

Las matemáticas en los tiempos del Covid-19

Autoría: Juan Luis Fernández-Martínez; Zulima Fernández Muñiz; José María Loché

Afiliación: Catedrático de Matemática Aplicada; Grupo de Problemas Inversos, Optimización y Aprendizaje Automático; Departamento de Matemáticas, Universidad de Oviedo | Profesora Titular de Matemática Aplicada; Grupo de Problemas Inversos, Optimización y Aprendizaje Automático; Departamento de Matemáticas, Universidad de Oviedo | Máster en Ingeniería de Telecomunicaciones; Grupo de Problemas Inversos, Optimización y Aprendizaje Automático; Departamento de Matemáticas, Universidad de Oviedo

TIPO	PÁGINAS	IDENTIFICADOR
Artículo de modelización matemática	41-57	TDL-77-2021-041-057

RESUMEN DOCUMENTAL

Artículo sobre el uso de modelos matemáticos para describir y predecir la evolución de la pandemia. Los autores explican problemas directos e inversos, estimación de parámetros, análisis de incertidumbre y combinación de modelos de corto y largo plazo. El objetivo es apoyar decisiones de salud pública y anticipar necesidades sanitarias como ingresos hospitalarios, pacientes críticos y unidades de cuidados intensivos.

PALABRAS CLAVE

COVID-19 · matemáticas · modelización · predicción · problemas inversos · incertidumbre · salud pública

CITA BIBLIOGRÁFICA

Juan Luis Fernández-Martínez; Zulima Fernández Muñiz; José María Loché. «Las matemáticas en los tiempos del Covid-19». Torre de los Lujanes, n.º 77, 2021, pp. 41-57. ISSN 1136-4343.

Las matemáticas en los tiempos del Covid-19

Juan Luis Fernández-Martínez
Catedrático de Matemática Aplicada
Director del Grupo de Problemas Inversos,
Optimización y Aprendizaje Automático.
Email: jlfm@uniovi.es
Departamento de Matemáticas.
Universidad de Oviedo

Zulima Fernández Muñiz,
Profesora Titular de Matemática Aplicada
Grupo de Problemas Inversos, Optimización
y Aprendizaje Automático. Departamento
de Matemáticas. Universidad de Oviedo.

José María Loché
Máster en Ingeniería de Telecomunicaciones
Grupo de Problemas Inversos, Optimiza-
ción y Aprendizaje Automático.
Departamento de Matemáticas.
Universidad de Oviedo.

Introducción

Según el informe titulado *Impacto socioeconómico de la investigación y la tecnología matemáticas en España* elaborado por Analistas Financieros Internacionales (AFI), las matemáticas aportan valor añadido a las empresas y representaron alrededor del 6% de empleo y del 10% del PIB en 2016. Según este informe, la informática, las actividades financieras, las telecomunicaciones y la energía son las áreas que obtienen beneficios directos, pero no podemos olvidarnos de la optimización de tareas y, en los últimos tiempos, de la biotecnología. En este último sector, la pandemia debida al coronavirus ayudó a demostrar, y a mostrar al gran público, que los modelos matemáticos son capaces de predecir su comportamiento en cuanto al número de infectados, de muertes y de ocupación de UCI, con la consiguiente ayuda para la organización de los recursos sanitarios.

Durante la crisis del Covid19 han salido a la luz muchos informes y estadísticas de muy diversa índole, tanto económicas como sanitarias o medioambientales. Gracias a las matemáticas se sabe que la Covid19 afecta más a familias sin ingresos, se puede hacer una estimación del tiempo que pasa desde que un enfermo es diagnosticado hasta que tienes que ser hospitalizado, ingresado en UCI, o incluso muere, con el fin de organizar turnos de campaña, dimensionar UCIs o llevar material sanitario donde más se necesita. Asimismo, también se investigó la relación de la pandemia con factores climáticos como la temperatura y la humedad y se dieron a conocer las predicciones de las distintas olas, así como las medidas sanitarias que se llevaron a cabo basadas en ese conocimiento.

Según el profesor Stephan Sturm de Worcester Polytechnic Institute, encerrar a la población o dejarla salir un día antes o después, influye notablemente en la expansión del virus. Para Tomoyuki Ichiba,

de la Universidad de California Santa Bárbara, es esencial combinar y comparar diferentes variables, integrando y equilibrando aspectos sanitarios y financieros, con el fin de plantear una mejor estrategia. La profesora Isabel Figuerola-Ferretti, de la Universidad Pontificia Comillas, expone que las bases de datos disponibles permiten combinar indicadores estructurales, como son la inflación o la tasa de desempleo, con otros coyunturales como búsquedas en Google o compras por internet, que se pueden ver prácticamente en tiempo real.

Los medios de comunicación se hicieron eco, durante este período, de los avances científicos y tecnológicos puestos al servicio de la sociedad con el fin de evitar los efectos de la pandemia, y muchos médicos, virólogos, inmunólogos y matemáticos pudieron exponer los avances de sus investigaciones, hasta ahora reservadas al ámbito académico. En todos los medios de comunicación aparecieron frases como *aplanar la curva*, *crecimiento exponencial*, *tasa de contagio y mortalidad*, *porcentaje de hospitalización*, *macrófagos*, *linfocitos*, *antígenos*, *vacunas ARNm*, *vacunas de subunidades proteicas*, *vacunas de vectores*, etc., que contribuyeron a normalizar el lenguaje científico.

Un poco de historia

Las matemáticas en las que hoy se fundamentan la modelización de enfermedades infecciosas se forjaron hace siglos con la aparición del cálculo diferencial. En los siglos XIV y XVII Europa se vio asolada por la peste negra y la peste de Londres, descrita por Daniel Defoe en su Diario de la peste, respectivamente. En el siglo XIV existió en Inglaterra un grupo vinculado al Merton College de Oxford conocido como los calculadores de Merton. Entre ellos se encontraba Thomas Bradwardine, posteriormente arzobispo de Canterbury, quien anticipó la idea de logaritmo y de función expo-

nencial tres siglos antes de que la desarrollaran algunos miembros de la familia Bernoulli y Leonard Euler.

En el siglo XVII, coincidiendo con la peste de Londres y el éxodo de la gente de las ciudades al campo, fue cuando Isaac Newton abandonó la universidad de Cambridge para refugiarse en su granja familiar y desarrolló la mayor parte de sus contribuciones científicas sobre mecánica, la teoría de la gravitación universal y los cimientos del cálculo diferencial utilizado en nuestros días.

Para viajar desde el cálculo diferencial de Newton a los actuales modelos epidemiológicos hicieron falta la intervención de la estadística de los ingleses Francis Galton y Karl Pearson, así como de las contribuciones de científicos especialistas en diferentes campos como Charles Darwin, Johann Mendel, James Watson, Francis Crick o Rosalind Franklin y de desarrollos tecnológicos como los microscopios ópticos y electrónicos y, por supuesto, los ordenadores.

El primer artículo conocido, con un modelo explícito para una enfermedad infecciosa basado en ecuaciones diferenciales, se debe a Daniel Bernoulli en 1760. El segundo se debe a sir Ronald Ross quien explicó el ciclo completo de la malaria incluyendo al mosquito como vector y el parásito *Plasmodium*, lo que le valió el premio Nobel en 1902. La siguiente contribución matemática se debe a Gray McKendrick y William Kermack, entre 1927 y 1939, con su teorema umbral según el cual individuos infecciosos en una población de susceptibles pueden originar una epidemia sólo si el número de susceptibles es mayor que un cierto valor crítico.

Modelo de Verhulfts

Intercalados en el tiempo con los modelos de pandemias, existieron modelos como el de Verhulfts (1804-1849), que es un modelo poblacional de crecimiento limitado, que supone que el crecimiento de una población está limitado por el tamaño y la fertilidad de la población y la cantidad de recursos disponibles, de manera que la población total converge hacia una situación estacionaria. Este modelo es una modificación del modelo de crecimiento exponencial de Malthus (1766-1834). El modelo de Verhulst también sirve para modelizar (o describir) la dinámica de una pandemia y se basa en una ecuación diferencial de primer orden que puede expresarse como sigue:

$$\frac{dP}{dt} = \alpha P(K - P),$$

donde P es el tamaño de la población, que depende del tiempo, y $\alpha = rK$ es un parámetro que proporciona una idea de la velocidad global de crecimiento de la pandemia, siendo r la constante de crecimiento (o decrecimiento) de la población y K la población máxima o capacidad de carga. Este modelo puede representar adecuadamente la propagación de una epidemia. Al principio la epidemia se propaga rápidamente ya que cada infectado es susceptible de contagiar a otros individuos con los cuales tenga relación, pero a medida que el número de infectados crece, es más difícil encontrar una persona que no haya estado en contacto con la enfermedad. Esta es la razón del crecimiento limitado, al cual también contribuyen las medidas de confinamiento. Este modelo tiene la ventaja de ser sencillo, robusto y de fácil comprensión.

De lo anteriormente escrito, se deriva que el modelo de Verhulst depende sólo de tres parámetros que deben ser identificados: la

población inicial de infectados P_0 que dio lugar a la pandemia, la tasa de crecimiento intrínseco r , y la capacidad de carga K . Conocidos estos parámetros es relativamente sencillo simular el crecimiento de una pandemia, en lo que se conoce como problema directo de predicción. A partir de aquí, lo interesante es poder calcular el valor de estos parámetros a partir de los datos observados de la pandemia que son medibles, como el número de infectados, el de muertes, el de ingresados en UCI, etc., con el fin de poder predecir el número de cada uno de estos grupos con cierta garantía de éxito.

La estrategia consiste en calcular un conjunto de modelos $\mathbf{m} = (K, P_0, r)$ que ajusten la serie temporal de datos observados de infectados $(t_1, P_1), (t_2, P_2), \dots, (t_n, P_n)$ con un error similar, menor que una cierta tolerancia. Puesto que es posible encontrar un amplio conjunto de modelos de estas características (modelos equivalentes), se hace un análisis del mismo (análisis de incertidumbre) con el objetivo de encontrar una muestra representativa. En este análisis, también se encuentra el modelo que predice los datos con el menor error, llamado modelo de máxima verosimilitud. A continuación, una vez que se ha obtenido el conjunto de modelos equivalentes, es posible extrapolar la curva de pandemia ad-futurum obtenida con todos ellos y determinar los diferentes percentiles de la distribución a posteriori de los valores de la pandemia. El percentil p de la predicción ad-futurum, en el día t , es el número de infectados que deja el $p\%$ de las predicciones por debajo. Es decir, si 500 infectados están en el percentil 75 quiere decir que el 75% de las predicciones son de menos de 500 infectados. Los percentiles más utilizados son 10, 25, 50, 75 y 90. De esta manera, la realidad (el número de infectados observado) en ese mismo día (predicho con anterioridad) es una curva más del modelo. El percentil 50 o mediana es el valor más probable de los valores predichos, dado que es el centro de la dis-

tribución a posteriori. Si la evolución de la pandemia es hacia percentiles inferiores a la mediana entonces la pandemia estaría bajo control. Inversamente, si la pandemia evoluciona hacia percentiles altos entonces se puede concluir lo contrario. Las curvas de percentiles recogen la contribución de diferentes escenarios plausibles, no sólo de aquel que posee el menor error de ajuste.

La figura 1 muestra la distribución a posteriori de la predicción del número de infectados diarios a nivel mundial. Como puede observarse la pandemia supera la curva de p50 (mediana) a finales del mes de Mayo del 2020, superando con creces el percentil 75 de la predicción y casi llegando al 90. Igualmente se puede hacer con el número de muertes (figura 2). Se observa que pese a mostrar una tendencia global decreciente a partir del 10 de Abril, la serie se encuentra lateralizando con un número de muertes diarias entre 4000 y 6000. La cifra de muertes diarias superaba el percentil 90 de la predicción, con lo cual la pandemia distaba de estar controlada.

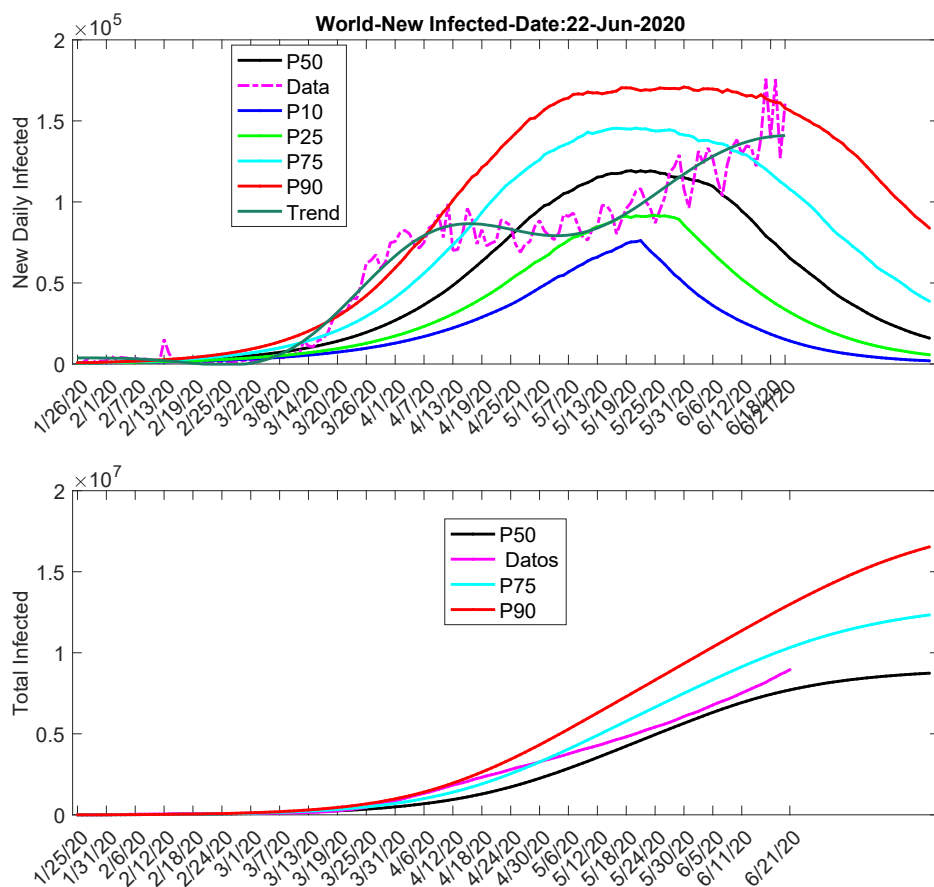


Figura 1: Distribución a posteriori de la predicción del número diario de infectados a nivel mundial.

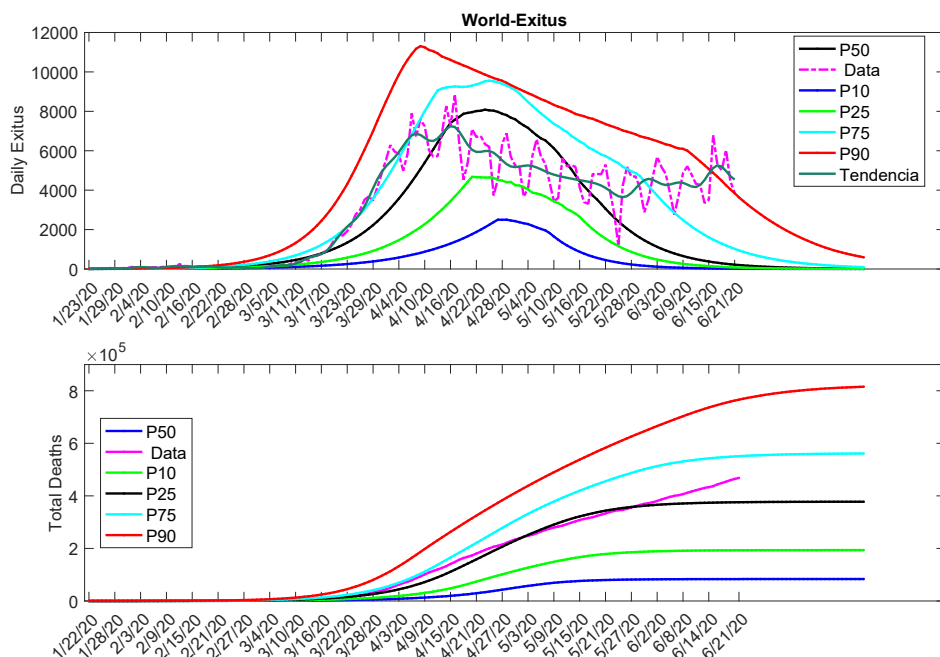
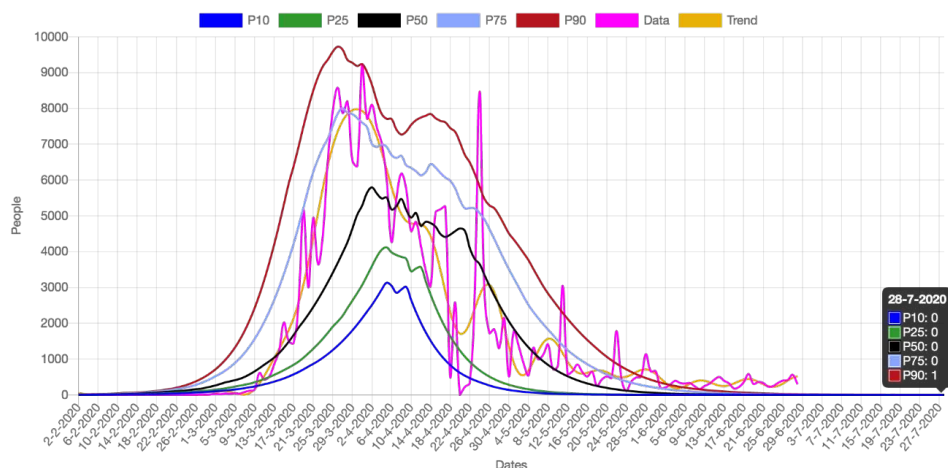


Figura 2: Distribución a posteriori de la predicción del número diario de muertes a nivel mundial.

En las figuras 3 y 4 se muestran los datos de infectados y muertes para España con fecha del 27 de junio de 2020.



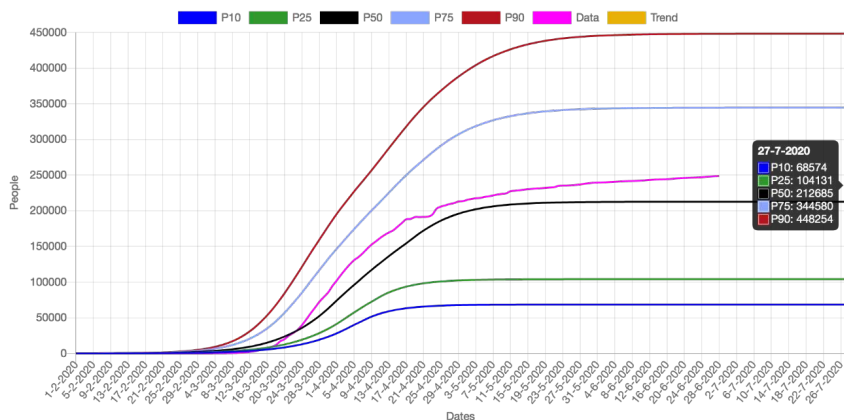


Figura 3: Predicción del número de infectados en España a fecha del 27 de Junio del 2020.

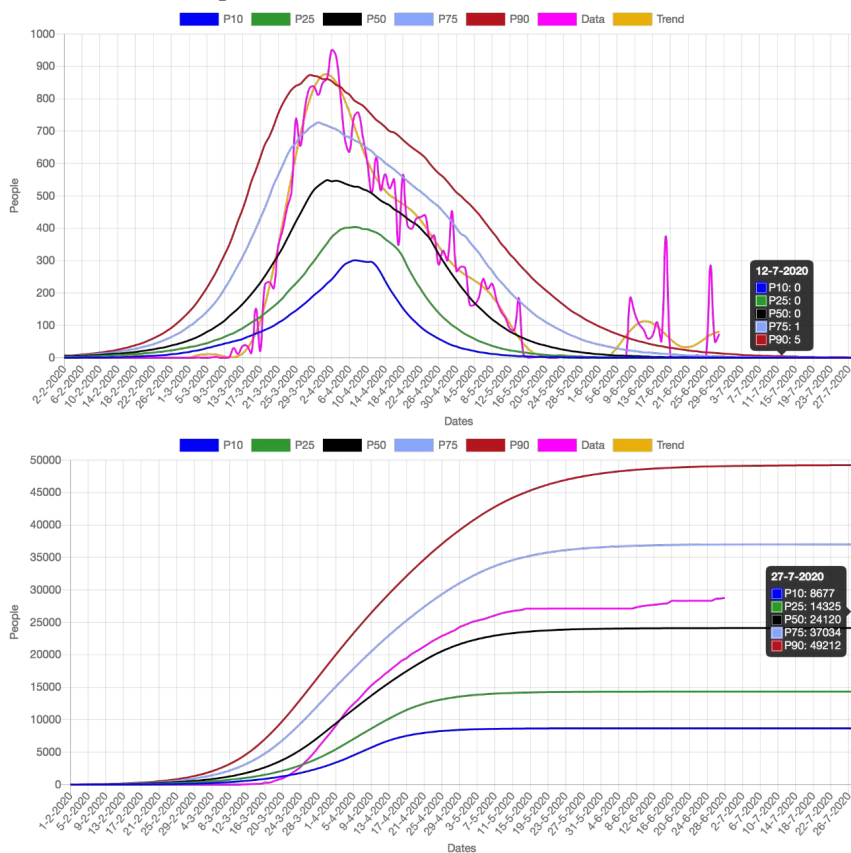


Figura 4: Predicción del número de muertes en España a fecha del 27 de Junio del 2020.

La predicción a largo plazo se puede complementar con una predicción a corto plazo (próximo día) mediante un método de análisis de series temporales. La figura 5 muestra los resultados de dicho método para la predicción a un día del número de infectados y de muertes en España. De esta manera el valor más probable es la mediana. Una cifra superior al percentil 90 se puede interpretar como una probabilidad de rebrote muy elevada.

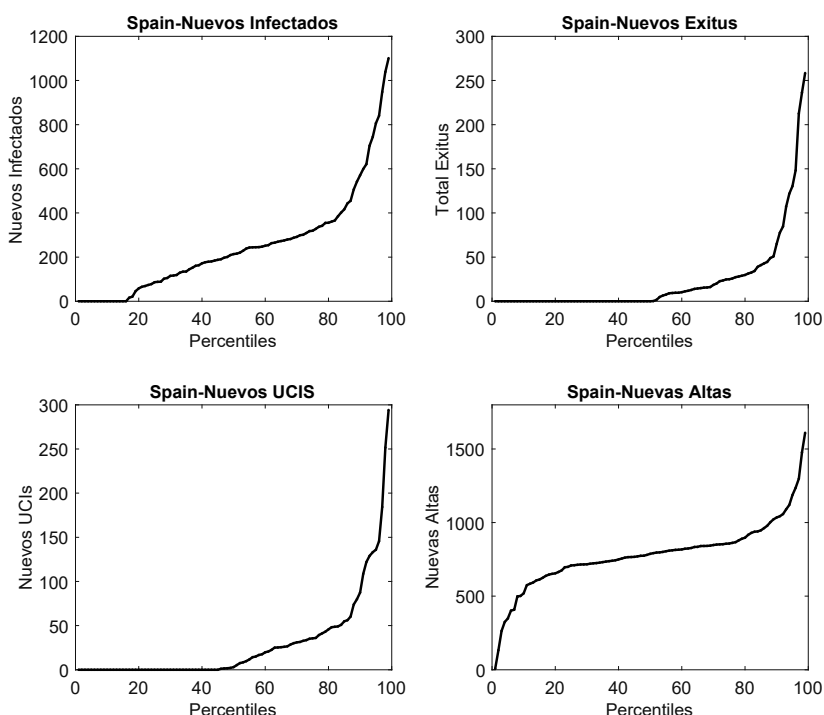


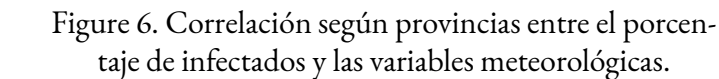
Figura 5: Predicción a corto plazo para España a fecha del 27 de Junio del 2020.

Estos métodos también se pueden utilizar para predecir las necesidades sanitarias (ingresos, Ucis, pacientes críticos, decesos). Estas predicciones se pueden regionalizar e incluso circunscribir a cada

hospital si se conocen los datos desagregados de la serie. Esto permitiría un control mucho más efectivo de la pandemia y una previsión inteligente y automatizada de las necesidades hospitalarias.

Influencia de las condiciones meteorológicas

Se ha encontrado que las variables meteorológicas son importantes en la propagación de la COVID-19. En general una mayor insolación, temperatura y humedad disminuyen las posibilidades de infección. La figura 6 muestra los porcentajes acumulados de infección COVID en las diferentes provincias españolas y su variación con la presión atmosférica y la temperatura media. Estas variables se pueden utilizar para condicionar la probabilidad de nuevas infecciones. La figura 7 muestra diferentes mapas de España donde se muestran la densidad de infectados, la temperatura media durante el mes de Abril y el porcentaje de anticuerpos detectados durante la primera ronda. Estos datos sirven para crear un modelo que permita condicionar el número de infectados mediante las variables climatológicas y establecer mapas inteligentes de inmunidad de grupo y de riesgos de infección y rebrotes.



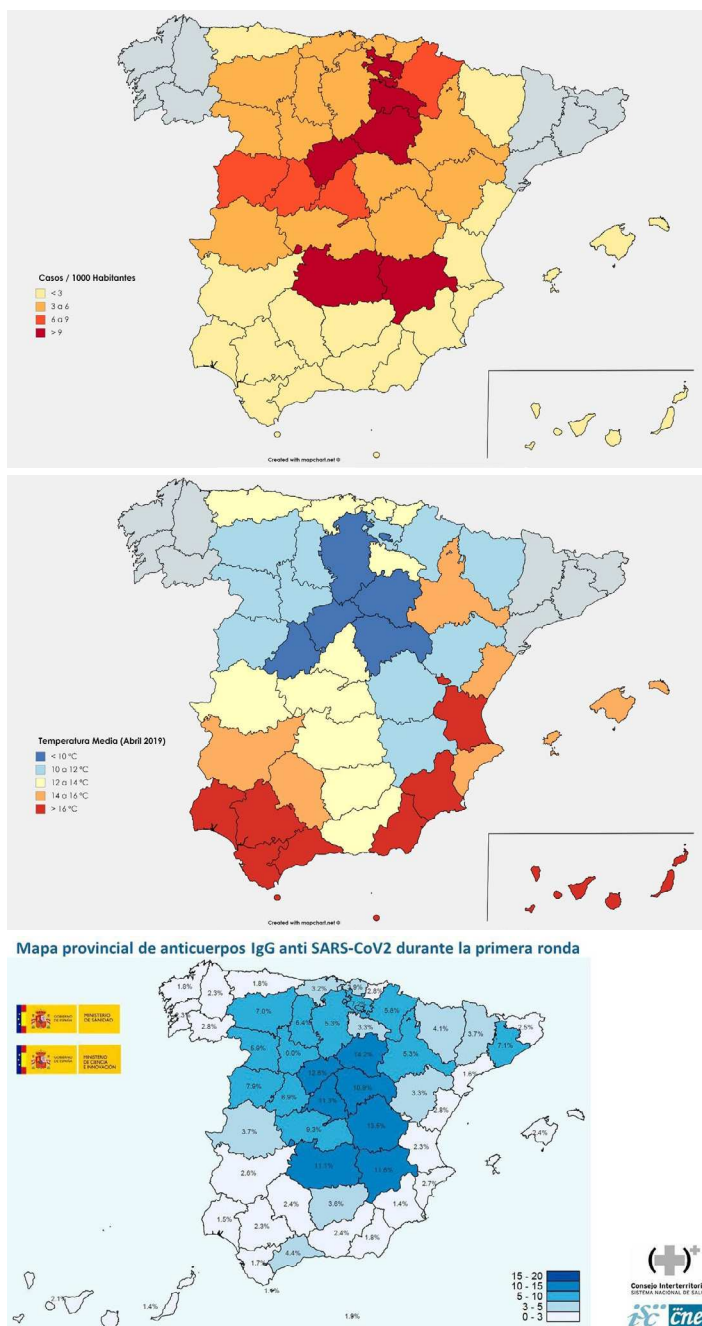


Figura 7: Mapa provinciales que muestran la similitud entre el número de casos Covid por cada 1000 habitantes, la temperatura media durante el mes de abril y el porcentaje de anticuerpos.

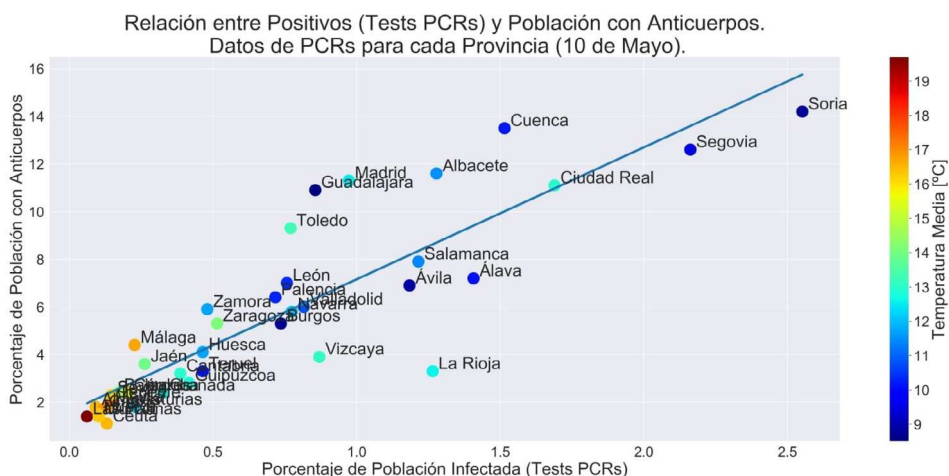


Figura 8: Porcentaje población con anticuerpos vs positivos PCR.

La figura 8 muestra la correlación existente entre los positivos detectados por PCR y la población con anticuerpos en las diferentes provincias de España. También se añade información relativa a la temperatura según los datos proporcionados por el Ministerio der Sanidad el 10 de Mayo.

Conclusiones

En este documento se presenta el modelo para la predicción de la pandemia Covid19 cara a la toma de decisiones, tanto de necesidades sanitarias como de salud pública. Se trata de la combinación de dos modelos de predicción a largo y corto plazo que utilizan como variables de entrada las series temporales relacionadas con la pandemia y con sus necesidades sanitarias asociadas. Varios fueron los modelos que se utilizaron para la predicción de la dinámica de la pandemia del Covid19 y de las necesidades sanitarias que esta pandemia genera (ingresos Covid, pacientes críticos, necesidad de unidades de cuidados intensivos, etc). Uno de los más utilizados, por

ser también de los más sencillos, se basa en la combinación de un modelo de crecimiento limitado (modelo logístico o de Verhulst) y un modelo predictivo a corto plazo, basado en el análisis de series temporales que permite hacer predicciones para el día siguiente. El modelo se puede regionalizar y enriquecer con otro tipo de variables -climatológicas, de movilidad, socioeconómicas, etc.- para ayudar a la toma de decisiones óptimas que permitan minimizar el impacto del Covid en la salud y en la actividad económica. En cualquier caso, un modelo de riesgo tiene que estar basado en modelos matemáticos y de inteligencia artificial robustos, es decir, que realicen un análisis de incertidumbre adecuado, cara a la cuantificación de riesgos. También se pueden utilizar otros modelos a largo plazo similares al de Verhulst (Gompertz, SIR) para realizar predicciones basadas en el consenso.

Referencias

1. <https://www.deepbioinsights.com/covid.html>. Trabajo realizado en colaboración con Zulima Fernández-Muñiz (profesora titular del Dpto. de Matemáticas) y Gonzalo Alonso (front-end).
2. Diekman O, Heesterbeek JAP, Metz J., 1990. On the definition and the computation of the basic reproduction ratio R_0 in models of infectious diseases in heterogeneous populations. *J Math Biol*, 28, 365-382.
3. Dietz K., 1967. Epidemics and Rumors: A survey. *JR Stat Soc [Ser A]* 130(4), 505-528.
4. Montesinos-López, O.A., Hernández-Suárez, C.M., 2007. Modelos matemáticos para enfermedades infecciosas. *Salud pública de México*, 49, 3, 218-226.

5. Juan Luis Fernández-Martínez, 2020. Un modelo robusto para la predicción ad-futurum de los efectos de la epidemia del Covid-19. Documento de Trabajo - 2020/07. Fedea
6. José MaríaLoché Fernández-Ahúja, Juan LuisFernández Martínez, 2021. Effects of climate variables on the COVID-19 outbreak in Spain. International Journal of Hygiene and Environmental Health, Volume 234, May 2021, 113723.
7. Pelayo Martínez Fernández, 2021. Modelos matemáticos en epidemiología y predicción de pandemias. TFG del doble Grado de Matemáticas y Física. Universidad de Oviedo. Directores: Juan Luis Fernández-Martínez y Zulima Fernández-Muñiz.