

Actualización de la predicción sobre la evolución del covid-19

5 de abril de 2020

Calibrado con los datos reales.

Por datos reales nos referimos a los registros que hace públicos el ministerio de sanidad. En la figura 1 se muestran los valores para los afectados (casos) hasta el 4 de abril

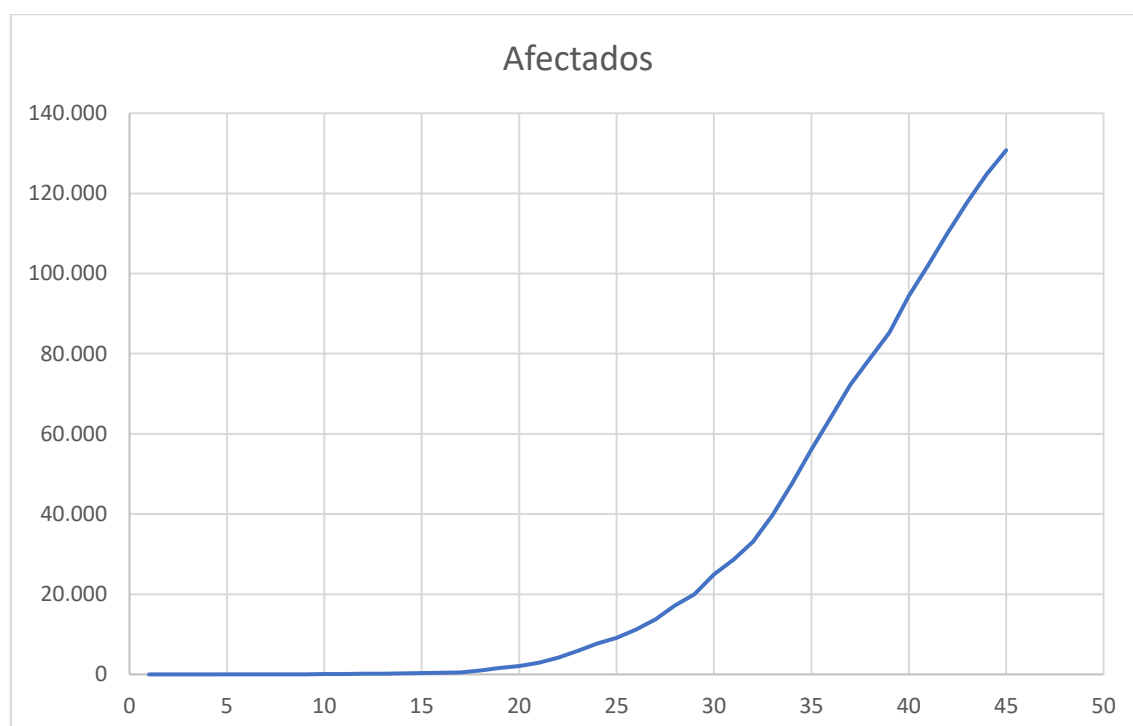


Figura 1. Evolución del número de afectados (casos) a 4 de abril.

La extrapolación de los datos reales debe tener en cuenta que la curva de la epidemia es una campana y con una curvatura determinada.

Por esto, el método será comparar la “forma” de la curva de los datos reales con la forma de la curva que predice el modelo. Para comparar hay que homogeneizar, lo que requiere cambiar el origen y la escala de los resultados del modelo de forma que las desviaciones respecto a los datos reales sean las mínimas. El cambio de origen que resulte se deberá, en buena parte, a que los valores del modelo se corresponden al día de la infección y los registrados por sanidad al día en que dio la cara la enfermedad, es decir, entre ellos media el periodo de incubación. La escala da cuenta de la diferencia entre la población simulada y la real.

El procedimiento para superponer las curvas es desplazar sucesivamente los resultados del modelo en 1,2,3, ..., n días y calcular el factor de escala que minimiza las diferencias entre realidad y modelo para cada uno de estos desplazamientos, hasta encontrar la pareja (desplazamiento, escala) que produce una menor diferencia. Como es habitual la

diferencia se calcula como la suma de los cuadrados de las diferencias entre todos los pares de puntos.

En los documentos previos se consideró que los datos reales aún no mostraban los efectos de las restricciones, pero era una aproximación que no se puede sostener, porque los datos reales presentan un evidente cambio de tendencia. Aunque el cambio a simple vista pueda parecer pequeño, su influencia es muy grande al realizar extrapolaciones, por ejemplo, si consideramos el 24 de marzo como la fecha de corte, la figura 2 muestra los dos periodos con sus correspondientes ajustes por mínimos cuadrados. Las tendencias son claramente diferentes y la extrapolación es manifiestamente divergente.

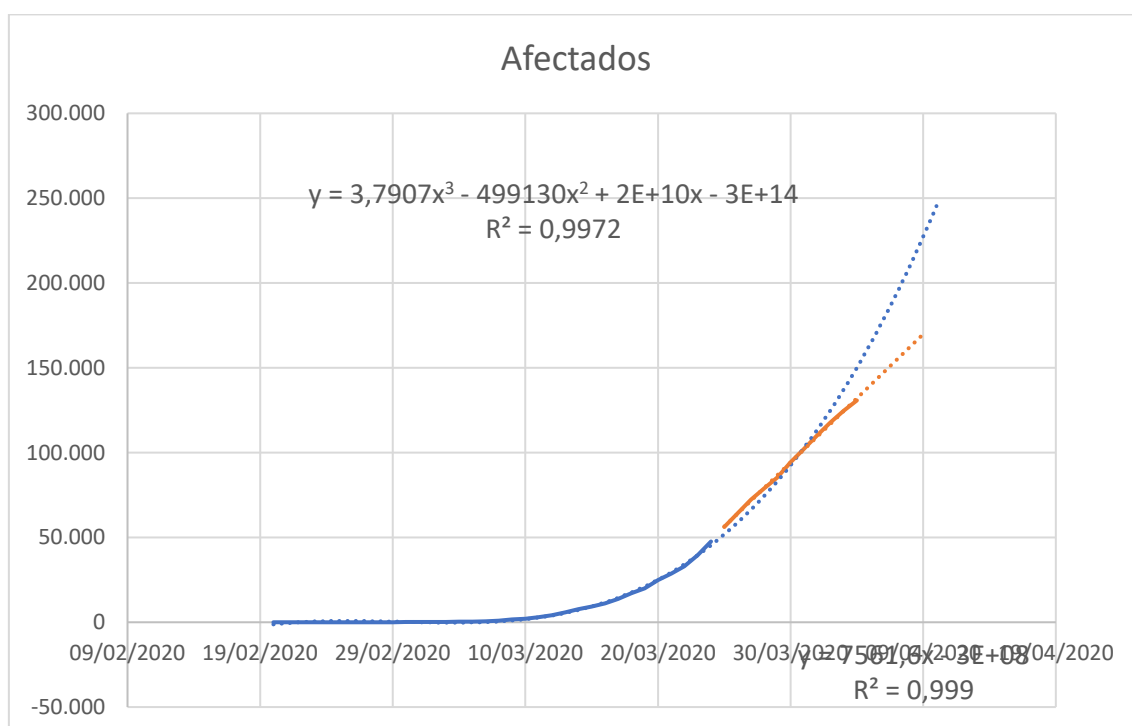


Figura 2. Cambio de tendencia en los datos reales del número de afectados a partir del día 24 de marzo.

Si se establece el punto de corte 4 días antes (20 de marzo) se ve como la tendencia pasa de convexa a cóncava (figura 3)

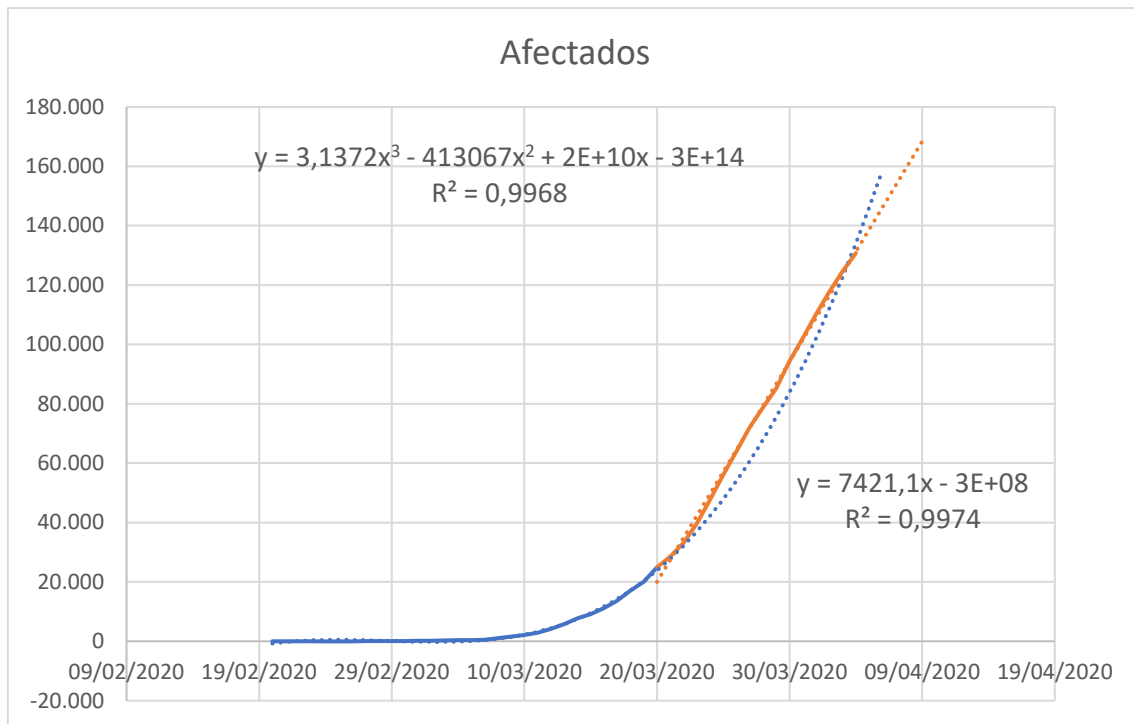


Figura 3. La curvatura pasa de cóncava a convexa a partir de un determinado momento.

Aunque lo apropiado es hacer el análisis mediante la derivada de la curva (casos diarios), el “ruido” de los datos lo hace poco evidente, figura 4.



Figura 4. Nuevos casos diarios.

La fecha de corte marca el punto en el que el efecto de las restricciones empieza a manifestarse y sería un error no tratar ambos periodos por separados.

En primer lugar procedemos a ajustar las curvas real y simulada con los datos previos al 24 de marzo obteniéndose los resultados de la figura 5 y tabla 1.

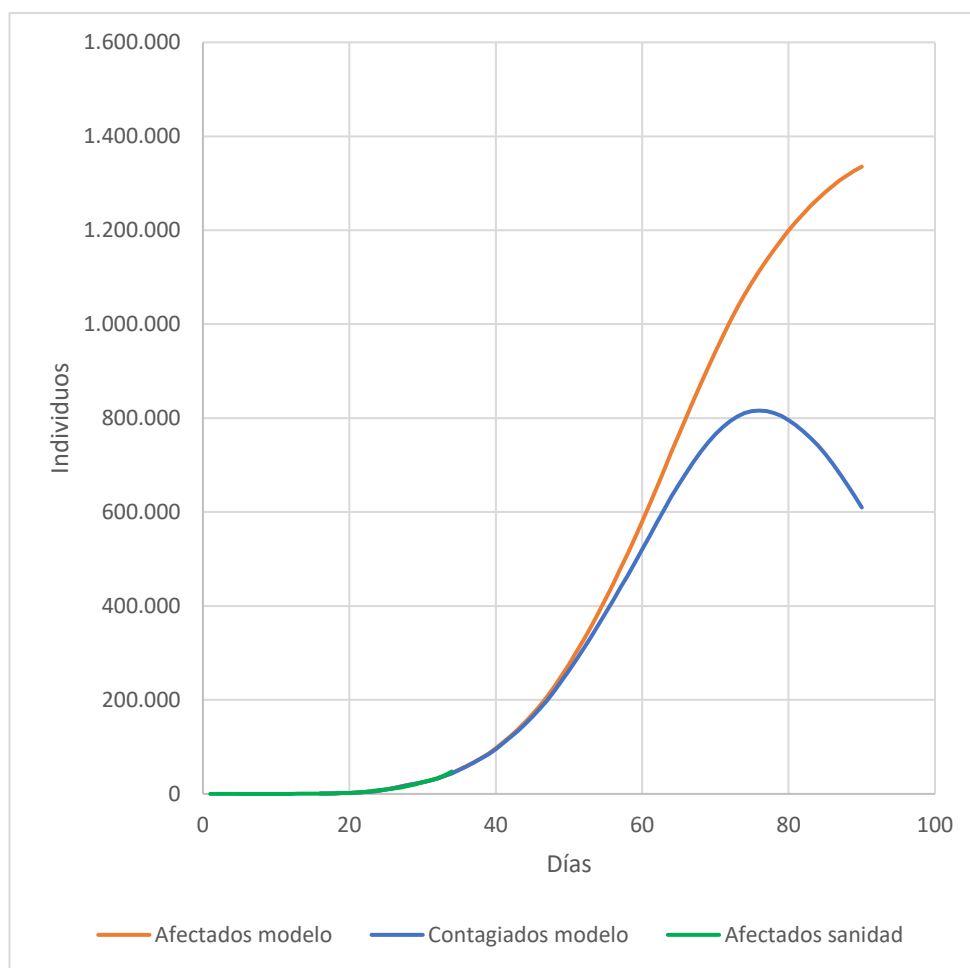


Figura 5. Predicciones del modelo calibrado con el período sin restricciones.

<u>Desplazamiento</u>	<u>Escala</u>	<u>SEC</u>	<u>N</u>	<u>SEC/N</u>	<u>Máximo Contagiados</u>	<u>Máximo Afectados</u>
16	93	14.911.566	18	828.420	816.034	1.113.658

Tabla 1. Predicciones del modelo calibrado con el período sin restricciones.

Los resultados de la tabla 1 los podemos interpretar como los valores a los que se habría llegado sin imponerse restricciones. El máximo de personas que el sistema de sanidad tendría que hacer atendido simultáneamente habría sido más de 800.000 y en ese momento los afectados sumarían más de 1.100.000 personas. Estos resultados hay que tomarlos con mucha precaución porque es una extrapolación a larga distancia. Como puede verse en la figura 5, los datos reales (línea verde) corresponden a un periodo muy pequeño en relación con el necesario para alcanzar la punta.

El calibrado del modelo para el período a partir del 24 de marzo (incluido) se muestra en la figura 6 y la tabla 2.

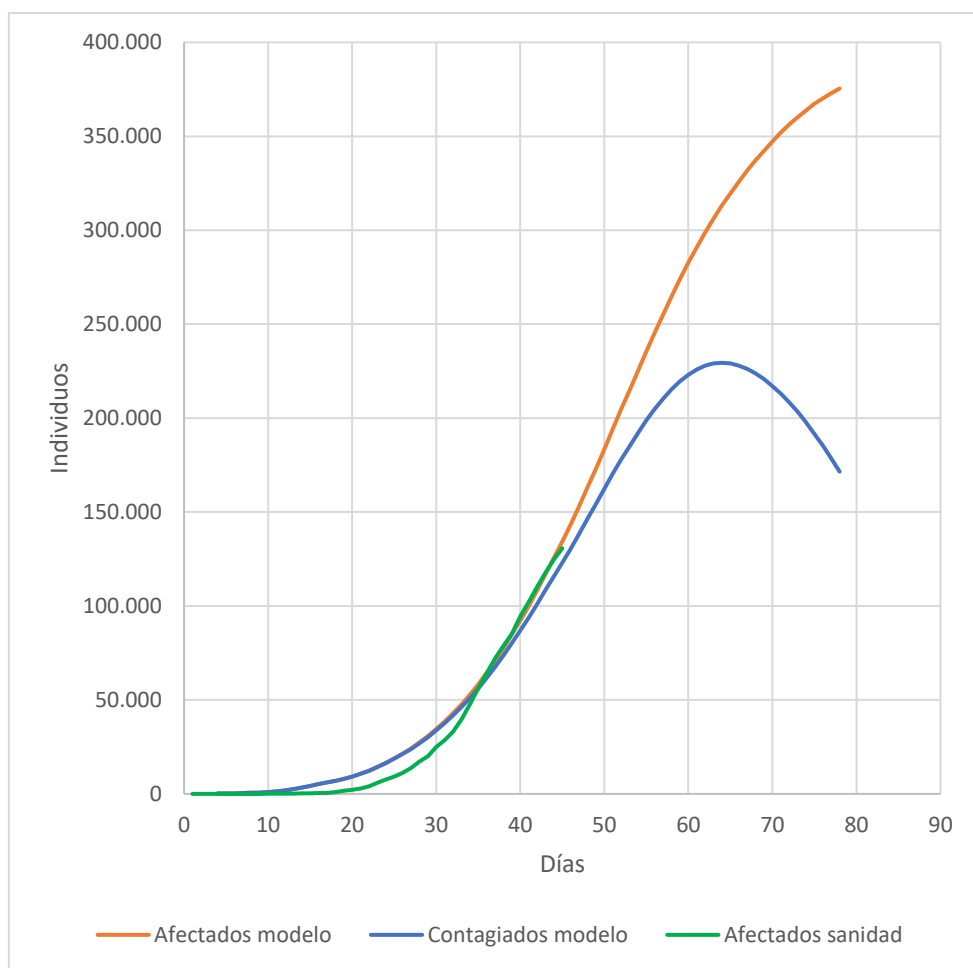


Figura 6. Predicciones del modelo calibrado con el período con restricciones.

Desplazamiento	Escala	SEC	N	SEC/N	Máximo Contagiados	Máximo Afectados
4	26	62.786.385	9	6.976.265	229.432	313.110

Tabla 2. Predicciones del modelo calibrado con el período con restricciones.

Estos resultados apuntan a que aún faltan 20 días para llegar al máximo, pero las cifras de contagiados simultáneamente y afectados son mucho menores que las que habrían sido sin las restricciones, además el pico se produce 14 días antes que sin restricciones.

En la figura 6 se observa que los datos reales (línea verde) se ajustan ligeramente por debajo del modelo en los últimos días, muy probablemente por un nuevo cambio de tendencia, derivado de las restricciones más exigentes que se impusieron la última semana, por lo que habría que esperar que las cifras que se alcancen en la realidad sean inferiores a las de la tabla 2.

Finalmente, hay que hacer una consideración sobre el procedimiento de superposición de curvas. Estamos comparando el inicio de la epidemia, que es muy sensible a las circunstancias que desencadenan la epidemia y que con seguridad han sido distintas de

unos lugares a otros. Si el foco de la infección surge en una zona densamente poblada, con transporte colectivo masivo (aeropuertos) o con motivo de concentraciones de personas (ferias) tendrá un inicio mucho más explosivo que si no se dan estas circunstancias. Una vez que la epidemia progresa el punto inicial pierde relevancia, pero nosotros estamos haciendo la comparación en la parte inicial y necesitamos que el modelo sea representativo del inicio real en promedio de la epidemia. El modelo se comporta como la realidad, si al situar los focos iniciales al azar, son pocos o ningunos los que son encuadrados en clusters, la epidemia progresará al principio lentamente. Con una sólo simulación se depende de la suerte. Para estimar el número de afectados y contagiados a la vez en el conjunto de la epidemia esto tendrá poca importancia pero no al comparar la fase inicial.

Para evitar este sesgo es necesario realizar más de una simulación del mismo escenario, lo cual es norma en el Método Montecarlo, el inconveniente es el tiempo de cálculo para un ordenador personal que es nuestra herramienta de cálculo.

Las figuras y tablas anteriores se han obtenido con la media de 30 simulaciones de una población de 50.000 individuos (figura 7).

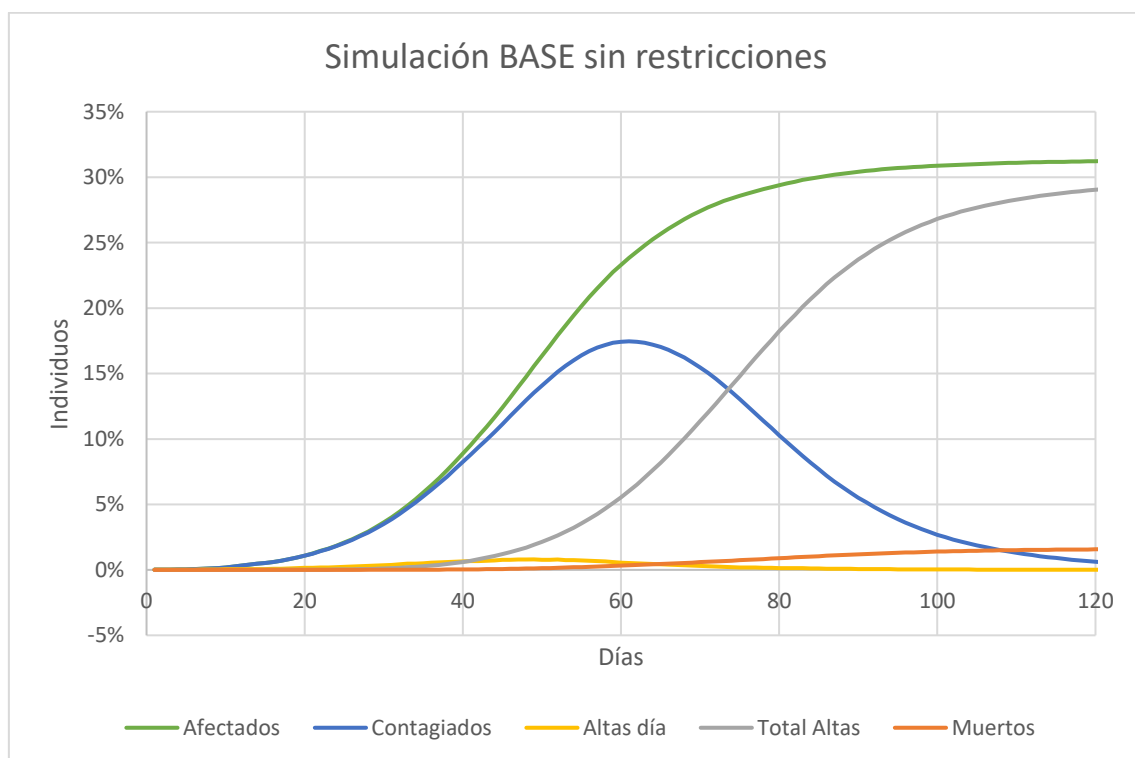


Figura 5. Valores medios de 30 simulaciones para una población de 50.000 individuos sin restricciones.