han elegido 16, principalmente en función de su población y del número de muertes por COVID-19.

Por otro lado, se han considerado únicamente medidas de intervención aplicadas en al menos 6 de los 16 países seleccionados, si bien los periodos con un efecto negativo sobre R_t se han tenido en cuenta independientemente del número de países afectados.

Los resultados quedan reflejados en gráficos como éste:

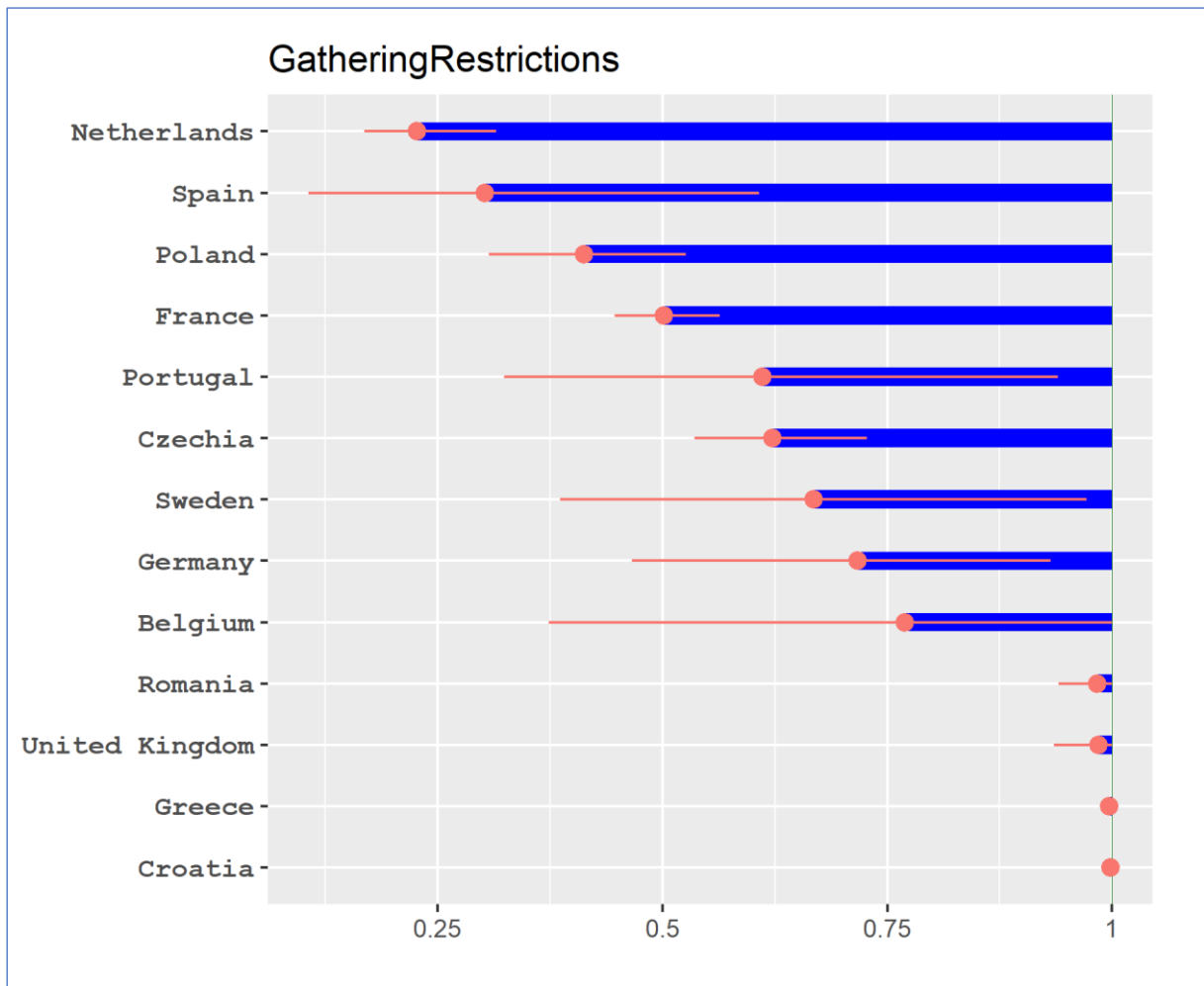


Figura 16 - Efecto de la limitación de reuniones y contactos sociales en distintos países

Los gráficos para cada una de las medidas consideradas para este análisis, de acuerdo con los criterios mencionados, se encuentran en [este enlace \[9\]](#) o pulsando sobre el gráfico anterior.

En el gráfico de la página anterior se documenta el efecto de las medidas relacionadas con la limitación de reuniones y contactos sociales (*GatheringRestrictions*) en los distintos países.

Se representa la media de las estimaciones para cada factor $e^{-\alpha}$ y el rango de valores que incluye el 95% de las predicciones. Este gráfico permite sacar, entre otras, las siguientes conclusiones:

- Este tipo de medidas tuvo un efecto significativo en países como Holanda, España y Polonia, si bien en el caso de España las estimaciones muestran un alto grado de incertidumbre.
- En otros países, el efecto fue menor, y en algunos de ellos (por ejemplo Rumanía y Reino Unido) fue irrelevante.

Los gráficos para periodos con un efecto negativo sobre R_t son similares, como se ve en el correspondiente al periodo navideño, que muestra los países para los que hubo un aumento de casos que, estadísticamente, no se puede explicar por la desactivación de medidas de intervención:

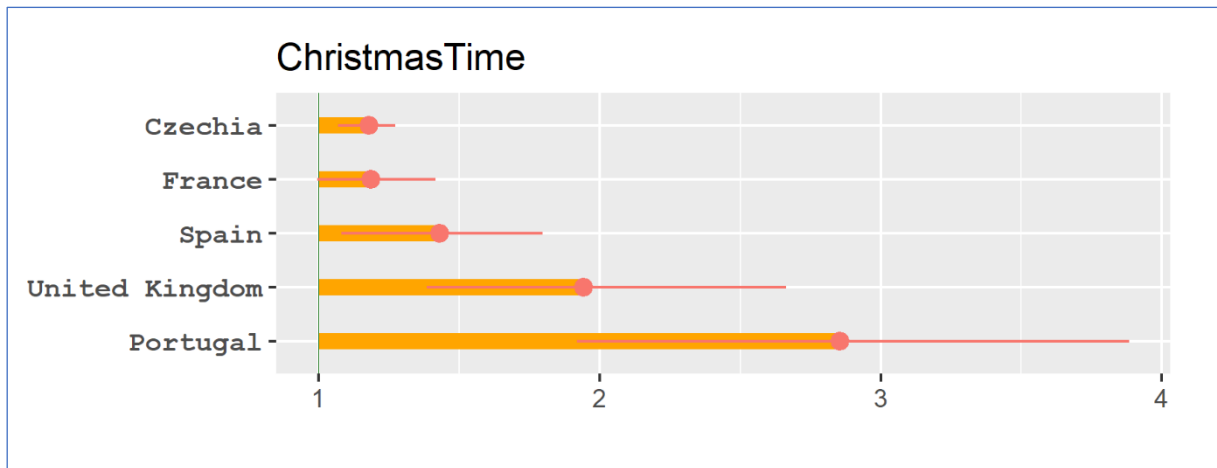


Figura 17 - Efecto del periodo navideño en distintos países

Como se ha comentado antes para el caso de España, se interpreta que el aumento de casos de COVID-19 en algunos países durante el periodo navideño es consecuencia de las reuniones familiares en esas fechas.

5.5.4 Medidas más efectivas en el conjunto de países en distintos periodos

Las siguiente tablas documentan las tres medidas más efectivas en los 16 países seleccionados en distintos periodos, incluyendo la información numérica sobre su efecto (valor del factor $e^{-\alpha}$). Pulsando en cada una de las tablas se accede a la versión disponible en el servidor *web* [9].

Periodo del 22.01.2020 al 31.07.2020 - primera ola de la pandemia

Most effective measures	1st	2nd	3rd
Belgium	StayHome (0.22)	MasksMandatory (0.34)	GatheringRestrictions (0.63)
Bulgaria	MajorInterventionMarch (0.73)	CloseEducation (0.76)	ClosPubAny (0.91)
Croatia	MajorInterventionMarch (0.57)	GatheringRestrictions (0.70)	MasksMandatory (0.78)
Czechia	MasksMandatory (0.27)	MajorInterventionMarch (0.68)	StayHome (0.86)
France	MajorInterventionMarch (0.29)	MasksMandatory (0.37)	StayHome (0.55)
Germany	StayHome (0.29)	MasksMandatory (0.51)	ClosPubAny (0.73)
Greece	StayHome (0.66)	CloseEducation (0.75)	MajorInterventionMarch (0.77)
Hungary	MasksMandatory (0.41)	MajorInterventionMarch (0.44)	-
Italy	MajorInterventionMarch (0.15)	MajorInterventionJuni (0.18)	CloseEducation (0.69)
Netherlands	MajorInterventionMarch (0.20)	ClosPubAny (0.60)	GatheringRestrictions (0.76)
Poland	StayHome (0.27)	GatheringRestrictions (0.52)	ClosureOfPublicTransport (0.69)
Portugal	ClosureOfPublicTransport (0.63)	ClosPubAny (0.67)	StayHome (0.69)
Romania	ClosPubAny (0.51)	StayHome (0.64)	CloseEducation (0.66)
Spain	MajorInterventionMarch (0.56)	StayHome (0.57)	CloseEducation (0.66)
Sweden	Teleworking (0.54)	StayHome (0.67)	GatheringRestrictions (0.76)
United Kingdom	StayHome (0.32)	MajorInterventionMarch (0.51)	ClosPubAny (0.56)

Tabla 8 - Medidas más efectivas en distintos países (22.01.2020 - 31.07.2020)

Los resultados parecen inequívocos en cuanto a las intervenciones más eficaces en la mayoría de los países durante la primera ola de la pandemia:

Intervención	Descripción
<i>StayHome</i>	Bajo este nombre genérico se han agrupado las medidas relacionadas con el confinamiento (ver detalles en "10.1 Anexo I - Nombres de las medidas de intervención"): <i>StayHomeOrder</i> , <i>RegionalStayHomeOrder</i> , <i>StayHomeGen</i> , <i>StayHomeRiskG</i> , <i>RegionalLockdown20</i>
<i>MajorInterventionMarch</i>	Conjunto de intervenciones aplicadas en marzo de 2020. Las medidas fueron distintas dependiendo del país.
<i>MasksMandatory</i>	Uso obligatorio de mascarillas en distintas variantes: <i>MasksMandatoryAllSpaces</i> , <i>MasksMandatoryClosedSpaces</i>

Tabla 9 - Medidas más efectivas en distintos países (22.01.2020 - 31.07.2020) - resumen

Periodo del 01.08.2020 al 18.04.2021 - después de la primera ola de la pandemia

Most effective measures	1st	2nd	3rd
Belgium	StayHome (0.11)	MasksMandatory (0.60)	MajorInterventionNovember (0.99)
Bulgaria	CloseEducation (0.41)	MajorInterventionNovember (0.57)	MasksMandatory (0.61)
Croatia	MajorInterventionNovember (0.80)	ClosPubAny (0.95)	RestaurantsCafes (0.99)
Czechia	MajorInterventionOctober (0.45)	ClosPubAny (0.74)	GatheringRestrictions (0.97)
France	CloseEducation (0.67)	StayHome (0.91)	MajorInterventionOctober (0.96)
Germany	GatheringRestrictions (0.65)	StayHome (0.76)	MajorInterventionDecember (0.80)
Greece	MajorInterventionNovember (0.92)	StayHome (0.99)	-
Hungary	ClosPubAny (0.77)	StayHome (0.92)	MajorInterventionNovember (0.99)
Italy	CloseEducation (0.42)	StayHome (0.46)	MajorInterventionJuni (0.77)
Netherlands	RestaurantsCafes (0.49)	GatheringRestrictions (0.51)	StayHome (0.69)
Poland	MasksMandatory (0.29)	GatheringRestrictions (0.76)	ClosureOfPublicTransport (0.90)
Portugal	GatheringRestrictions (0.52)	StayHome (0.56)	MajorInterventionMai (0.61)
Romania	CloseEducation (0.34)	GatheringRestrictions (0.87)	StayHome (0.93)
Spain	CloseEducation (0.32)	MajorInterventionOctober (0.53)	MajorInterventionMarch (0.81)
Sweden	StayHome (0.40)	GatheringRestrictions (0.62)	CloseEducation (0.72)
United Kingdom	StayHome (0.59)	GatheringRestrictions (0.99)	-

Tabla 10 - Medidas más efectivas en distintos países (01.08.2020 - 18.04.2021)

En este periodo, las tres medidas que destacan entre las más efectivas son:

Intervención	Descripción
StayHome	Medidas relacionadas con el confinamiento: StayHomeOrder, RegionalStayHomeOrder, StayHomeGen, StayHomeRiskG, RegionalLockdown20
GatheringRestrictions	Medidas relacionadas con la limitación de reuniones y contactos sociales: BanOnAllEvents, PrivateGatheringRestrictions, MassGatherAll, IndoorOverNNN, OutdoorOverNNN, SocialCircle
CloseEducation	Cierre de centros educativos: ClosDaycare, ClosPrim, ClosSec, ClosHigh

Tabla 11 - Medidas más efectivas en distintos países (01.08.2020 - 18.04.2021) - resumen

Periodo del 22.01.2020 al 18.04.2021 - periodo completo desde el inicio de la pandemia

Most effective measures	1st	2nd	3rd
Belgium	StayHome (0.14)	MasksMandatory (0.32)	GatheringRestrictions (0.77)
Bulgaria	NonEssentialShops (0.47)	CloseEducation (0.84)	ClosPubAny (0.87)
Croatia	MajorInterventionMarch (0.62)	StayHome (0.78)	MasksMandatory (0.79)
Czechia	MajorInterventionMarch (0.20)	MajorInterventionOctober (0.23)	ClosPubAny (0.50)
France	MajorInterventionMarch (0.08)	MasksMandatory (0.10)	GatheringRestrictions (0.50)
Germany	StayHome (0.29)	MasksMandatory (0.51)	CloseEducation (0.54)
Greece	CloseEducation (0.68)	MajorInterventionMarch (0.73)	StayHome (0.74)
Hungary	MajorInterventionMarch (0.25)	MasksMandatory (0.34)	StayHome (0.89)
Italy	MajorInterventionMarch (0.18)	StayHome (0.39)	MasksMandatory (0.56)
Netherlands	MajorInterventionMarch (0.11)	RestaurantsCafes (0.17)	GatheringRestrictions (0.23)
Poland	StayHome (0.13)	GatheringRestrictions (0.41)	MasksMandatory (0.61)
Portugal	StayHome (0.35)	MajorInterventionMarch (0.43)	MajorInterventionMai (0.44)
Romania	StayHome (0.33)	ClosPubAny (0.45)	CloseEducation (0.92)
Spain	StayHome (0.23)	GatheringRestrictions (0.30)	MajorInterventionOctober (0.40)
Sweden	StayHome (0.48)	GatheringRestrictions (0.67)	CloseEducation (0.70)
United Kingdom	StayHome (0.30)	ClosPubAny (0.49)	MajorInterventionMarch (0.59)

Tabla 12 - Medidas más efectivas en distintos países (22.01.2020 - 18.04.2021)

Considerando los datos desde el inicio de la pandemia hasta fechas actuales, los resultados, si bien distribuidos de forma ligeramente distinta son, en conjunto, muy similares a los de la primera ola.

En este periodo, el caso de **Bulgaria**, único país cuyas medidas más efectivas no incluyen la intervención de marzo de 2020 ni las medidas de confinamiento, merece un análisis más detallado:

- Si bien en marzo de 2020 se aplicaron en Bulgaria un conjunto amplio de medidas (un total de 17 entre el 8 y el 13 de marzo - ver detalles [aquí](#)), su efecto sobre R_t (*número de reproducción*) no es relevante en comparación a periodos posteriores, ya que en esa época el número de casos en Bulgaria era muy reducido.

Por tanto, el impacto de esta intervención, que pudo ser efectiva como medida preventiva, no se ve reflejado en los resultados. Esto puede considerarse una limitación del modelo.

- El hecho de que las medidas relacionadas con el confinamiento (*StayHome*) no aparezcan entre las más efectivas tiene una explicación mucho más sencilla: de acuerdo con la información disponible en el [European Centre for Disease Prevention and Control](#) [22], en Bulgaria no se han aplicado medidas de confinamiento.

5.5.5 Efecto de las medidas de intervención - Conclusiones

Como se ha comentado, el hecho de que en algunos periodos se hayan aplicado varias medidas simultáneamente hace que no siempre sea posible *aislar* el efecto de medidas individuales.

En otros casos, como el ejemplo de las medidas *MajorInterventionMarch* y *StayHome* en algunos países, mencionado antes, los resultados pueden parecer poco plausibles.

Además, el ejemplo de Bulgaria, analizado en páginas anteriores, ha mostrado que el modelo no siempre es capaz de detectar el efecto de medidas preventivas en periodos de baja incidencia.

Todos estos aspectos son muestras de las limitaciones del modelo, algunas de ellas causadas por la naturaleza de los datos.

A pesar de estas limitaciones, los resultados permiten obtener algunas conclusiones que, como se ve en las páginas anteriores, son plausibles y consistentes con la información conocida sobre la evolución de la enfermedad, entre otras:

- Analizando los datos de la primera ola de la pandemia para un grupo de 16 países europeos, se ha llegado a la conclusión de que, en general, las medidas más efectivas contra el *COVID-19* han sido las relacionadas con el confinamiento, los conjuntos de medidas aplicadas simultáneamente en marzo de 2020 y el uso de mascarillas.
- Teniendo en cuenta el periodo a partir del 1 de agosto de 2020, después de la primera ola de la pandemia, se identifican como medidas más eficaces las relacionadas con el confinamiento, la limitación de reuniones y contactos sociales y el cierre de centros educativos.
- Las conclusiones anteriores, válidas como resumen para el conjunto de los 16 países, cambian para algunos de ellos si se analizan de forma individual. Se ha comentado como ejemplo el caso de Bulgaria, cuyas medidas más efectivas en el periodo más extenso analizado difieren de las de la mayoría de los otros países.
- Para algunos países, entre ellos España, se observa un aumento de casos en el periodo navideño que no se explica por la desactivación de alguna de las medidas de intervención. Esto se ha interpretado como consecuencia de las reuniones familiares en el periodo festivo.

Por un lado, los resultados confirman algo que, con la perspectiva de los meses pasados, parece obvio: que algunas de las medidas de intervención han tenido un impacto fundamental en la limitación de la propagación de la enfermedad.

Por otro lado, se ha comprobado que la metodología implementada en este trabajo, extensión de trabajos anteriores [6], [7], es adecuada para describir la evolución de casos y muertes por *COVID-19* y para evaluar el efecto de las medidas de intervención.

Los trabajos anteriores [6], [7], por el momento en que se crearon, limitaban el análisis a un periodo reducido de tiempo y a un máximo de tres periodos de intervención. Con los cambios desarrollados en este trabajo, se ha puesto de manifiesto que la metodología es adecuada también para periodos largos de tiempo, con datos de la enfermedad que pueden seguir patrones muy distintos y con un número alto de medidas de intervención.

5.6 Predicciones en distintos escenarios

Se ha utilizado la herramienta informática para obtener predicciones de dos formas distintas:

- Con modelos creados a partir de datos hasta el 18.02.2021, se han hecho predicciones de la evolución de la enfermedad hasta el 18.04.2021. Los resultados se han comparado con los datos reales disponibles hasta esta última fecha, para comprobar la bondad de las predicciones.
- Con modelos creados a partir de datos hasta fechas más actuales (18.04.2021), se han hecho predicciones sobre la evolución de la enfermedad hasta el 18.06.2021 bajo distintos supuestos.

La siguiente tabla resume las dos variantes e incluye enlaces para acceder directamente a todos los resultados:



	Datos de COVID-19		Intervenciones	Predicciones hasta	Número de países	Resultados
	Países / Regiones	Rango de fechas				
1	Países europeos	22.01.2020 - 18.02.2021	Reducidas	18.04.2021	11	
2	Países europeos	22.01.2020 - 18.04.2021	Reducidas	18.06.2021	16	

Tabla 13 - Datos utilizados para el cálculo de predicciones

Las fechas elegidas están relacionadas con el periodo en el que se ha desarrollado este trabajo y con la disponibilidad de algunos datos:

- El 22.01.2020 es la fecha a partir de la que hay datos de COVID-19 disponibles en la fuente utilizada en este trabajo [24].
- El 18.02.2021 coincide aproximadamente con la fecha de inicio del trabajo. Los datos de COVID-19 disponibles hasta entonces son los utilizados para los primeros análisis.
- Los datos del 18.04.2021 son más actuales y han permitido extender el alcance de los análisis a periodos más largos de tiempo.
Esta fecha coincide en el tiempo con una actualización del 15.04.2021 de las medidas de intervención documentadas en el [European Centre for Disease Prevention and Control](#) [22].
- El 18.06.2021 se ha elegido para algunas de las predicciones tanto por englobar un espacio de dos meses (igual que el periodo anterior, entre el 18.02. y el 18.04.) como por coincidir, aproximadamente, con la finalización de este trabajo.

5.6.1 Modelos con datos hasta el 18.02.2021 / Predicciones hasta el 18.04.2021

En un primer paso, se han utilizado modelos creados con datos hasta el 18.02.2021 con un conjunto reducido/optimizado de medidas de intervención.

Se puede acceder a la información de estos modelos a través de [este enlace](#), donde los modelos utilizados para hacer predicciones son los que ofrecen el opción **Predictions based on this model** para acceder a ellas. El acceso a todas las predicciones es también posible a través de [este otro enlace](#).

El proceso de análisis ha sido el siguiente:

- Se han completado las medidas de intervención utilizadas para la creación de los modelos con nuevos periodos de intervención para medidas aplicadas a partir del 18.02.2021, utilizando los datos del [European Centre for Disease Prevention and Control](#) [22].
- Si entre los nuevos periodos se encuentran medidas aplicadas en el pasado, el proceso de predicción utilizará automáticamente la información disponible en el modelo para estimar su efecto.
- Si entre los nuevos periodos se encuentran medidas no aplicadas antes, la herramienta no tiene información para calcular su efecto. En este caso, se hace uso de la opción de asignar un valor predeterminado del mismo, para lo que se utiliza información de otros modelos.

Con los modelos utilizados, por tanto, el periodo entre el 18.02.2021 y el 18.04.2021 es un periodo *futuro*, por lo que las predicciones obtenidas en este rango de fechas no son un ajuste (*fit*) sobre datos conocidos, sino predicciones futuras (*forecast*).

En las siguientes páginas se comentan, como ejemplo, los resultados de las predicciones para Francia. Se describe el procedimiento de introducción de nuevas medidas para hacer las predicciones y se explica la forma de interpretar los resultados.

El [Anexo III](#) incluye resultados para algunos otros países y, como se ha dicho, es posible acceder directamente a todos los resultados a través de [este enlace](#).

Las predicciones son, en general, razonablemente buenas en cuanto que sus valores medios se adaptan bien a la tendencia de los datos reales.

En casos puntuales, las predicciones no se adaptan bien a los datos, lo que puede deberse tanto a limitaciones del método utilizado como a la existencia de hechos que han influido en la evolución de la enfermedad pero que no están reflejados en las medidas de intervención utilizadas para las predicciones.

Predicciones para Francia hasta el 18.04.2021

Entre el 18.02.2021 y el 18.04.2021, de acuerdo con la información disponible [22], en Francia se activaron algunas medidas de intervención: *RegionalStayHomeOrderPartial* entre el 26.02.2021 y el 19.03.2021 y varias medidas relacionadas con el cierre de centros educativos a partir del 06.04.2021. Estas últimas se han agrupado, siguiendo los criterios definidos en “5.2 Agrupación y reducción de las medidas de intervención”, en una medida conjunta que se ha denominado *ClosSecPrimDayCare*.

Además, por los resultados de otros modelos, se tiene la información de que el periodo de la Semana Santa supuso un aumento de casos de *COVID-19*, por lo que se define un evento adicional llamado *EasterTime* entre el 26.03.2021 y el 11.04.2021.

La siguiente tabla documenta la forma de introducir las nuevas medidas (recuadradas en verde) para calcular las predicciones. Las otras medidas son las que se utilizaron para la creación del modelo:

Country	Response_measure	date_start	date_end	MeasureAdditionalInfo
France	MassGatherAll	29/02/2020	17/09/2020	
France	OutdoorOver1000	13/03/2020	11/07/2020	
France	MajorInterventionMarch	16/03/2020	11/05/2020	
France	StayHomeOrderPartial	12/05/2020	02/06/2020	
France	MasksMandatoryClosedSpaces	18/05/2020	NA	
France	MassGatherAllPartial	18/09/2020	18/10/2020	
France	RegionalStayHomeOrderPartial	17/10/2020	28/10/2020	
France	MajorInterventionOctober	19/10/2020	NA	
France	StayHomeOrder	28/10/2020	14/12/2020	
France	ChristmasTime	10/12/2020	10/01/2021	X
France	StayHomeOrderPartial	15/12/2020	NA	
France	RegionalStayHomeOrderPartial	26/02/2021	19/03/2021	
France	EasterTime	26/03/2021	11/04/2021	1.3
France	ClosSecPrimDaycare	06/04/2021	NA	0.69

Tabla 14 - Francia - Medidas de intervención hasta el 18.04.2021

La medida *RegionalStayHomeOrderPartial* ya se ha aplicado en periodos anteriores, por lo que la herramienta informática utilizará la información sobre su efecto pasado para estimar su impacto en el periodo futuro.

Para el periodo de Semana Santa y para las medidas relacionadas con el cierre de centros educativos, se utiliza información de otros modelos para asignarles un efecto predefinido (1.3 y 0.69, respectivamente).

La activación y desactivación de las nuevas medidas se verán reflejadas en cambios en el valor de R_t , que a su vez influirán en las curvas de las predicciones para el número de casos y de muertes por *COVID-19*. Esto se puede ver en los gráficos de la página siguiente.

El efecto de las nuevas medidas queda reflejado en los *saltos* en la curva de predicciones para el **número de reproducción (R_t)** a partir del 18.02.2021:

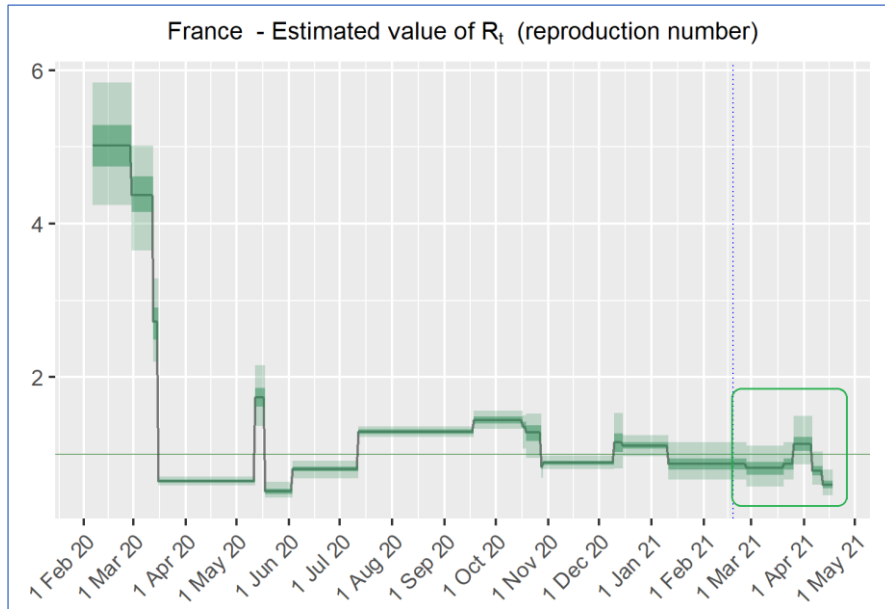


Figura 18 - Francia - Predicciones para el número de reproducción (R_t) hasta el 18.04.2021

Con los resultados para R_t , se obtienen las **predicciones para el número de muertes** (gráfico de la izquierda). Éstas se pueden comparar con los datos reales (gráfico de la derecha).

Se puede ver que los valores medios de las predicciones *futuras* se adaptan razonablemente bien a los datos reales, pero la anchura de la franja del 95% muestra un alto grado de incertidumbre.

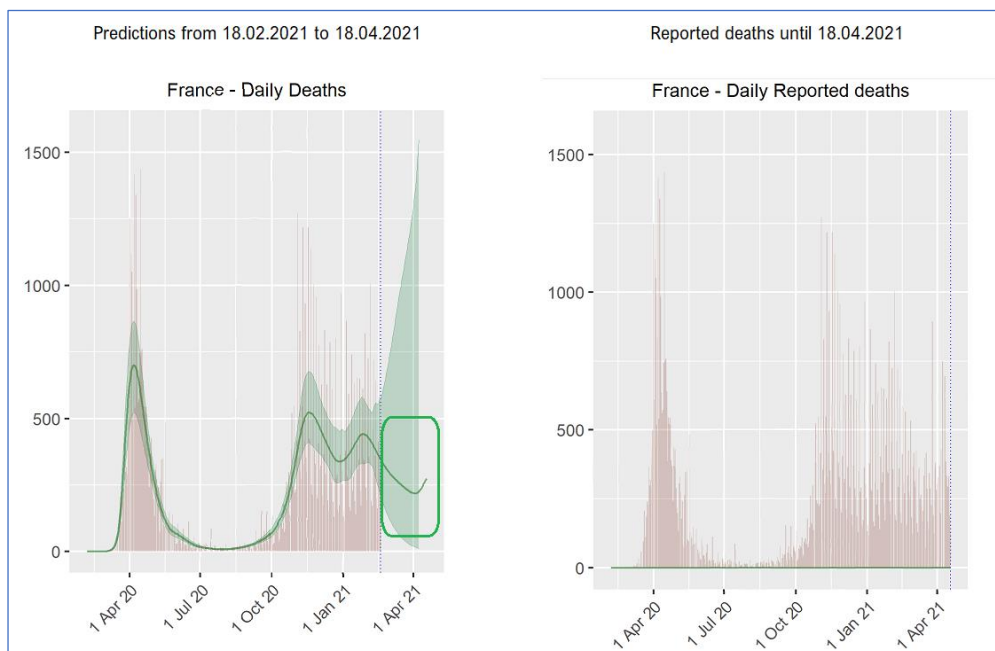


Figura 19 - Francia - Predicciones y datos reales hasta el 18.04.2021

5.6.2 Modelos con datos hasta el 18.04.2021 / Predicciones hasta el 18.06.2021

En este paso, se han obtenido predicciones utilizando modelos creados con datos de *COVID-19* hasta el 18.04.2021.

Se han definido varios escenarios, relacionados con la activación y desactivación de distintas medidas de intervención, y se han hecho predicciones a partir de ellos para algunos países europeos. Los resultados de estas predicciones están disponibles a través de [este enlace](#).

El punto de partida son los modelos creados con datos hasta el 18.04.2021 y con las medidas de intervención reducidas/optimizadas. Se puede acceder a ellos [aquí](#), donde los modelos utilizados para hacer predicciones son los que ofrecen el enlace ***Predictions based on this model*** para acceder a ellas. El acceso a todas las predicciones es también posible a través de [este otro enlace](#).

Al igual que para análisis previos, de los 32 países para los que se han creado modelos, se han elegido 16, principalmente en función de su población y del número de muertes por *COVID-19*, para hacer predicciones en distintos escenarios.

Primero, se han definido dos escenarios sencillos que se han aplicado a los 16 países para obtener las predicciones hasta el 18.06.2021:

- Escenario 1: no hay cambios en las medidas de intervención entre el 18.04.2021 y el 18.06.2021
- Escenario 2: el 31.05.2021 se desactivan todas las medidas de intervención

Después, se han definido escenarios menos *extremos* para uno de los países analizados, Alemania, consistentes en la activación y desactivación de medidas en el periodo entre el 18.04.2021 y el 18.06.2021.

Los resultados se resumen en las siguientes páginas.

Predicciones para el escenario 1 - Las medidas de intervención no cambian hasta el 18.06.2021

En este escenario, se utilizan las mismas medidas de intervención con las que se crearon los modelos.

Esto hará que no haya cambios en el valor de R_t (número de reproducción) y, por tanto, la evolución del número de casos y muertes por COVID-19 seguirá la misma tendencia de los días previos al 18.04.2021.

Hay algunas excepciones para países, para los que se ha estimado que el periodo de Semana Santa ha supuesto un aumento del valor de R_t . En estos casos, se observa un pico de casos y muertes a principios de mayo que se *corrige* después para seguir la tendencia anterior al 18.04.2021.

Por tanto, se puede concluir que el no modificar las medidas de intervención a partir del 18.04.2021 podría ser una estrategia adecuada para países en los que se observa una tendencia descendente de casos y muertes hasta esa fecha (por ejemplo, España), pero será insuficiente en países con una tendencia ascendente (por ejemplo, Grecia):

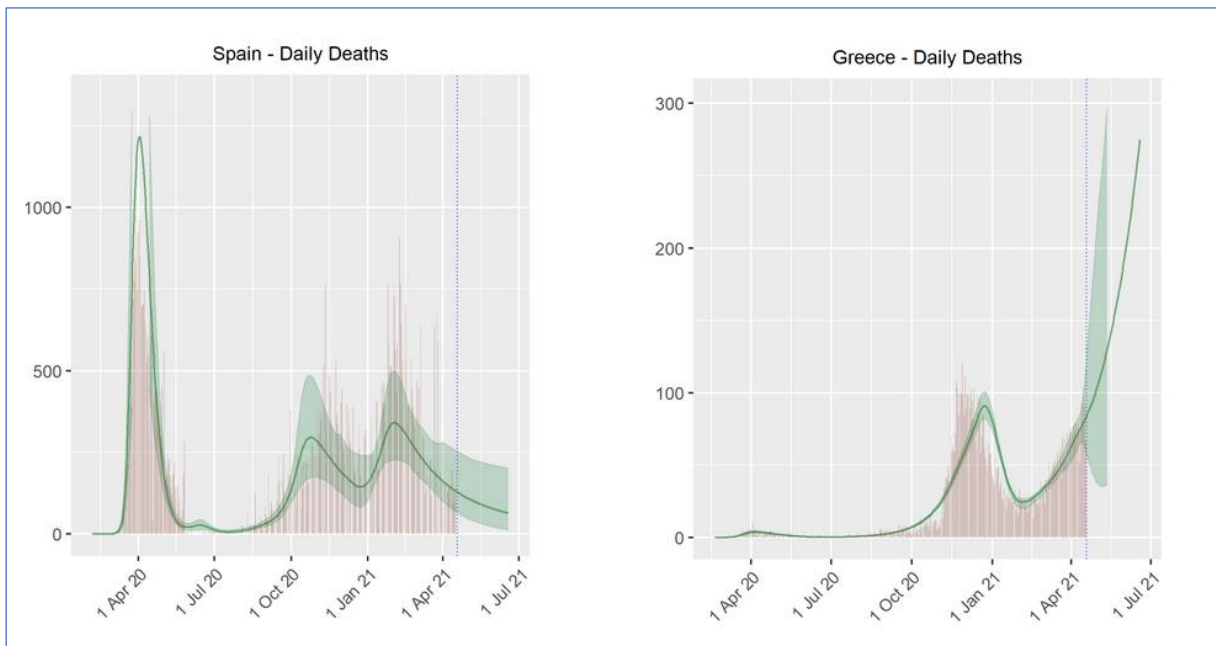


Figura 20 - Predicciones hasta el 18.06.2021 (escenario 1) - España, Grecia

Se puede acceder a las predicciones para los demás países a través de [este enlace](#), donde los resultados para este escenario son los correspondientes a la primera fila de la tabla para cada país.

A partir de ellos, se puede concluir que el no modificar las medidas de intervención sería adecuado para los siguientes países: Bélgica, República Checa, Alemania, Francia, Italia, Holanda, Portugal, España, Suecia, Reino Unido.

Sin embargo, en estos otros países habrá un aumento del número de casos y de muertes por COVID-19 si no se introducen medidas adicionales: Bulgaria, Croacia, Grecia, Hungría, Polonia, Rumanía.

Predicciones para el escenario 2 - Se desactivan todas las medidas el 31.05.2021

Por simplicidad, se asume que el desactivar todas las medidas en una fecha dada, en este caso el 31.05.2021, supondría la vuelta al valor inicial de R_t (número de reproducción) que, para los 16 países analizados en este apartado, es en todos los casos un valor mayor de 1. Esto es una simplificación no del todo correcta ya que, como se ha explicado en “4.1 Número de reproducción (R_t)”, para calcular el número de reproducción efectivo habría que tener en cuenta la proporción de población ya inmunizada.

En países con una incidencia de la enfermedad irrelevante el 31.05.2021, la vuelta al valor inicial de R_t podría no tener impacto en el aumento del número de casos. Pero en países con un número significativo de casos, aunque éste no sea muy grande, esto significaría un aumento de contagios y muertes por COVID-19.

De acuerdo con los resultados obtenidos, únicamente Portugal tendría, a fecha del 31.05.2021, un número tan reducido de casos de COVID-19 que la desactivación de todas las medidas no tendría un efecto apreciable. Por el contrario, en todos los demás países, en mayor o menor medida, el efecto de desactivar todas las medidas el 31.05.2021 se reflejaría en un aumento de casos y muertes por COVID-19. En algunos casos, los gráficos muestran un aumento probablemente desproporcionado del número de casos y muertes, resultado de la elección de un periodo relativamente amplio para las predicciones (2 meses).

Como ejemplo, se muestra la evolución del número de muertes previstas hasta el 18.06.2021 en este escenario para Portugal y para Alemania:

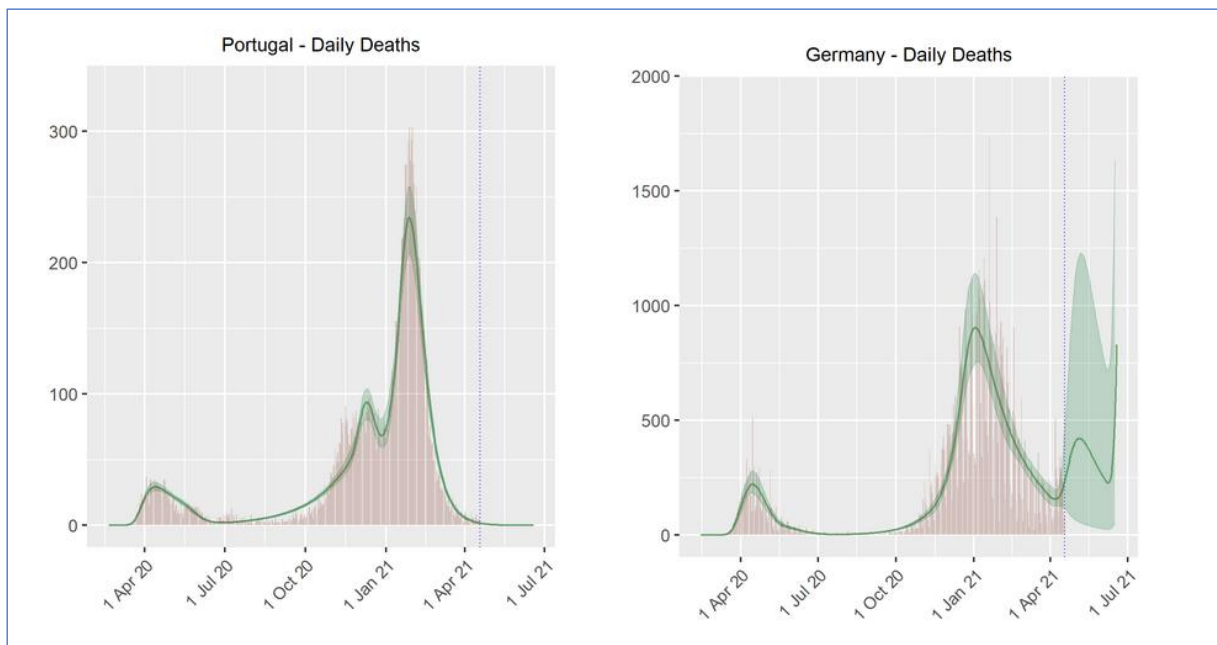


Figura 21 - Predicciones hasta el 18.06.2021 (escenario 2) - Portugal, Alemania

Se puede acceder a las predicciones para los demás países a través de [este enlace](#), donde los resultados para este escenario son los correspondientes a la segunda fila de la tabla para cada país.

Predicciones para Alemania en otros escenarios

En el momento de escribir este documento, esta información todavía no está documentada en el [European Centre for Disease Prevention and Control \[22\]](#), pero el 23 de abril de 2021 el gobierno federal alemán acordó un paquete de medidas que entraron en vigor el 24 de abril [26]. Estas medidas sustituyen o extienden las medidas activas en Alemania hasta esa fecha.

Se han hecho predicciones simulando el efecto de estas medidas por medio de una nueva medida de intervención, definida como *MajorInterventionApril2021* para el periodo del 24.04.2021 al 05.06.2021. La fecha final se ha definido de forma arbitraria con el fin de ilustrar el funcionamiento de la herramienta informática ya que, en el momento de escribir este documento, la fecha *real* todavía no es conocida.

Para este conjunto de medidas, que son algo más restrictivas que medidas anteriores, se ha asumido un efecto igual al de la medida de mayor efecto en el pasado.

A partir de estos datos, se han hecho predicciones en dos escenarios:

- Escenario 1: la intervención *MajorInterventionApril2021* es la única nueva medida a partir del 18.04.2021
- Escenario 2: además de esta intervención, se introduce una medida adicional que simula, a partir de una fecha concreta, el efecto de la vacuna. Esto es una simplificación que se incluye con la única finalidad de mostrar el funcionamiento de la herramienta en distintos escenarios.

Los resultados de las predicciones, que pueden verse [aquí](#), son plausibles:

- En el primer caso, se observa un descenso de casos, que se ve frenado con la desactivación de la intervención *MajorInterventionApril2021*.
- En el segundo caso, el aumento de casos por la desactivación de la medida anterior queda compensado por el efecto de la medida que simula el efecto de la vacunación, para el que se ha asignado un valor conveniente con fines ilustrativos.

Lo que se pretende mostrar con estos ejemplos es que, con la información adecuada, la herramienta puede utilizarse para obtener predicciones sobre la evolución futura de la enfermedad.

Los siguientes gráficos resumen los resultados de los escenarios descritos en la página anterior.

El de la izquierda refleja la predicción de un aumento de las muertes como consecuencia de la desactivación de la intervención *MajorInterventionApril2021*, mientras que el de la derecha muestra cómo la desactivación de esta medida queda compensada por el evento relacionado con el efecto la vacunación:

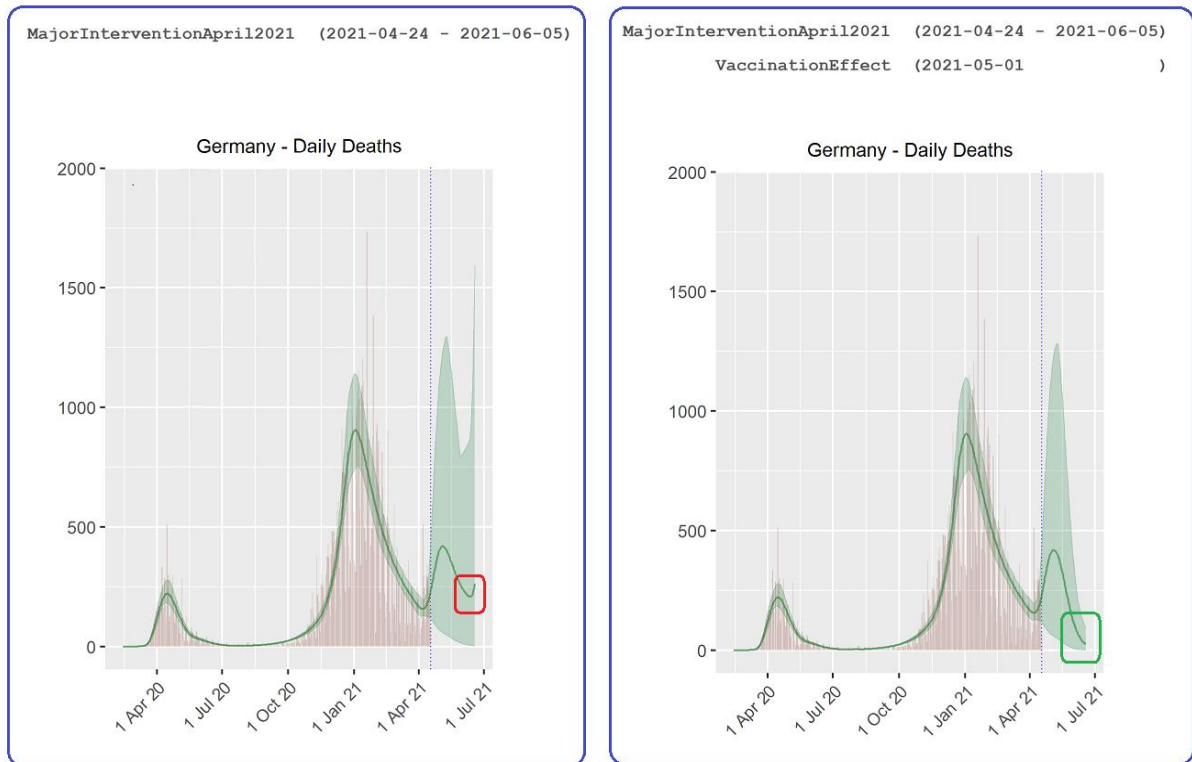


Figura 22 - Predicciones adicionales hasta el 18.06.2021 (escenarios 1 y 2) - Alemania

5.6.3 Resultados de las predicciones - Conclusiones

En un primer paso, se han obtenido predicciones hasta el 18.04.2021 a partir de modelos creados con datos de *COVID-19* hasta el 18.02.2021, lo que ha permitido evaluar la bondad de las mismas comparándolas con datos reales.

Con algunas excepciones, los resultados han sido satisfactorios, obteniendo predicciones que replican bastante bien la tendencia de los datos reales. Los resultados menos buenos, además de hacer visibles algunas limitaciones del método elegido, pueden estar causados por una parametrización no del todo óptima para las medidas de intervención aplicadas a partir del 18.02.2021.

A pesar de las limitaciones, se ha llegado a la conclusión de que la metodología desarrollada en este trabajo permite utilizar los modelos generados, cuyo ajuste, como se ha visto en “5.3 Evaluación del modelo - Bondad del ajuste” es, en general, muy bueno, para hacer predicciones sobre la evolución futura de la enfermedad.

En un segundo paso, se han utilizado modelos creados con datos más recientes de *COVID-19*, hasta el 18.04.2021, para hacer predicciones sobre la evolución de la enfermedad hasta el 18.06.2021. Estas predicciones, si bien plausibles, pueden ser demasiado pesimistas, al tenerse en cuenta únicamente el efecto de medidas no farmacológicas y de eventos limitados en el tiempo, pero no el efecto de la vacuna, cuya posible influencia en la evolución de la enfermedad no está incluida en el modelo estadístico, que es uno de los aspectos a desarrollar en el futuro.

A partir de los resultados obtenidos para las predicciones hasta el 18.06.2021, se pueden sacar las siguientes conclusiones:

- La vuelta a la normalidad, que se ha simulado mediante la desactivación de todas las medidas de intervención el 31.05.2021, sólo sería posible en el caso de Portugal.
Para los otros países analizados, la desactivación de todas las medidas en un futuro próximo supondría un aumento del número de casos y muertes por *COVID-19*.
- En algunos países, no sólo es desaconsejable desactivar las medidas activas, sino que sería necesaria la aplicación de medidas adicionales para evitar un aumento de casos a partir del 18.04.2021.

6 Discusión

Con este *Trabajo Fin de Máster*, se ha extendido un modelo existente de transmisión del *COVID-19* [6] basado en el análisis bayesiano.

Se han desarrollado funcionalidades con las que se mejora el análisis de múltiples periodos de diferente velocidad de transmisión y se posibilita la inclusión de un número potencialmente ilimitado de medidas de intervención, lo que permite una descripción más realista en rangos de tiempo más extensos que abarcan varias olas de la pandemia.

La calidad de los resultados depende del uso de datos adecuados y consistentes, por lo que se han utilizado datos de fuentes que se consideran fiables.

Aun así, en la fuente utilizada para obtener información sobre el número de casos y muertes por *COVID-19* [24], se han identificado datos anómalos (*outliers*), que tienen su origen en la forma en que se mantiene y se corrige la información. Salvo excepciones puntuales para el caso de España, no se ha invertido esfuerzo en la identificación y tratamiento de datos anómalos.

Del mismo modo, la información sobre las medidas de intervención [22] contiene, en algunos casos, información redundante. En “5.2 Agrupación y reducción de las medidas de intervención” se han definido criterios para eliminar inconsistencias y reducir el número de medidas de intervención, de forma que se posibilite la obtención de conclusiones sobre el efecto de las distintas medidas con un grado de incertidumbre aceptable.

Los otros datos de los que depende el modelo (ver “4.5 Datos necesarios para el análisis”) tienen que ver con parámetros que describen la forma en que se propaga la enfermedad. Éstos se han obtenido del estudio en el que se basa este trabajo [6] y, posiblemente, necesitan una revisión y actualización de cara a futuras versiones del trabajo.

Los resultados obtenidos, resumidos en los capítulos anteriores, permiten concluir que, a pesar de ciertas simplificaciones y limitaciones, que se comentan en el siguiente capítulo, la metodología utilizada es adecuada para describir la evolución de la enfermedad y obtener conclusiones sobre el efecto de las distintas medidas de intervención.

Otros trabajos (por ejemplo [6], [7], [13], [16], [17]) limitan el análisis a periodos breves de tiempo y con un número limitado de medidas de intervención.

Con este trabajo, se ha desarrollado una herramienta que funciona bien también con periodos largos de tiempo y con un número alto de medidas de intervención, que era el objetivo principal desde el punto de vista metodológico.

La parte del trabajo relacionada con el análisis de datos ha permitido obtener resultados a partir de los que se pueden extraer conclusiones sobre el desarrollo de la enfermedad y sobre el efecto de las distintas medidas de intervención.

Como se describe en el capítulo siguiente, las conclusiones obtenidas son plausibles y consistentes con las de otros trabajos.

7 Conclusiones

7.1 Grado de cumplimiento de los objetivos del trabajo

Los objetivos definidos inicialmente para este *Trabajo Fin de Máster* se han cumplido totalmente.

Se considera que el planteamiento del trabajo ha sido el adecuado, definiendo objetivos realistas y concretos que se han ido cumpliendo de acuerdo con la planificación prevista.

Se ha desarrollado una **herramienta informática**, extendiendo un trabajo previo [6], con la que es posible analizar datos de *COVID-19* en periodos extensos y con un número potencialmente ilimitado de medidas de intervención.

Utilizando esta herramienta, se han analizado, en distintas variantes, datos desde el inicio de la pandemia hasta el 18.04.2021, obteniendo conclusiones sobre el **efecto de las distintas medidas de intervención** que son plausibles y consistentes con las de otros estudios.

Asimismo, se han hecho **predicciones** sobre la evolución futura de la enfermedad. Esto ha servido para confirmar la capacidad de la herramienta de utilizar información sobre periodos pasados para hacer estimaciones futuras. Pero en cuanto a los resultados obtenidos, ésta es la parte menos segura de todo el trabajo debido a una de sus limitaciones: en su versión actual, el modelo no considera el efecto de la vacunación en el desarrollo de la enfermedad.

Al inicio del trabajo, se identificaron distintos **riesgos** que se han ido confirmando durante las distintas fases del proyecto. Por un lado, se mencionó el riesgo relacionado con la actualidad y la dinámica cambiante del tema objeto de estudio. Se identificaron dos aspectos recientes que no se han considerado en este trabajo, como es el efecto de nuevas variantes del virus y el efecto de la vacuna, que empezó a suministrarse en diciembre de 2020.

El no considerar el efecto de la vacuna, cuya efectividad ya tiene un impacto visible en el descenso de infecciones y muertes por *COVID-19*, hace que los resultados relacionados con predicciones futuras sean dudosos. Ésta es la única limitación con un efecto significativo en los resultados que no se ha resuelto en la versión actual del trabajo.

Además, se identificaron riesgos relacionados con la variedad de los datos y la gran cantidad de medidas de intervención aplicadas en algunos países, que también se han confirmado. Esto ha hecho necesaria la extensión de algunos aspectos metodológicos, entre otros:

- Reducción y agrupación de medidas en periodos, en los que se han aplicado simultáneamente varias medidas de intervención
- Opciones para ajustar los datos asumiendo que una misma medida tiene el mismo efecto en todos sus periodos de aplicación (opción por defecto) o que el efecto de una medida pueda ser diferente en distintos periodos
- Definición de distintos parámetros que se pueden utilizar para cada medida individual por separado: asumir que el efecto de una medida es positivo (opción por defecto), que puede ser negativo o que éste tiene un valor fijo predeterminado

Por último, se calificó como posible riesgo el carácter no determinista de la metodología utilizada para la estimación de los parámetros que describen el efecto de las distintas medidas de intervención.

Esto se ha solventado a través de la parametrización del *muestreo de Monte Carlo Hamiltoniano*:

- Se ha definido una configuración por defecto para el número de iteraciones que ha funcionado bien para la mayoría de los países, produciendo resultados de ajuste buenos y reproducibles.
- Para algunos países, el ajuste con los parámetros por defecto no ofrecía buenos resultados o éstos variaban mucho entre diferentes ajustes. Para estos casos, se ha definido una configuración individual, incrementando el número de iteraciones, con lo que se han obtenido resultados estables.

Los detalles, que se han considerado demasiado técnicos para desarrollarlos en este documento, se pueden ver en la [documentación de la herramienta informática \[9\]](#).

En **resumen**, se puede decir que se han solucionado las consecuencias de los riesgos con extensiones en la metodología, sin por ello poner en riesgo la consecución de los objetivos.

La única limitación con un impacto relevante que no se ha solucionado ha sido la de no incluir en el modelo estadístico el efecto de la vacunación. Esta limitación influye en los resultados relacionados con las predicciones futuras, pero no en las conclusiones del análisis sobre datos pasados.

Por tanto, de todos los resultados, los relacionados con la obtención de predicciones futuras son los menos fiables. Sin embargo, tanto los aspectos metodológicos como las estimaciones sobre el efecto de las distintas medidas de intervención en periodos pasados ofrecen resultados muy satisfactorios.

7.2 Seguimiento de la planificación

Como se ha comentado en el apartado anterior, se considera que el planteamiento del trabajo ha sido el adecuado, definiendo objetivos realistas y concretos que se han ido cumpliendo de acuerdo con la planificación prevista.

La metodología definida inicialmente ha resultado ser adecuada para la consecución de los objetivos, si bien han sido necesarios algunos cambios para mitigar parte de los riesgos que se identificaron al inicio del proyecto y que se han confirmado durante el desarrollo del trabajo.

Estos cambios no han supuesto cambios en la planificación ni tampoco un riesgo para la consecución de los objetivos.

7.3 Conclusiones obtenidas a partir de los resultados del trabajo

Los resultados de este trabajo permiten obtener las siguientes conclusiones:

- La metodología utilizada, basada en el análisis bayesiano, y adaptada para tener en cuenta un número potencialmente ilimitado de medidas de intervención, es adecuada para describir la evolución del COVID-19 y obtener conclusiones sobre el efecto de las distintas medidas.

Tanto los estudios en los que se basa este trabajo [6], [7] como otros estudios posteriores [17] han obtenido buenos resultados con análisis basados en este tipo de metodología.

El estudio de *Bryant et col.* [20] es una voz crítica con los métodos de análisis bayesiano. Sin embargo, los argumentos críticos tienen que ver más con los supuestos del modelo analizado [7] que con la metodología en sí misma.

- El método utilizado funciona bien en general, salvo en algunos periodos para países como Bulgaria, donde se ha visto que no es capaz de detectar el efecto preventivo de medidas aplicadas en periodos de baja incidencia de la enfermedad.
- Los resultados confirman algo que, pasado el tiempo, parece obvio: que la aplicación de medidas de intervención ha sido fundamental para frenar la extensión de la pandemia.
- Los valores para el número de reproducción básico (R_0), parámetro relacionado con el ritmo inicial de propagación de la enfermedad, son muy distintos entre los distintos países europeos y, en promedio, mayores de los estimados en algunos estudios previos.

Esto tiene como consecuencia que el porcentaje de población que debe estar inmunizada para llegar a la situación en la que la enfermedad está controlada (*inmunidad de rebaño*) varía dependiendo del país y que, en algunos países, este porcentaje es algo mayor del estimado en otros estudios.

Estos resultados contradicen, en parte, los de algunos trabajos anteriores (por ejemplo, *Locatelli et col.* [11]), pero son consistentes con resultados de otros estudios [13], [27], entre ellos uno muy reciente [27], publicado el 21 de mayo de 2021.

- Analizando los datos de la primera ola de la pandemia para un grupo de 16 países europeos, seleccionados tanto por su población como por la incidencia de la enfermedad, se ha llegado a la conclusión de que, en general, las medidas más efectivas contra el COVID-19 han sido las relacionadas con el confinamiento, los conjuntos de medidas aplicadas simultáneamente en marzo de 2020 y el uso de mascarillas.
 - Teniendo en cuenta el periodo a partir del 1 de agosto de 2020, después de la primera ola de la pandemia, se identifican como medidas más eficaces las relacionadas con el confinamiento, la limitación de reuniones y contactos sociales y el cierre de centros educativos.
 - Para algunos países, entre ellos España, se observa un aumento de casos en el periodo navideño que no se explica por la desactivación de alguna de las medidas de intervención. Esto se ha interpretado como consecuencia de las reuniones familiares en el periodo festivo.
-

- La vacunación ha empezado a mostrar su efecto positivo en el descenso de casos y de muertes por *COVID-19*. Dado que la versión actual del modelo estadístico no considera este efecto, las predicciones sobre el desarrollo futuro de la enfermedad serán excesivamente pesimistas.
- Para el análisis de datos futuros, se hace necesaria la extensión del modelo para incluir factores adicionales que influyen en el desarrollo de la enfermedad, siendo el más relevante el efecto de la vacunación.

7.4 Limitaciones de este trabajo

El objetivo principal de este trabajo, que era eliminar una de las limitaciones estructurales del modelo original [6], permitiendo procesar un número potencialmente ilimitado de periodos de intervención, se ha cumplido.

A pesar de esta mejora, el modelo sigue teniendo algunas limitaciones, relacionadas con algunas simplificaciones o con la falta de información disponible. Éstas no impiden que los resultados obtenidos sean buenos, y que se hayan podido obtener conclusiones razonables sobre el efecto de las distintas medidas de intervención en periodos pasados. Sin embargo, son aspectos que deben tenerse en cuenta para trabajos futuros basados en la metodología de este *Trabajo Fin de Máster*.

Limitaciones:

- Se consideran los datos de *COVID-19* de forma global para cada país, a pesar de que en algunos países el desarrollo de la enfermedad y las medidas aplicadas han variado significativamente entre distintas regiones o provincias.
- Las distribuciones y los parámetros que describen la probabilidad de infección y de fallecimiento por *COVID-19* se definen de forma global o restringidos a un único país, pero no se tienen en cuenta posibles diferencias relacionadas con factores como pueden ser la edad o la distribución geográfica de la población.
- El modelo no es capaz de detectar el efecto de medidas preventivas en periodos de baja incidencia de la enfermedad
- Se asume que el efecto de una intervención permanece constante a lo largo del tiempo mientras ésta está activa.
- A la hora de hacer predicciones, se asume que el efecto de intervenciones futuras sobre el valor de R_t (*número de reproducción*) es el mismo que el observado en el pasado para las mismas intervenciones.
- No se considera el posible impacto de factores no relacionados con las medidas de intervención, pero que pueden tener un efecto significativo en el desarrollo de la enfermedad.

Algunos ejemplos son: la mejoría en la capacidad de detección de la enfermedad, las diferentes variantes del virus y, especialmente, el efecto de la vacunación.

- El modelo no tiene en cuenta el porcentaje de población inmunizada, bien por haber pasado la enfermedad, bien por estar vacunada, para calcular el *número de reproducción efectivo*.
- Si bien el modelo permite el procesamiento de un número arbitrario de intervenciones, no siempre es posible sacar conclusiones sobre el efecto de intervenciones individuales.
Esto, más que una limitación del modelo, es una restricción originada por la forma en que se han aplicado las medidas de intervención en los diferentes países.
- Por la naturaleza de la fuente de los datos relacionados con las medidas de intervención [22], el análisis, en la versión actual, se ha limitado a los países europeos para los que hay información en el *European Centre for Disease Prevention and Control* [22]
- El modelo no incluye funciones específicas para identificar y tratar datos anómalos (*outliers*) en los datos de casos y muertes por COVID-19.
- Se ha mostrado, con ejemplos para España y Alemania, que es posible utilizar el modelo para hacer ajustes de datos a nivel regional, pero no se ha identificado una fuente de datos que ofrezca información completa sobre las medidas de intervención aplicadas en cada una de las regiones, lo que limita mucho la capacidad de este tipo de análisis.
- En la configuración por defecto, siguiendo la metodología y los datos de trabajos previos [6], [7], se ha utilizado una distribución *prior* para el valor del *número de reproducción básico* común a todos los países. Los resultados muestran que este valor varía mucho entre países, por lo que sería más conveniente poder definir una distribución *prior* de forma individual para cada país.
En la versión actual ya es posible definir un valor predefinido para el *número de reproducción básico*, en lugar de estimar este valor durante el proceso de ajuste. Pero, sin hacer uso de esta opción de forma explícita, la distribución *prior* es común a todos los países y tiene un valor medio, obtenido de estudios previos, que, a la vista de los resultados de este trabajo, es demasiado pequeño.

7.5 Líneas de futuro

Teniendo en cuenta los buenos resultados obtenidos, se ha llegado a la conclusión de que la metodología utilizada, basada en el análisis bayesiano, es adecuada para describir la evolución del *COVID-19*.

A pesar de sus limitaciones, documentadas en el apartado anterior, se han obtenido buenos resultados de ajuste y conclusiones consistentes sobre el efecto de las distintas medidas de intervención en diferentes países.

Por tanto, se considera que tiene sentido utilizar este trabajo como base para trabajos futuros, tanto para el análisis de datos de *COVID-19* como para el análisis de otros fenómenos que sigan patrones de propagación similares (otras enfermedades, patógenos de plantas y animales, etc.).

La herramienta resultante de este trabajo se ha diseñado de forma modular, y es posible extenderla de forma sencilla para incluir fuentes de datos adicionales, otros factores que influyan en la evolución de la enfermedad, para utilizar un modelo estadístico distinto, etc.

En relación con el análisis de datos del *COVID-19*, las posibles líneas de futuro están relacionadas principalmente con la mejora de algunas de las simplificaciones y limitaciones del modelo y se describen en los siguientes puntos.

1. Calidad de los datos utilizados y obtención de datos adicionales

Como se ha comentado, la fuente de los **datos de infecciones y muertes por COVID-19** [24] incluye datos anómalos. Esto puede suponer, en algunos casos, que los resultados del análisis no sean del todo correctos o estén sesgados a causa de valores extremos en fechas concretas. Una posible mejora sería el análisis de datos anómalos (outliers) y su tratamiento.

Por otro lado, se ha utilizado como fuente para las **medidas de intervención** el *European Centre for Disease Prevention and Control* [22], donde se documenta información sobre las medidas aplicadas en 32 países europeos. Una posible mejora sería la inclusión de otras fuentes para posibilitar el análisis de datos de otros países o de datos a nivel regional.

Las distribuciones **Serial Interval (SI)** e **Infection-to-Death (ITD)** y el valor para la probabilidad media de fallecimiento de un infectado (**Infection Fatality Ratio (IFR)**) se han obtenido del estudio original [6] y, de acuerdo con su descripción, se basan en datos del inicio de la pandemia en China. Sería necesario revisar estos datos y comprobar si requieren una actualización.

2. Mejoras para corregir algunas de las simplificaciones y limitaciones del modelo

El apartado “7.4 Limitaciones de este trabajo” describe las limitaciones del modelo actual.

Futuros trabajos deberían eliminar algunas de las limitaciones metodológicas.

La mejora más necesaria es la opción de considerar factores adicionales, aparte de las medidas de intervención, en el modelo estadístico que describe la evolución de la enfermedad. De forma más específica, sería necesario incluir tener en cuenta la influencia de la vacunación para el cálculo del número de reproducción (R_t).

Otra mejora necesaria, relacionada con el punto anterior, es el considerar en el cálculo de R_t el porcentaje de población inmunizada, bien por haber pasado la enfermedad, bien por haber sido vacunada.

3. Extensión de las tareas de análisis a partir de los resultados generados

Con la creación de los modelos para cada uno de los países, la herramienta informática produce gran cantidad de datos, tanto gráficos como numéricos.

El análisis de la bondad del ajuste se ha hecho exclusivamente a partir de los gráficos, con criterios más o menos subjetivos. Sería posible definir criterios objetivos basados en los datos numéricos para la evaluación del ajuste de los distintos modelos. Una opción sería comparar, de forma automática, las predicciones del modelo con los datos reales y evaluar la bondad del ajuste, por ejemplo, en función del *RMSE* (*root mean square error*).

Asimismo, sería posible utilizar los resultados de los distintos modelos para hacer análisis más exhaustivos de forma numérica, por ejemplo en relación con el efecto de las distintas medidas o con la comparación de este efecto entre distintos países.

4. Extensión de la funcionalidad del servidor web [9]

Durante el progreso del trabajo, dada la gran cantidad de datos generados, se consideró adecuada la creación de un servidor web [9] para almacenar los distintos resultados y ofrecer un acceso a los mismos de forma sencilla e intuitiva.

Si bien ésta es una parte accesorio del trabajo, se trata de un componente de uso sencillo que podría extenderse para presentar resultados de trabajos futuros o, incluso, para permitir la interacción del usuario con la herramienta informática y posibilitar el análisis de datos y la creación de nuevos modelos de forma interactiva.

8 Glosario

Término / acrónimo	Descripción
<p><i>Análisis bayesiano</i> <i>Inferencia bayesiana</i></p>	<p>De forma muy simplificada, se puede decir que los métodos de inferencia bayesiana, como el que se utiliza en este trabajo, permiten sacar conclusiones, con la distribución <i>posterior</i> que se obtiene para ciertos parámetros, a partir de conocimientos o supuestos previos (<i>prior</i>) y de evidencias o datos conocidos [8], [23].</p>
<p>ARIMA Modelo de crecimiento logístico</p>	<p><i>Autoregressive integrated moving average</i> <i>Logistic growth model</i></p> <p>Dos métodos para el análisis de series temporales (<i>time series analysis</i>) distintos del utilizado en este Trabajo Fin de Máster. Se utilizan en uno de los trabajos mencionados en “3 Estado del arte” [18].</p>
<p><i>Coronavirus</i></p>	<p>Extensa familia de virus que pueden causar enfermedades tanto en animales como en humanos.</p> <p>En los humanos, se sabe que varios coronavirus causan infecciones respiratorias que pueden ir desde el resfriado común hasta enfermedades más graves [1].</p>
<p>COVID-19</p>	<p><i>Coronavirus disease 2019</i></p> <p>Enfermedad causada por un nuevo tipo de coronavirus llamado SARS-CoV-2 [1].</p> <p>Aunque hay estudios y organismos (por ejemplo, la OMS) que se refieren a esta enfermedad en femenino (<i>la COVID-19</i>), en este trabajo se ha utilizado el artículo masculino (<i>el COVID-19</i>).</p> <p>Ambas variantes son aceptadas por la <i>Real Academia Española</i> [30], por lo que se ha elegido la variante masculina al considerarla más natural.</p>

Término / acrónimo	Descripción
<i>Inmunidad de rebaño</i>	<p>También denominada inmunidad de grupo o colectiva.</p> <p>Situación en la que un porcentaje suficiente de la población de una comunidad es inmune a una enfermedad (por ejemplo, gracias a la vacunación), de forma que ésta está bajo control y deja de extenderse [28].</p> <p>Matemáticamente, esta situación se alcanza cuando la inmunización de parte de la población hace que el <i>número de reproducción efectivo</i> llegue a valores por debajo de 1 sin necesidad de la aplicación de medidas no farmacológicas (ver “4.1 Número de reproducción (R_t)”).</p>
<i>Intervenciones no farmacológicas</i> <i>Medidas de intervención</i>	En el contexto de este trabajo, medidas destinadas a prevenir y controlar la transmisión del COVID-19.
<i>Modelo mecanístico</i>	<p>En un modelo mecanístico, se asumen conocidas las leyes que describen un fenómeno.</p> <p>La finalidad del análisis estadístico es encontrar combinaciones de parámetros que hagan que, aplicando las leyes o fórmulas que describen el fenómeno, los resultados se adapten lo mejor posible a los datos conocidos [32].</p> <p>En este trabajo, se asume que la evolución del COVID-19 se puede describir a través de ciertas fórmulas matemáticas (ver “4.3 Fundamento estadístico - Modelo mecanístico”), y la finalidad es encontrar los valores de los parámetros desconocidos (básicamente, los que describen el efecto de las distintas medidas de intervención) con los que se consigue un mejor ajuste a los datos de muertes por COVID-19.</p>
<i>Muestreo de Monte Carlo Hamiltoniano</i>	<p>Método de simulación con el que se obtienen series de valores para ciertos parámetros que se adaptan a una distribución de probabilidad dada.</p> <p>En el contexto de este trabajo, este método se utiliza para la obtención de estimaciones de los parámetros que describen la evolución del COVID-19.</p>
<i>Ola</i>	En relación con la propagación de enfermedades, “número creciente de personas infectadas por una enfermedad, que alcanza un máximo para, a continuación, descender más lentamente.” [31]. Se recomienda su uso en lugar del término <i>oleada</i> [31].

Término / acrónimo	Descripción
OMS	Organización Mundial de la Salud
<i>prior</i> <i>posterior</i>	Conceptos de inferencia bayesiana, que definen las distribuciones iniciales supuestas de los parámetros a optimizar (<i>prior</i>) y las distribuciones resultantes del proceso de inferencia (<i>posterior</i>) [8], [23]
R	Lenguaje de programación para el cálculo estadístico [33]
Número de reproducción	<p>El <i>número de reproducción</i> (<i>reproduction number</i> en inglés) representa el promedio de casos secundarios que se producirán en un grupo de población no inmune a partir de un caso primario [11].</p> <p>Valores mayores de 1 indican una tendencia creciente en el número de casos, mientras que un descenso por debajo de 1 se asocia con la contención en la expansión de la pandemia [10], [11].</p>
SARS-CoV-2	<p>Nuevo tipo de coronavirus.</p> <p>La Organización Mundial de la Salud (OMS) tuvo conocimiento de su existencia a finales del año 2019, a partir de un informe sobre un grupo de casos de <i>neumonía viral</i> en Wuhan (China) [1].</p>
Stan	Lenguaje de programación para el desarrollo de modelos estadísticos [25].
SEIR SIDARTHE	<p>Dos métodos para el análisis de series temporales (<i>time series analysis</i>) distintos del utilizado en este Trabajo Fin de Máster.</p> <p>Se trata de dos ejemplos de los llamados <i>modelos compartimentales</i> (<i>compartmental models</i>).</p> <p>Se utilizan en dos de los trabajos mencionados en “3 Estado del arte” [13], [16].</p>

Tabla 15 - Glosario

9 Bibliografía

- [1] World Health Organization (2020). *Coronavirus Disease COVID-19 Q&As*
Versión en línea disponible en: <https://www.who.int/es/emergencies/diseases/novel-coronavirus-2019/question-and-answers-hub/q-a-detail/coronavirus-disease-covid-19>
(fecha de consulta: 25 de febrero de 2021).
- [2] World Health Organization (2020). *WHO Director-General's opening remarks at the media briefing on COVID-19 - 11 March 2020*
Versión en línea disponible en: <https://www.who.int/director-general/speeches/detail/who-director-general-s-opening-remarks-at-the-media-briefing-on-covid-19---11-march-2020>
(fecha de consulta: 25 de febrero de 2021).
- [3] World Health Organization (2021). *Weekly operational update on COVID-19 - 1 February 2021*
Versión en línea disponible en: <https://www.who.int/publications/m/item/weekly-operational-update-on-covid-19---1-february-2021>
(fecha de consulta: 25 de febrero de 2021).
- [4] World Health Organization (2021) *Weekly operational update on COVID-19 - 31 May 2021*
Versión en línea disponible en: <https://www.who.int/publications/m/item/weekly-operational-update-on-covid-19-31-may-2021>
(fecha de consulta: 1 de junio de 2021).
- [5] Organización Mundial de la Salud (2020) *Preguntas y respuestas sobre la transmisión de la COVID-19*
Versión en línea disponible en: <https://www.who.int/es/news-room/q-a-detail/coronavirus-disease-covid-19-how-is-it-transmitted>
(fecha de consulta: 16 de mayo de 2021).
- [6] Fernández-Recio, J. (2020) *Modelling the Evolution of COVID-19 in High-Incidence European Countries and Regions: Estimated Number of Infections and Impact of Past and Future Intervention Measures*. 2020 *J. Clin. Med.* 9, no. 6, 1825
- [7] Flaxman, S., Mishra, S., Gandy, A. et al. (2020) *Estimating the effects of non-pharmaceutical interventions on COVID-19 in Europe*. 2020 *Nature* 584, pp. 257-261
- [8] Kruschke, J. (2015) *Doing Bayesian Data Analysis*
Versión en línea disponible en: <https://sites.google.com/site/doingbayesiandataanalysis/>
(fecha de consulta: 20 de febrero de 2021).
-

-
- [9] Blecua, J. (2021) *COVID-19 transmission model - Master thesis*
Versión en línea disponible en <http://jblecua.spdns.org:9095/> (acceso restringido)
(fecha de consulta: 8 de junio de 2021).
- [10] Hethcote, H. W. (2000) *The mathematics of infectious diseases*. 2020 SIAM REVIEW Vol. 42, No. 4, pp. 599-653
- [11] Locatelli, I., Trächsel, B., Rousson, V.(2021) *Estimating the basic reproduction number for COVID-19 in Western Europe*. 2021 PLoS ONE 16(3): e0248731
- [12] Haug, N., Geyrhofer, L., Londei, A., Dervic, E., Desvars-Larrive, A., Loreto, V., Pinior, B., Thurner, S., Klimek, P. (2020) *Ranking the effectiveness of worldwide COVID-19 government interventions*. 2020 Nat Hum Behav 4, pp. 1303-1312
- [13] Linka, K., Peirlinck, M., Kuhl, E. (2020) *The reproduction number of COVID-19 and its correlation with public health interventions* [published online ahead of print, 2020 Jul 28]. 2020 Comput Mech.;1-16. doi:10.1007/s00466-020-01880-8
- [14] Getz, W. M., Salter, R., Muellerklein, O., Yoon, H.S., Tallam, K. (2018) *Modeling Epidemics: A Primer and Numerus Software Implementation*. 2018 Epidemics, Volume 25, pp. 9-19
- [15] Català, M., Pino, D., Marchena, M., Palacios, P., Urdiales, T., Cardona, P.J., Alonso, S., López-Codina, D., Prats, C., Alvarez-Lacalle, E. (2021) *Robust estimation of diagnostic rate and real incidence of COVID-19 for European policymakers*. 2021 PLoS ONE 16(1): e0243701
- [16] Giordano, G., Blanchini, F., Bruno, R., Colaneri, P., Di Filippo, A., Di Matteo, A., Colaneri, M. (2020) *Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy*. 2020 Nat Med 26, pp. 855-860
- [17] Brauner, J.M., Mindermann, S., Sharma, M., Johnston, D., Salvatier, J., Gavenčiak, T., Stephenson, A.B., Leech, G., Mikulik, V., Norman, A.J., Teperowski Monrad, J., Besiroglu, T., Ge, H., Hartwick, M.A., The, Y.W., Chindelevitch, L., Gal, Y., Kulveit, J. (2021) *Inferring the effectiveness of government interventions against COVID-19*. 2021 Science Vol. 371, Issue 6531, eabd9338
- [18] Elhassan, T., Gaafar, A. (2020) *Mathematical modeling of the COVID-19 prevalence in Saudi Arabia*. 2020 medRxiv: n. pag.
Versión en línea disponible en: <https://doi.org/10.1101/2020.06.25.20138602>
-

-
- [19] Bendavid, E., Oh, C., Bhattacharya, J., Ioannidis, J.P.A (2021) *Assessing mandatory stay-at-home and business closure effects on the spread of COVID-19*. 2021 *European Journal of Clinical Investigation* Vol. 54, no. 4: e13484
- [20] Bryant, P., Elofsson, A (2020) *The limits of estimating COVID-19 intervention effects using Bayesian models*. 2020 medRxiv: n. pag.
Versión en línea disponible en: <https://doi.org/10.1101/2020.08.14.20175240>
- [21] Pozo-Martin, F., Cristea, F., El Bcheraoui Robert, C. - Koch-Institut, Zentrum für Internationalen Gesundheitsschutz (ZIG) (2020) *Rapid Review der Wirksamkeit nicht-pharmazeutischer Interventionen bei der Kontrolle der COVID-19-Pandemie*
Versión en línea disponible en:
https://www.rki.de/DE/Content/InfAZ/N/Neuartiges_Coronavirus/Projekte_RKI/Rapid-Review-NPIs.pdf?__blob=publicationFile
(fecha de consulta: 19 de mayo de 2021).
- [22] European Centre for Disease Prevention and Control (2021) *Data on country response measures to COVID-19*
Versión en línea disponible en: <https://www.ecdc.europa.eu/en/publications-data/download-data-response-measures-covid-19>
(fecha de consulta: 19 de abril de 2021).
- [23] Gelman, A., Carlin, J.B., Stern, H.S., Dunson, D.B., Vehtari, A., Rubin, D.B. (2020) *Bayesian Data Analysis (third edition)*
Versión en línea disponible en: <http://www.stat.columbia.edu/~gelman/book/>
(fecha de consulta: 1 de abril de 2021).
- [24] Johns Hopkins University (2021) *COVID-19 Data Repository by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University*
Versión en línea disponible en: https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series
(fecha de consulta: 19 de abril de 2021).
- [25] Stan Development Team (2018) *Stan User's guide*
Versión en línea disponible en: https://mc-stan.org/docs/2_18/stan-users-guide/index.html
(fecha de consulta: 20 de febrero de 2021).
-

[26] *Información del gobierno alemán sobre nuevas medidas a partir del 24 de abril de 2021*

Versión en línea disponible en: <https://www.bundesregierung.de/breg-de/suche/bundesweite-notbremse-1888982>

(fecha de consulta: 10 de mayo de 2021).

[27] Ke, R., Romero-Severson, E., Sanche, E., Hengartner, R. (2021) *Estimating the reproductive number R_0 of SARS-CoV-2 in the United States and eight European countries and implications for vaccination*. 2021 *Journal of Theoretical Biology* Vol. 517, 110621

[28] *Johns Hopkins Bloomberg School of Public Health (2021) What is Herd Immunity and How Can We Achieve It With COVID-19?*

Versión en línea disponible en: <https://www.jhsph.edu/covid-19/articles/achieving-herd-immunity-with-covid19.html>

(fecha de consulta: 26 de mayo de 2021).

[29] *worldometer (2021) Countries in the world by population (2021)*

Versión en línea disponible en: <https://www.worldometers.info/world-population/population-by-country/>

(fecha de consulta: 26 de mayo de 2021).

[30] *Real Academia Española (2021) Dudas rápidas*

Versión en línea disponible en: <https://www.rae.es/duda-linguistica/es-el-covid-19-o-la-covid-19>

(fecha de consulta: 27 de mayo de 2021).

[31] *Fundéu RAE (2021) Recomendaciones*

Versión en línea disponible en: <https://www.fundeu.es/recomendacion/brote-y-rebrote-ola-y-oleada-matices/>

(fecha de consulta: 5 de junio de 2021).

[32] *GoSilico (2021) Mechanistic vs. statistical models*

Versión en línea disponible en: <https://gosilico.com/technology/mechanistic-vs-statistical-models/>

(fecha de consulta: 27 de mayo de 2021).

[33] *The R Project for Statistical Computing*

Versión en línea disponible en: <https://www.r-project.org/>

(fecha de consulta: 27 de mayo de 2021).

10 Anexos

10.1 Anexo I - Nombres de las medidas de intervención

La información sobre las medidas de intervención aplicadas en los distintos países se ha obtenido del [European Centre for Disease Prevention and Control \[22\]](#).

Los nombres de las medidas individuales están documentados en fuente de los datos: [Data dictionary non-pharmaceutical measures \[22\]](#), y se resumen en la siguiente tabla:

Nombre de la medida	Significado
<i>StayHomeOrder</i>	Confinamiento obligatorio para toda la población (<i>lockdown</i>)
<i>RegionalStayHomeOrder</i>	Confinamiento obligatorio para toda la población (<i>lockdown</i>) en al menos una de las regiones del país
<i>StayHomeGen</i>	Recomendación de confinamiento para toda la población
<i>StayHomeRiskG</i>	Recomendación de confinamiento para grupos de riesgo y población vulnerable (por ejemplo, gente mayor o personas con enfermedades previas)
<i>SocialCircle</i>	Limitación de contactos sociales, por ejemplo a un número limitado de hogares
<i>PrivateGatheringRestrictions</i>	Restricciones para reuniones privadas
<i>ClosDaycare</i>	Cierre de centros educativos: guarderías
<i>ClosPrim</i>	Cierre de centros educativos: escuelas de enseñanza primaria
<i>ClosSec</i>	Cierre de centros educativos: escuelas de enseñanza secundaria
<i>ClosHigh</i>	Cierre de centros educativos: instituciones de enseñanza superior
<i>MassGatherAll</i>	Limitación de las concentraciones masivas/públicas (se incluye cualquier intervención sobre las concentraciones masivas de hasta 1000 participantes)
<i>BanOnAllEvents</i>	Limitación de todas las reuniones masivas/públicas en interiores/exteriores
<i>IndoorOverNNN</i>	Limitación de las reuniones masivas o públicas en interiores de más de <i>NNN</i> participantes.
<i>OutdoorOverNNN</i>	Limitación de las concentraciones masivas o públicas al aire libre de más de <i>NNN</i> participantes
<i>ClosPubAny</i>	Cierre de espacios públicos de cualquier tipo (incluyendo restaurantes, lugares de ocio, tiendas no esenciales, cierre parcial o total del transporte público, gimnasios y centros deportivos, etc.)

Nombre de la medida	Significado
<i>EntertainmentVenues</i>	Cierre de locales de ocio
<i>ClosureOfPublicTransport</i>	Cierre del transporte público
<i>GymsSportsCentres</i>	Cierre de gimnasios y centros deportivos
<i>HotelsAccommodation</i>	Cierre de hoteles y otros servicios de alojamiento
<i>NonEssentialShops</i>	Cierre de tiendas no esenciales
<i>PlaceOfWorship</i>	Cierre de lugares de culto
<i>RestaurantsCafes</i>	Cierre de restaurantes y cafés/bares
<i>MasksVoluntaryAllSpaces</i>	Uso voluntario de mascarillas en todos los espacios públicos
<i>MasksVoluntaryClosedSpaces</i>	Uso voluntario de mascarillas en todos los espacios cerrados y en el transporte público
<i>MasksMandatoryAllSpaces</i>	Uso obligatorio de mascarillas en todos los lugares
<i>MasksMandatoryClosedSpaces</i>	Uso obligatorio de mascarillas en todos los espacios cerrados y en el transporte público
<i>Teleworking</i>	Recomendación de teletrabajo
<i>AdaptationOfWorkplace</i>	Adaptación de los lugares de trabajo (por ejemplo, para reducir el riesgo de contagio)
<i>WorkplaceClosures</i>	Cierre de centros de trabajo

Tabla 16 - Significado de los nombres de las distintas medidas de intervención

En algunos periodos, a los nombres de las medidas de la tabla anterior se les añade el sufijo *Partial* (por ejemplo, *StayHomeOrderPartial*), lo que indica que la aplicación de la medida ha sido relajada o menos estricta (*partially relaxed measure*).

Para algunos de los análisis, conjuntos de medidas que han estado activas simultáneamente se han agrupado en lo que se ha documentado como *major interventions*.

Para las comparaciones del efecto de distintas medidas entre distintos países, se han agrupado algunas de las medidas de significado similar en categorías genéricas para obtener resultados más compactos. Por ejemplo, todas las medidas relacionadas con el confinamiento (*StayHomeOrder*, *StayHomeGen*, *StayHomeRiskG*, ...) se agrupan bajo el nombre genérico *StayHome*.

10.2 Anexo II - Análisis de los datos a nivel regional para España y Alemania

Si bien no estaba previsto inicialmente, como complemento a los análisis de los distintos países, se han analizado también los datos de cada región de España y Alemania por separado.

Los datos de infecciones y muertes por *COVID-19* a nivel regional están disponibles en la [Johns Hopkins University \[24\]](#). En el caso de las regiones españolas y alemanas, los primeros datos son del 14.05.2020. Este análisis incluye, por tanto, datos desde esta fecha hasta el día en que se descargaron los datos, el 09.04.2021.

Se han utilizado, para cada región, las medidas de intervención del país correspondiente, ya que no se ha encontrado información completa sobre las medidas aplicadas en cada una de las regiones. A pesar de esta simplificación, se puede decir que los ajustes para las regiones son bastante buenos, si bien es necesario matizar esta afirmación.

Para algunas regiones, especialmente en el caso de España, el ajuste no es del todo bueno. Esto tiene varias causas:

- El uso de las medidas de todo el país puede no ser adecuado para todas las regiones. Esto es más relevante en el caso de España, donde ha habido diferencias significativas entre las distintas regiones en relación con las medidas aplicadas.
- En algunas regiones, el número de casos y de muertes por *COVID-19* es muy reducido, lo que dificulta el ajuste.
- Por la forma en que se mantienen los datos de *COVID-19* en la fuente utilizada [24], hay días que incluyen valores anómalamente altos o bajos de infecciones y muertes. Esto se debe a la forma en que se aplican algunas correcciones sobre datos acumulados de fechas pasadas. No se ha invertido tiempo en el tratamiento de *outliers*, por lo que algunos de los ajustes se desvían bastante de los datos reales.

El ajuste para las regiones alemanas es, en general, mejor. Esto es consistente con el hecho de que en Alemania, si bien ha habido diferencias a nivel regional, las medidas principales se han aplicado de forma homogénea en todo el país. Por eso, el uso de las medidas nacionales para el ajuste de los datos regionales es una buena aproximación.

Los modelos para todas las regiones de los dos países se han generado utilizando, en todos los casos, la configuración por defecto. Es posible que, en algunos casos, el ajuste mejore modificando algunos de los parámetros, pero no se ha invertido tiempo en comprobarlo.

La siguiente página muestra algunos ejemplos, tanto de ajustes buenos como de algunos menos buenos por algunas de las causas descritas anteriormente.

A través de los siguientes enlaces se accede a todos los resultados:

- Análisis de los datos de las regiones españolas: [🔗](#)
- Análisis de los datos de las regiones alemanas (*Bundesländer*): [🔗](#)

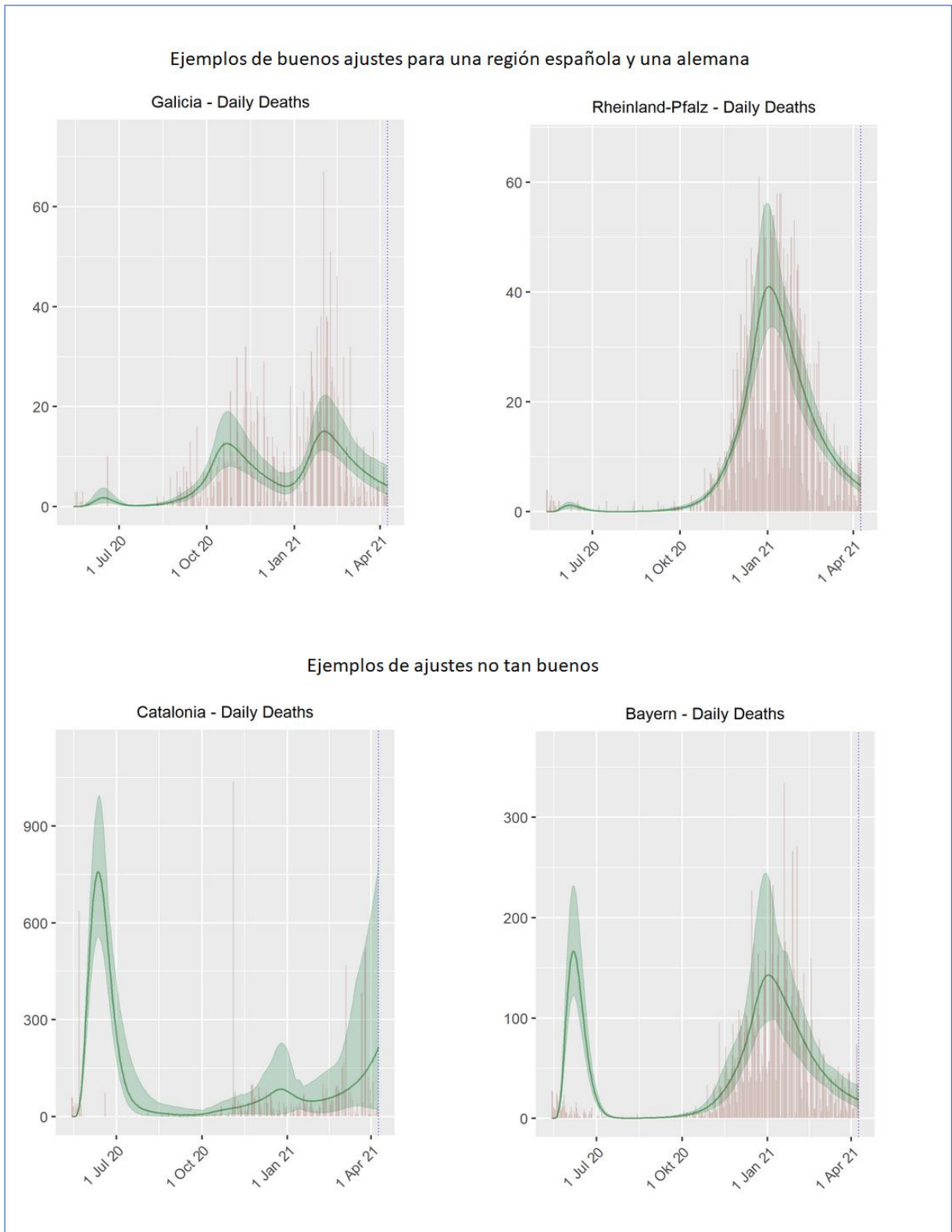


Figura 23 - Resultados del análisis para regiones españolas y alemanas

10.3 Anexo III - Ejemplos de los resultados de predicciones hasta el 18.04.2021

Este anexo completa la información del apartado “5.6.1 Modelos con datos hasta el 18.02.2021 / Predicciones hasta el 18.04.2021” con algunos ejemplos adicionales que se comentan brevemente.

Predicciones para Austria y Bélgica

En ambos casos, el gráfico superior representa los datos reales y el recuadro verde del gráfico inferior las predicciones *futuras* obtenidas a partir del modelo creado con datos hasta el 18.02.2021 que, como se puede ver, se adaptan bien a la tendencia de los datos reales.

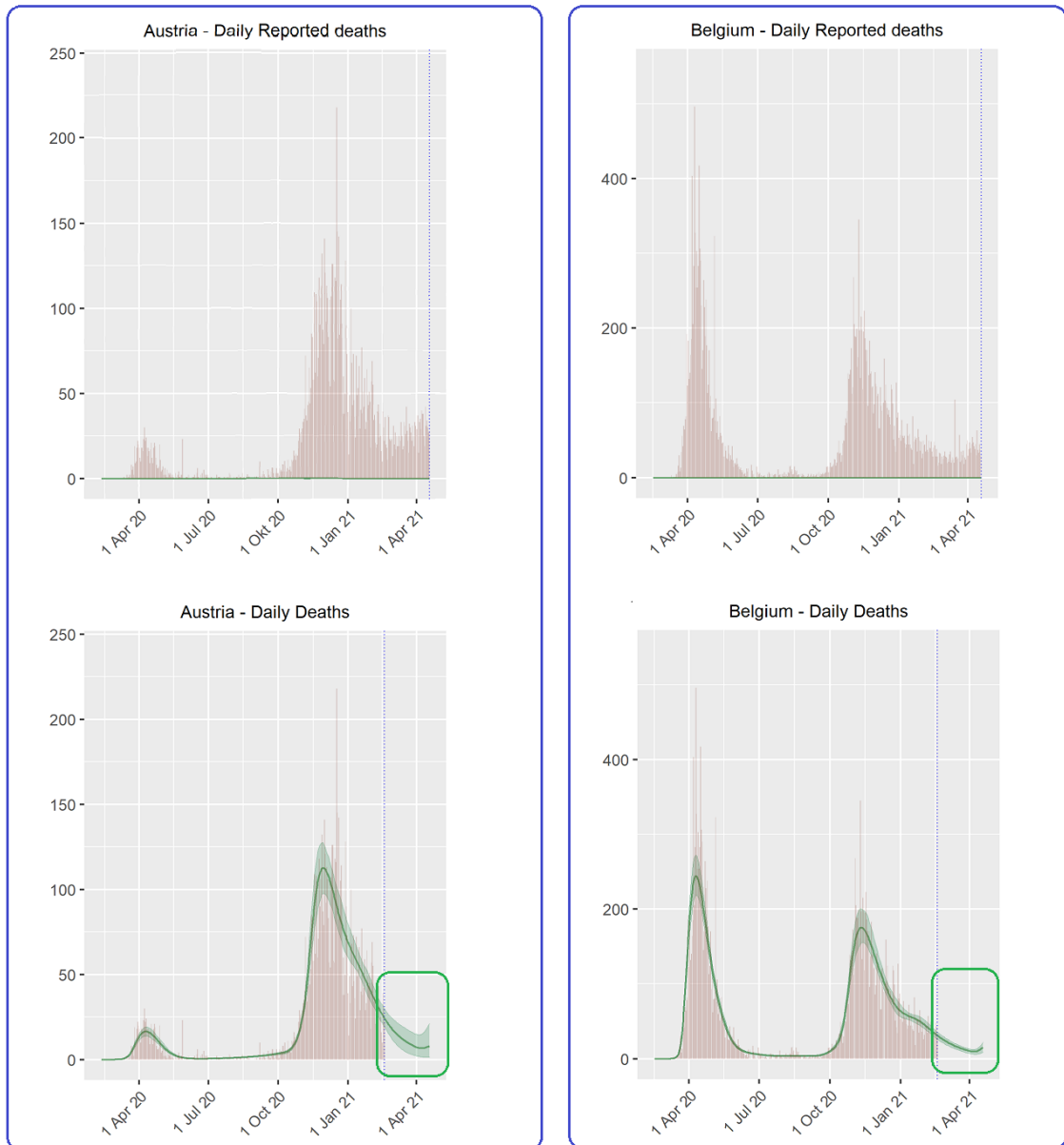


Figura 24 - Predicciones hasta el 18.04.2021 - Austria, Bélgica

Predicciones para Croacia y República Checa

En el caso de Croacia, las predicciones siguen también la tendencia de los datos reales, aunque el aumento real de muertes es más pronunciado que el reflejado en las predicciones.

En el caso de República Checa, las predicciones no son tan buenas, al no reflejar un pico claramente visible en los datos reales, si bien la tendencia final de descenso del número de muertes queda reflejada correctamente.

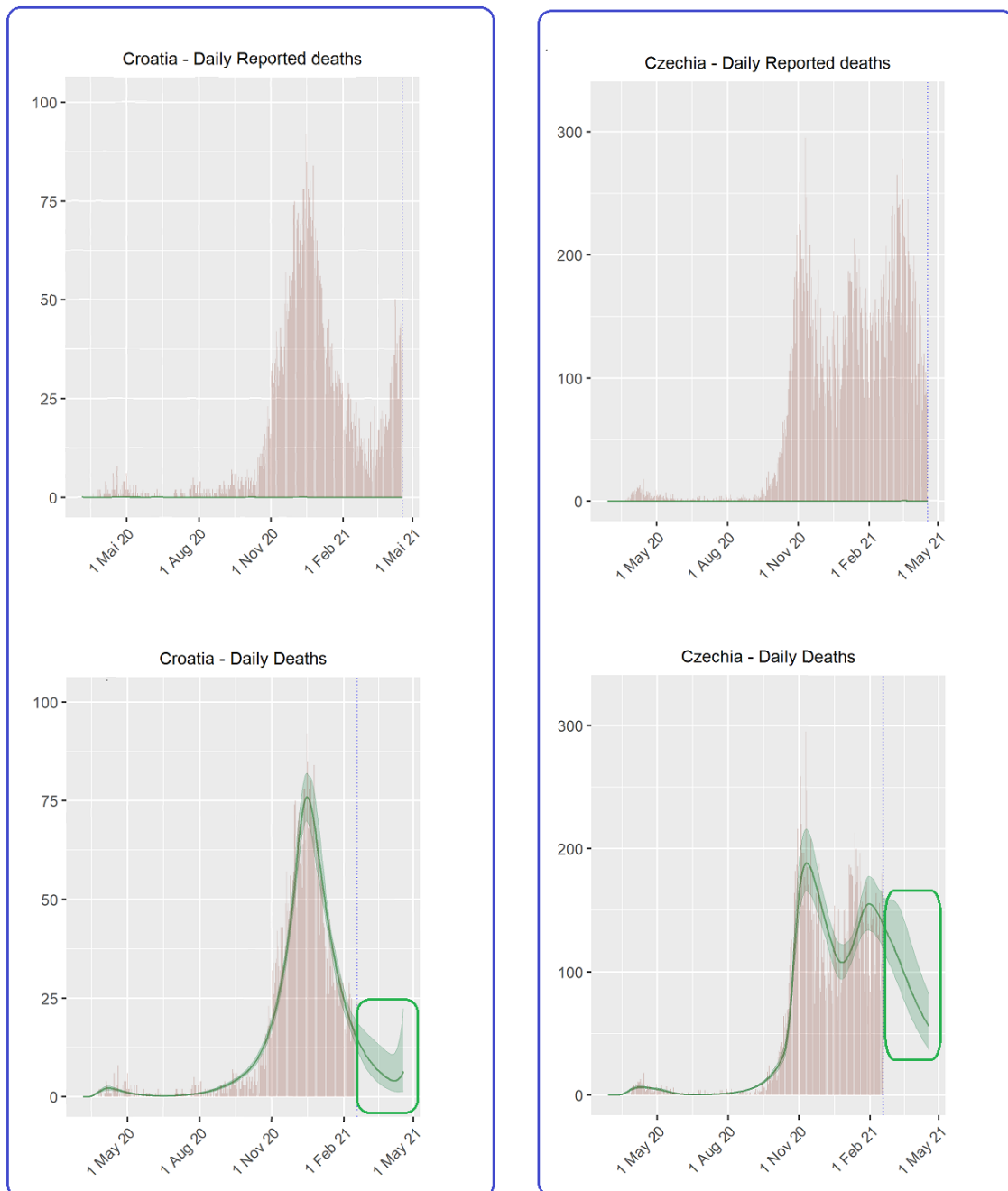


Figura 25 - Predicciones hasta el 18.04.2021 - Croacia, República Checa

Predicciones para Francia e Italia

Los valores medios de las predicciones para Francia se adaptan razonablemente bien a los datos reales, pero la anchura de la franja del 95% muestra un alto grado de incertidumbre.

Para Italia, las predicciones no reflejan el ascenso de casos que muestran los casos reales. Esto puede deberse a una elección poco adecuada de los parámetros relacionados con las medidas a partir del 18.02.2021.

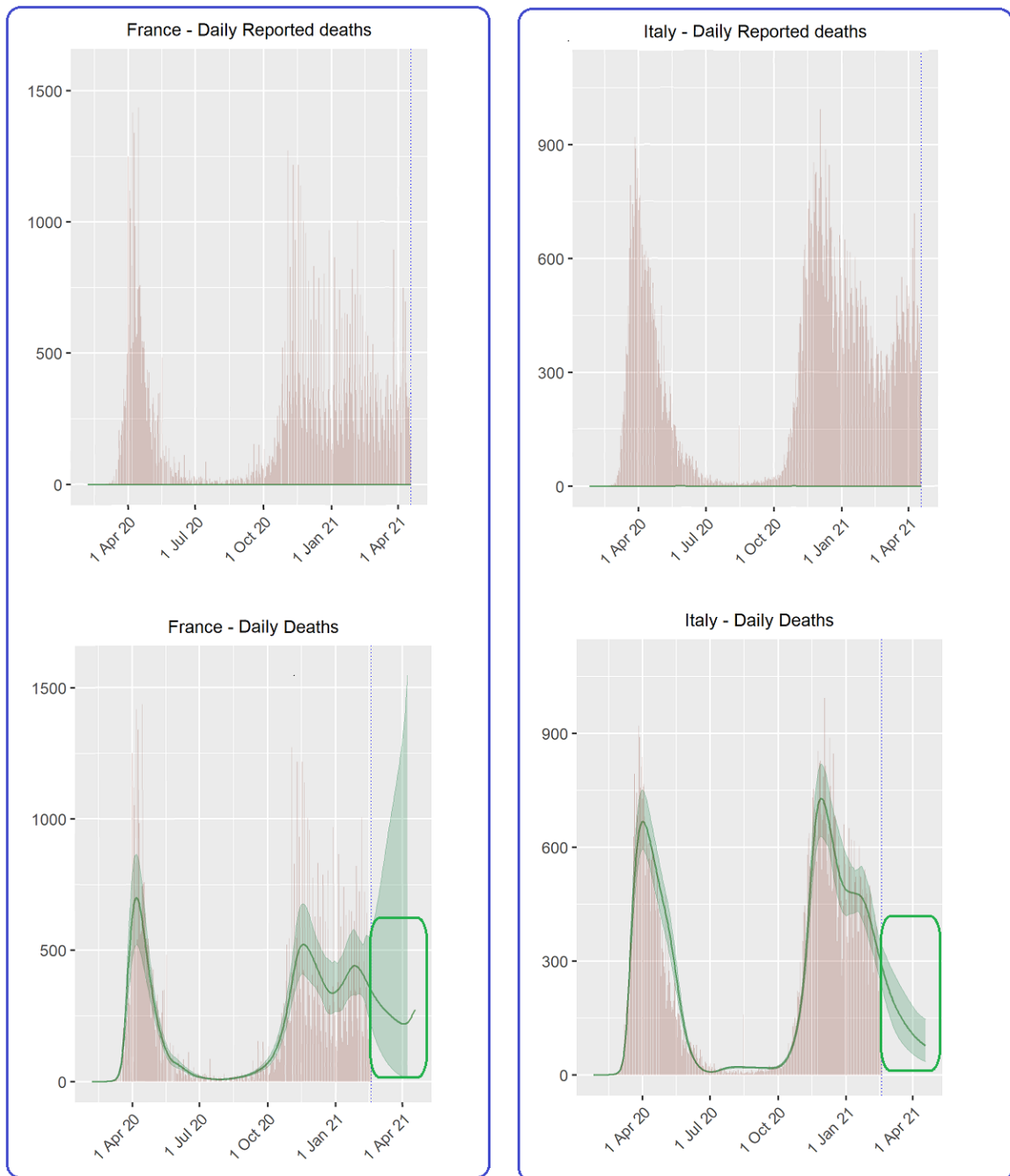


Figura 26 - Predicciones hasta el 18.04.2021 - Francia, Italia

Predicciones para España y Alemania

En ambos casos, los valores medios de las predicciones reflejan razonablemente bien la tendencia de los datos reales, si bien con un alto grado de incertidumbre en la franja de valores del 95%.

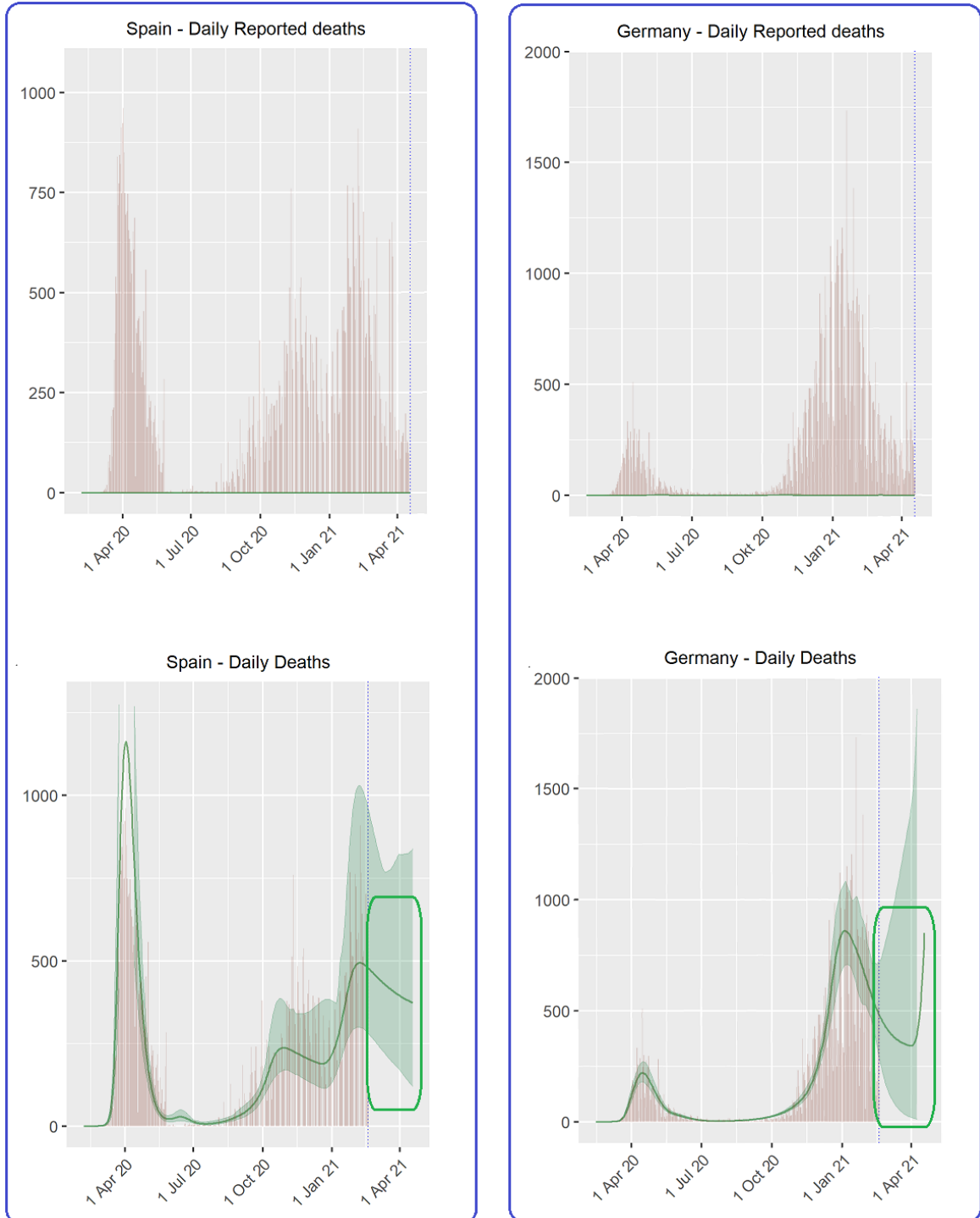


Figura 27 - Predicciones hasta el 18.04.2021 - España, Alemania

10.4 Anexo IV - Ejemplos de resultados de ajuste poco satisfactorios

El apartado “5.3 Evaluación del modelo - Bondad del ajuste” documenta los resultados obtenidos para una gran cantidad de modelos generados con datos que siguen patrones muy distintos. Se llega a la conclusión de que, salvo unas pocas excepciones, los resultados de los ajustes son muy buenos.

Este anexo se incluye para mostrar ejemplos de algunos de los resultados menos satisfactorios. Esto se hace para dejar constancia de posibles debilidades del modelo en ciertas situaciones, y de la necesidad de utilizar datos consistentes y una parametrización adecuada si se pretende sacar conclusiones de los resultados obtenidos.

Las causas para desviaciones en los ajustes son, entre otras:

- Datos incompletos o con valores anómalos (*outliers*)
- Factores no relacionados con la aplicación de medidas de intervención, pero que influyen de forma significativa en la evolución de la enfermedad.
Ejemplo: la vacunación, cuyo efecto no está incluido en la versión actual del modelo
- Parametrización poco conveniente para los datos objeto de análisis.
Ejemplo: número insuficiente de iteraciones para el *muestreo de Monte Carlo Hamiltoniano*.

Estos dos gráficos muestran resultados de ajuste poco satisfactorios. Otros ejemplos pueden encontrarse en [este enlace](#).

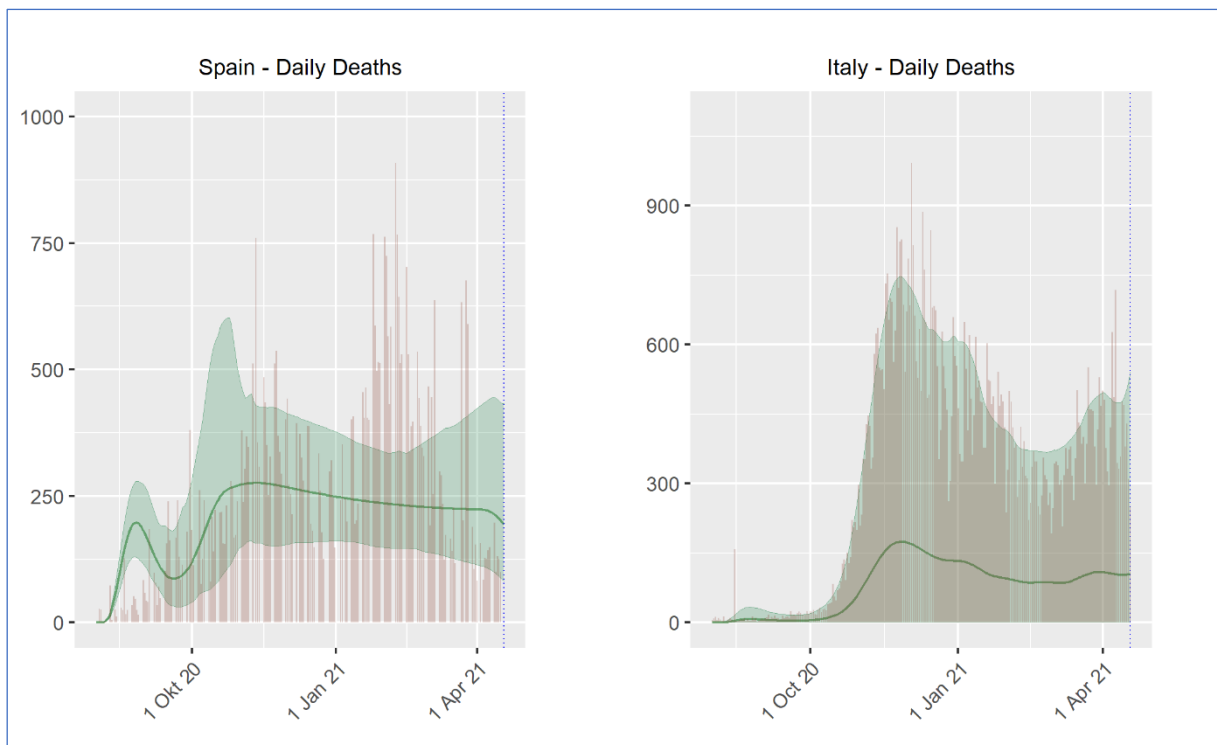


Figura 28 - Ejemplos de ajuste poco satisfactorio - España, Italia

10.5 Anexo V - Enlaces para el acceso a todos los resultados

Todos los resultados de los modelos generados, del análisis del efecto de las distintas medidas así como las predicciones obtenidas están disponibles en un servidor *web* [9]:

<http://jblecua.spdns.org:9095/> (usuario: *tfm* / password: *tfm*).

En distintas partes de este documento se han mencionado enlaces concretos.

La siguientes tablas resumen las partes principales del proceso de análisis y proporcionan los enlaces para acceder a los resultados correspondientes:

Modelos generados a partir de datos de COVID-19










	Datos de COVID-19		Intervenciones	Número de modelos resultantes	Resultados
	Países / Regiones	Rango de fechas			
1	Países europeos	22.01.2020 - 18.04.2021	Reducidas	32	
2	Países europeos	22.01.2020 - 18.02.2021	Reducidas	32	
3	Países europeos	22.01.2020 - 31.07.2020	Reducidas	32	
4	Países europeos	01.08.2020 - 18.04.2021	Reducidas	32	
5	Países europeos	22.01.2020 - 18.02.2021	Todas	32	
6	Países europeos	22.01.2020 - 31.07.2020	Todas	32	
7	Países europeos	01.08.2020 - 18.04.2021	Todas	32	
8	Regiones españolas	14.05.2020 - 09.04.2021	Reducidas	19	
9	Regiones alemanas	14.05.2020 - 09.04.2021	Reducidas	16	

Tabla 17 - Enlaces al servidor web - modelos

Análisis del efecto de las medidas de intervención en distintos países y regiones

	Datos de COVID-19		Intervenciones	Número de modelos resultantes	Resultados	
	Países / Regiones	Rango de fechas				
1	Países europeos	22.01.2020 - 18.04.2021	Reducidas	32		
2	Países europeos	22.01.2020 - 18.02.2021	Reducidas	32		
3	Países europeos	22.01.2020 - 31.07.2020	Reducidas	32		
4	Países europeos	01.08.2020 - 18.04.2021	Reducidas	32		
8	Regiones españolas	14.05.2020 - 09.04.2021	Reducidas	19		
9	Regiones alemanas	14.05.2020 - 09.04.2021	Reducidas	16		

Tabla 18 - Enlaces al servidor web - efecto de las medidas de intervención

Comparación del efecto de las principales medidas de intervención entre distintos países

	Datos de COVID-19		Intervenciones	Número de países comparados	Resultados	
	Países / Regiones	Rango de fechas				
1	Países europeos	22.01.2020 - 18.04.2021	Reducidas	16		
2	Países europeos	22.01.2020 - 18.02.2021	Reducidas	16		
3	Países europeos	22.01.2020 - 31.07.2020	Reducidas	16		
4	Países europeos	01.08.2020 - 18.04.2021	Reducidas	16		

Tabla 19 - Enlaces al servidor web - comparación del efecto de las medidas de intervención

Predicciones

	Datos de COVID-19		Intervenciones	Predicciones hasta	Número de países	Resultados
	Países / Regiones	Rango de fechas				
1	Países europeos	22.01.2020 - 18.02.2021	Reducidas	18.04.2021	11	
2	Países europeos	22.01.2020 - 18.04.2021	Reducidas	18.06.2021	16	

Tabla 20 - Enlaces al servidor web - predicciones

10.6 Anexo VI - Herramienta informática - aspectos técnicos

El software está escrito en el lenguaje de programación *R* [33].

Se utiliza, además, el lenguaje probabilístico *Stan* [25] para la implementación del modelo mecánico que describe la evolución del número de infecciones y de muertes (ver detalles en “4.3 Fundamento estadístico - Modelo mecánico”), y que permite obtener las estimaciones de los parámetros a través de un *muestreo de Monte Carlo Hamiltoniano* [8], [23].

Toda la información sobre el *Trabajo Fin de Máster* está disponible en un servidor *web* [9]: <http://jblecua.spdns.org:9095/> (usuario: *tfm* / password: *tfm*).

Información sobre el entorno de trabajo para el desarrollo de la herramienta informática:

Componente	Información
Sistema operativo	<i>Windows 10</i>
Lenguajes de programación	<i>R, Stan</i>
Entorno de desarrollo (<i>IDE</i>)	<i>RStudio (versión 1.4.1103)</i>
Instalación de <i>R</i>	<p><i>Versión 4.0.3</i> con las siguientes librerías adicionales, entre las que se encuentra la implementación de <i>Stan</i> para <i>R</i> (<i>rstan</i>):</p> <ul style="list-style-type: none"> <i>bayesplot</i> <i>cowplot</i> <i>data.table</i> <i>dplyr</i> <i>EnvStats</i> <i>gdata</i> <i>ggplot2</i> <i>ggpubr</i> <i>grid</i> <i>gridExtra</i> <i>htmlwidgets</i> <i>lubridate</i> <i>matrixStats</i> <i>png</i> <i>rstan</i> <i>scales</i> <i>tidyr</i> <i>webshot</i> <i>wordcloud2</i>
Documentación de los resultados	Servidor http://jblecua.spdns.org:9095/
Sistema operativo	<i>Raspberry Pi OS (Linux)</i>
Servidor <i>http</i>	<i>Apache httpd (versión 2.4.38)</i>
Lenguajes de programación	<i>php, html</i>

Tabla 21 - Herramienta informática - aspectos técnicos