




ARTÍCULO ORIGINAL

**Método estadístico matemático para identificar el estado de la COVID-19 con relación al pico epidémico**

**Statistical and mathematical method to monitor COVID-19 status in relation to the peak of the epidemic**

Javier Pérez-Capdevila<sup>1\*</sup> , Yanelis Cobas-Ortiz<sup>2</sup> , Alejandro Javier Pérez-Cobas<sup>3</sup> 

<sup>1</sup> Doctor en Ciencias. Especialidad Matemática y Ciencias Económicas. Profesor e Investigador Titular. Delegación de Ciencia, Tecnología y Medio Ambiente de Guantánamo. Guantánamo. Cuba.

<sup>2</sup> Máster en Ciencias. Especialidad Geografía y Psicología. Profesora Auxiliar. Investigadora Agregada. Universidad de Guantánamo. Guantánamo. Cuba.

<sup>3</sup> Estudiante de Segundo Año de la carrera de Medicina. Universidad de Ciencias Médicas Guantánamo. Guantánamo. Cuba.

\*Autor para la correspondencia: [capdevila@gtmo.inf.cu](mailto:capdevila@gtmo.inf.cu)

Recibido: 6 de noviembre de 2020 Aprobado: 13 de noviembre de 2020

**RESUMEN**

**Introducción:** varios modelos han intentado pronosticar y evaluar el estado actual de la pandemia que ha generado el coronavirus SARS-CoV2, siendo la evaluación la base fundamental para la toma de decisiones. Por ello, la importancia de identificar el estado de la COVID-19 en un lapso seleccionado es un proceso de gestión de información muy importante, el cual puede realizarse a través de métodos estadísticos y matemáticos con el fin de tomar decisiones para controlar la epidemia. **Objetivo:** proponer un método estadístico matemático para identificar el estado de la COVID-19 con relación al pico epidémico en lapsos escogidos. **Método:** se utilizaron métodos teóricos entre los que destacan el análisis, la síntesis y la abstracción, y otros de carácter puramente matemáticos. **Resultados:**

como resultado de la aplicación práctica del método se generan gráficos que ofrecen información válida y confiable para un eficaz proceso de toma de decisiones. **Conclusiones:** esta propuesta muestra robustez teórica y eficacia práctica que, aunque se elabora teniendo en cuenta los datos de Cuba, es extrapolable a cualquier otro país, e incluso a provincias y municipios.

**Palabras clave:** análisis de datos, coronavirus; COVID-19; estadística y datos numéricos; gráficos estadísticos; matemática; toma de decisiones



## ABSTRACT

**Introduction:** several models have tried to predict and evaluate the current status of the pandemic that the new coronavirus, labeled SARVS-CoV2, has caused. This evaluation would be the basis for decision making. Therefore, the importance of monitoring the COVID-19 status in a selected period of time is very important for the process of information management, which can be done through statistical and mathematical methods in order to make big decisions to control the epidemic. **Objective:** to propose a mathematical and statistical method to monitor COVID-19 status in contrast to the peak of the epidemic in a selected period of time. **Method:** several theoretical methods

were used, specially: analysis, synthesis, abstract; and other purely mathematical methods. **Results:** as a result of the practical application of the methods used, valid and reliable information was generated in charts, supporting an effective process of decision making. **Conclusions:** this proposal shows the robustness of its theoretical aspects and a practical effectiveness that, even if elaborated to submit Cuban-generated national data, it could be used in other countries, and even in a provincial or municipal level.

**Keywords:** data analysis; coronavirus; COVID-19; statistical and numerical data; statistical graphs; mathematics; decision making

### Cómo citar este artículo:

Pérez-Capdevila J, Cobas-Ortiz Y, Pérez-Cobas AJ. Método estadístico matemático para identificar el estado de la COVID-19 con relación al pico epidémico. Rev Inf Cient [en línea]. 2020 [citado día mes año]; 99(6):553-562. Disponible en: <http://www.revinfcientifica.sld.cu/index.php/ric/article/view/3147>

## INTRODUCCIÓN

El nuevo coronavirus ha sido catalogado por la Organización Mundial de la Salud primeramente como una emergencia en salud pública de jerarquía internacional y, posteriormente como, una pandemia global.<sup>(1)</sup>

La enfermedad denominada COVID-19 provocada por el nuevo coronavirus causa daños significativos a la salud humana, de ahí, la importancia de aplanar la curva, y una enunciación útil -no la única- al respecto es que el número de casos activos deje de incrementarse y comience a disminuir, es decir, el gráfico de casos activos crece, se aplanan y luego disminuye. Lamentablemente, desde el punto de vista matemático, en Cuba no se ha seguido una modelación algorítmica encaminada a aplanar la curva.

Diferentes modelos estadísticos matemáticos y de pronósticos intentan mostrar en qué estado se encuentra la epidemia destacando las curvas relativas a los casos activos, los que no son la única variable digna de estudio durante la epidemia, pero son muy importantes no solo porque son los que están sometidos al sufrimiento de una enfermedad que puede provocarles la muerte, sino también porque son fuentes potenciales de nuevas infecciones. Cuando los casos activos disminuyen, la probabilidad de que se incrementen también disminuye.



Los casos activos se aplanan cuando los casos nuevos son menores que la suma de recuperaciones y fallecimientos. Se trata de una aritmética simple con tres variables.

En Cuba, la matemática ha definido esencialmente tres escenarios (favorable, medio y crítico), los cuales se han mantenido bastante constantes durante toda la epidemia. También se han modelado curvas de pronósticos que hasta la fecha se han comportado mayoritariamente por debajo del escenario favorable.

La principal problemática teórico práctica se halla en que el comportamiento de la epidemia, según estos modelos, se ha mantenido en un escenario favorable, a pesar de:

- Tener el 80 % de casos activos con respecto al pico epidémico.
- Más de 20 personas en estado crítico o grave.
- 14 fallecidos en 12 días.

Estos datos reflejan una situación desfavorable, lo cual es contradictorio con los modelos.

Dada la contradicción mencionada anteriormente, el propósito esencial de este artículo es proponer un método estadístico matemático para identificar el estado de la COVID-19 con relación al pico epidémico. Este método, aunque se elabora teniendo en cuenta los datos de Cuba, es extrapolable a cualquier otro país.

## DESARROLLO

Disponer de datos precisos y accesibles sobre una epidemia es transcendental porque estos facilitan información válida y confiable acerca de la propagación de la enfermedad y permiten a las autoridades competentes el proceso de toma de decisiones para su control.

En Cuba ha existido desde el comienzo de la epidemia transparencia total en la divulgación de los datos, lo cual facilita el trabajo de científicos e investigadores porque se pueden realizar estimaciones de los cambios en la transmisión a través del tiempo, que proveen una comprensión de la situación epidemiológica<sup>(2)</sup>, y la estimación identifica si las medidas de control están teniendo un efecto deseado.<sup>(3,4)</sup>

Basados en esos datos se ha publicado recientemente el gráfico de los escenarios y seis curvas de pronóstico, el cual se muestra en la Figura 1 y se analiza a continuación, críticamente, teniendo en cuenta que la crítica científica es esencial para que el conocimiento crezca, y se puedan encontrar elementos que no fueron tomados en cuenta en una investigación, cabos sueltos, otras opiniones, puntos de vistas diferentes e incluso errores.

Además, ese tipo de crítica es una modalidad de publicación, y es tan respetable y respetada que las revistas más prestigiosas del mundo, y las de mayor impacto la incluyen desde el siglo XVII.<sup>(5)</sup>



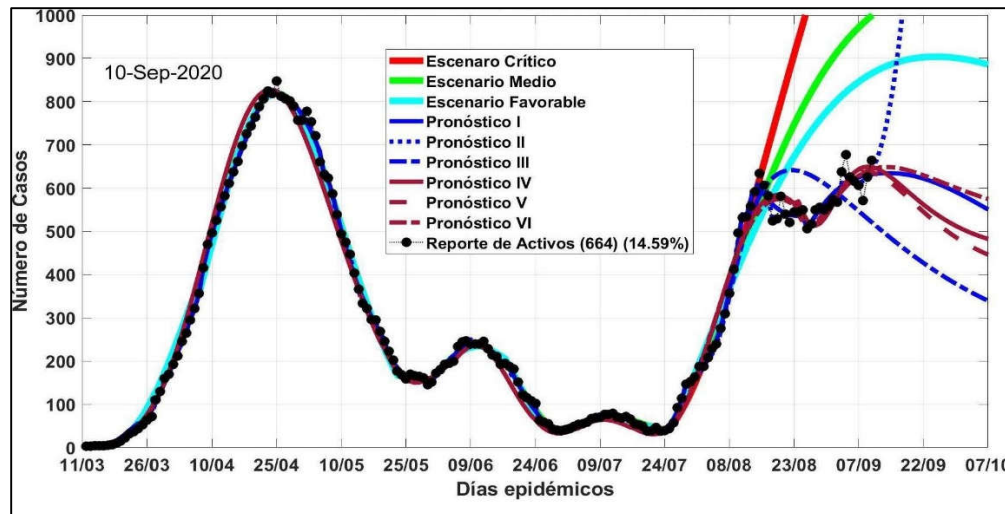


Fig. 1. Estimación de los casos activos en Cuba.

Primero que todo hay que considerar que la ciencia tiene una dimensión moral que plantea por un lado, el ascenso de los derechos humanos y, por el otro, la forma en que el trabajo científico debe llevarse a cabo.<sup>(6)</sup> Con respecto al gráfico, el análisis desde el punto de vista axiológico se fundamenta en que bajo condiciones de una epidemia las personas no se encuentran nunca en un escenario favorable sinónimo de un estado propicio, apacible o benévolo, incoherente con una situación real, donde prima el miedo a una enfermedad que se propaga rápidamente; nadie puede considerar la muerte en un escenario favorable.

Por otra parte, y ya desde el punto de vista científico, es falible construir una curva para el caso del rebrote con un máximo valor superior al pico epidémico, lo que ha provocado que, a pesar de una gran cantidad de casos activos entre el 61 y el 80 % del pico epidémico durante los días que van desde el 10 de agosto hasta el 14 de septiembre, con el fallecimiento de 14 personas en solo 12 días (lo cual no ocurría desde el 12 de mayo), se haya considerado que estamos en un escenario favorable.

Por suerte es obvio que las autoridades cubanas no tomaron decisiones basadas en este gráfico, pues de haberse considerado que estábamos en una situación favorable no se hubieran adoptado las medidas precisas, coherentes y severas que se han llevado a efecto.

Es evidente que estos modelos no brindan información válida para una alerta necesaria y suficiente para que se realice un proceso de toma de decisiones coherente. También, la utilización de seis pronósticos distintos es una muestra de falta de precisión sobre todo cuando el rango entre ellos supera fácilmente decenas de miles de casos, provocadas por el pronóstico II, que además parece irse casi asintóticamente al infinito.

En esencia, el problema no radica en lo desacertado de los pronósticos, sino en que se obtienen modelos de escenarios que no facilitan información para la toma de decisiones en momentos donde está en riesgo la vida humana.

A partir de esta crítica y, con basamento en que esta área de investigación es todavía incipiente, y, por lo tanto, es difícil confiar en un modelo único para la predicción<sup>(7)</sup>, se propone un nuevo método estadístico matemático, como modelo que genere información confiable para tomar decisiones, identificando el estado real en que se encuentra la COVID-19 (con datos de Cuba), a partir de una curva que compara los datos de activos con el pico epidémico.

En otras palabras, se trata de que los análisis del estado del rebrote sean realizados por comparación con nuestra propia realidad en la dinámica de la epidemia. Un marco referencial atinado sería el pico epidémico cuyo valor ascendió a 847 casos.

Siguiendo la idea anterior, se utilizará el porcentaje que es usado para definir relaciones entre dos cantidades, en este caso serán relacionados el número de casos activos diarios con respecto al pico epidémico para medir los niveles alcanzados de los datos consultados como fundamento de decisiones, así como también para entender la magnitud de cambios de la medida utilizada.

De esta manera, la estimación de los cambios en la transmisión a través del tiempo puede proveer una comprensión dentro de la situación epidemiológica<sup>(2)</sup>, y la estimación identifica si las medidas de control están teniendo un efecto deseado.<sup>(8,9)</sup>

Por todo lo anterior, el cálculo del porcentaje de activos diarios con respecto al pico epidémico será el primer paso a seguir en este método (todos los datos tomados del tablero cubano en internet<sup>(10)</sup>). Este cómputo puede hacerse para toda la serie, pero es más importante para los últimos días. En Cuba, generalmente se refieren con preponderancia los datos de los últimos 15 días. Con estos porcentajes se puede generar la curva, pero ello no es suficiente para visualizar el estado de la epidemia.

Entonces, con el objetivo de ir completando el método, será útil revelar el gráfico con una curva que evidencie el porcentaje que representan los casos activos con respecto a dicho pico y fijar umbrales que delimiten zonas interpretables. No puede evitarse que dichos umbrales tengan una carga de subjetividad, pero al menos responden al sentido común y pueden, a su vez, ser debatidos y comprendidos por cualquier persona. Las zonas representarían distintos estados de la epidemia y serían denominadas de la siguiente manera:

*Estado tolerable* (zona por debajo de la curva verde): donde hay una marcada confianza en que se tiene bajo control la epidemia porque se puede soportar y hacerle resistencia.

*Estado de alerta* (zona por debajo de la curva azul): es aquella que obliga al incremento de la vigilancia y de las medidas restrictivas. Disminuye el nivel de tolerancia.

*Estado de alarma* (zona por debajo de la curva rojo claro): identificada como la que exige cambios profundos en la estrategia y que pondría de manifiesto que se está en un marco de mucha gravedad.

*Estado de crisis* (zona por debajo de la curva rojo oscuro): es la que indica una inminente regresión al peor momento de la epidemia.



*Estado de emergencia* (zona por encima de la curva rojo oscuro coloreada en negro): situación de peligro o desastre que requiere una acción inmediata para decidir la restricción o suspensión de algunos derechos esenciales de forma parcial o total para garantizar el orden en el estricto cumplimiento de las normas dictadas por el Estado para evitar que la situación empeore y genere un caos aún mayor.

Consiguientemente, deben determinarse valores numéricos que delimiten los estados que se propongan. Es evidente que deben determinarse cuatro límites en el gráfico, como cotas superiores de los cuatro primeros estados (representados en zonas del gráfico) ya que el quinto estado estaría por encima del límite del cuarto.

Ya una vez conocida la forma de calcular y construir la curva, además de las cinco zonas a delimitar en el gráfico y con el objetivo de determinar los valores que acotarán a cada zona o estado, se debe pasar al cálculo de la media de los casos activos durante 135 días de la epidemia, ya que el pico fue alcanzado en 45 días y se debe tener en cuenta que la demora en llegar al pico puede ser la mitad de la demora en bajar.

La media es el valor más importante cuando los datos son dispersos, sin seguir un patrón típico. La moda puede identificar aquellos valores que se producen con más frecuencia mientras la mediana es sesgada cuando los datos son dispersos y hay mayor cantidad de valores altos o bajos, pero la media captura todos los valores.

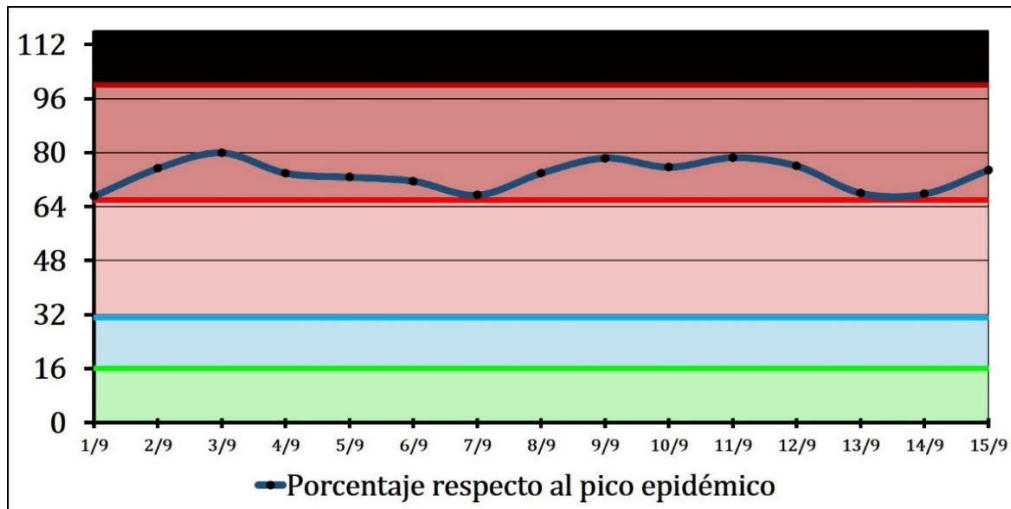
La epidemia muestra patrones no típicos y, en ellos, el número de casos varía de acuerdo a factores externos muy relacionados con el comportamiento humano. La media o promedio genera un valor intermedio influido por todos los datos y forma una buena base para la estimación de los valores futuros, siempre y cuando los factores externos sigan siendo los mismos, aun cuando se comporten de maneras diferentes.

El resultado del cálculo de la media es 273,92 casos, pero como la curva debe estar en porcentajes, hay que calcular qué porcentaje representa este promedio con respecto al pico, obteniéndose como resultado el 32,33 % del pico epidémico. Aproximando por defecto, entonces se tiene como media de las ordenadas del gráfico 32 %.

A partir de lo anterior se deben generar dos zonas hacia abajo entre 32 y 0 %, cuya media es 16 %. Por tanto, ya se tienen dos límites: 16 y 32 %. Ahora corresponde determinar el límite para generar dos zonas entre 32 y 100 % y la media será 66 %. Así quedan calculados y definidos los cuatro límites: 16, 32, 66 y 100 %.

Con los datos de casos activos diarios, el porcentaje que representa cada uno con respecto al pico epidémico, la definición de las cinco zonas que representan los cinco estados posibles de la epidemia y el cálculo y definición de sus límites, se obtiene el siguiente gráfico para 15 días de la epidemia en Cuba comprendidos entre el 1ro y el 15 de septiembre (Figura 2).

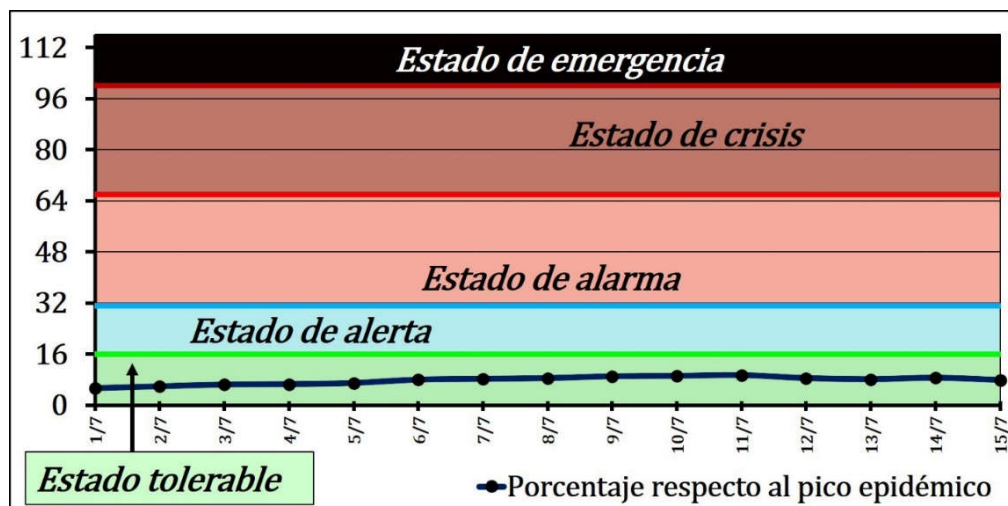




**Fig. 2.** Comportamiento de la curva de porcentajes en la primera quincena de septiembre.

Se observa que la curva se ha mantenido todo el tiempo en un estado de crisis, haciendo una meseta entre el 64 y el 80 % del pico epidémico, de donde se deduce que las medidas para hacer bajar la epidemia no han sido eficaces, y esto puede tener dos posibles causas esenciales, una porque las medidas hayan sido concebidas y, otra, por causa de su incumplimiento. En el segundo caso quedan implicados entre otros varios actores relacionados con: la comunicación social, el orden público, el comercio interior, el control del cumplimiento de las medidas implementadas por el gobierno sobre la base de indicadores medibles e independientes entre sí y la correcta aplicación de la ciencia. Por tanto, la información que ofrece el gráfico conduce a una revisión de las medidas y del comportamiento en general con respecto a ellas.

Sin embargo, aunque esa situación es la más actualizada concluida la escritura de este documento, con solo dos meses de antelación la curva tenía un comportamiento similar en cuanto a que igualmente hacía una meseta, pero en otro estado, y se muestra a continuación en la Figura 3.



**Fig. 3.** Comportamiento de la curva de porcentajes durante la primera quincena de julio.

Es evidente que en este período el gráfico informa un control sobre la epidemia, seguramente basado en una mayor disciplina en el comportamiento humano en general, es decir, tanto como de los decisores, así como de los ejecutores. Pero puede que un exceso de confianza haya conllevado a disminuir el rigor de las medidas para el enfrentamiento de la epidemia y, entonces, sea esa la causa de lo ocurrido unos días después, lo cual se muestra seguidamente (Figura 4).

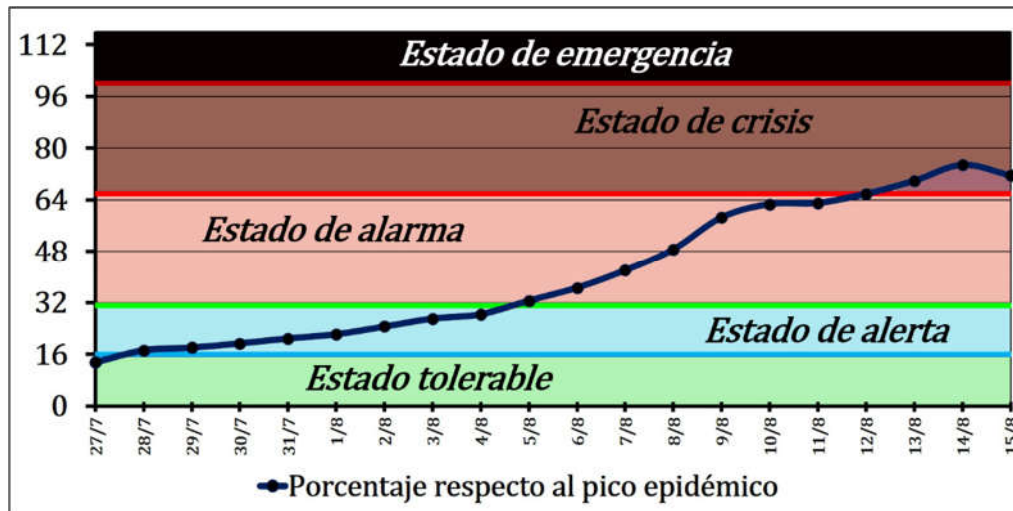


Fig. 4. Comportamiento de la curva de porcentajes entre el 27 de julio y el 15 de agosto.

El gráfico muestra un crecimiento en solo 20 días desde el estado tolerable hasta el de crisis. Entonces hay que volver a dudar si las decisiones fueron buenas o es que no se cumplieron a cabalidad.

El crecimiento desde un escenario mínimo hasta el máximo antes de declarar un estado de emergencia, habla en favor de que, de haberse analizado con este método, las decisiones habrían sido otras, pues los decisores tendrían información con la que no contaban en ese momento utilizando un método diferente al que se ha presentado.

De manera general, el ajuste a la realidad de este método y los gráficos que genera, tiene un basamento dialéctico que es imprescindible referir.

Hay un ascenso de la realidad a la abstracción para visualizar y conceptualizar los cinco estados en que puede encontrarse la epidemia en los lapsos que se escojan y la elección de los porcentajes con respecto al pico epidémico, como medida estándar para medir el comportamiento epidémico.

Los datos cuantitativos se transforman en información cualitativa para la toma de decisiones, además, el método presentado y las definiciones de los conceptos relativos a los cinco estados constituyen generación de nuevo conocimiento. Aquí se exterioriza el papel clave de la gestión de información para la toma de decisiones, ahora para el caso de la COVID-19.



## CONCLUSIONES

Han quedado demostradas las inexactitudes de los modelos utilizados para conocer el estado de la COVID-19 en Cuba, lo que condujo a que en este artículo se haya logrado proponer un nuevo método estadístico matemático, como modelo generador de información confiable para tomar decisiones, identificando el estado real en que se encuentra la COVID-19 (con datos del país), a partir de una curva que compara los datos de activos con el pico epidémico.

Los gráficos generados a partir del método brindan la información referida en el párrafo anterior y, por tanto, se convierten en un instrumento útil para identificar de manera más exacta tendencias de la epidemia, y así poder aplicar medidas correctivas para controlarla.

## REFERENCIAS BIBLIOGRÁFICAS

1. Díaz-Pinzona JE. Precisión del pronóstico de la propagación del COVID-19 en Colombia. *Repert Méd Cir* 2020; 29(Supl.1):27-33.
2. Camacho A, Kucharski A, Aki-Sawyer Y, White MA, Flasche S, Baguelin M, Pollington T, Carney JR, Glover R, Smout E, Tiffany A, Edmunds WJ, Funk S. Temporal Changes in Ebola Transmission in Sierra Leone and Implications for Control Requirements: A Real-time Modelling Study. *PLoS Curr* [en línea]. 2015 Feb. [citado 11 Ago 2020]. Disponible en: <http://currents.plos.org/outbreaks/index.html%3Fp=55052.html>
3. Jaramillo O. Pertinencia del perfil de los profesionales de la información con las demandas del mercado laboral. *Rev Int Bibliotecol* [en línea]. 2015 [citado 11 Ago 2020]; 38(2):111-120. Disponible en: <http://www.scielo.org.co/pdf/rib/v38n2/v38n2a3.pdf>
4. Rivera Z, León M, García T. El mercado laboral para el profesional de la información en Cuba: ¿qué piensan los empleadores al respecto? *Alcance: Rev Cubana Inf Comun* [en línea]. 2017 [citado 11 Ago 2020]; 7(15):6-27. Disponible en: <http://scielo.sld.cu/pdf/ralc/v7n15/ralc02118.pdf>
5. Dorta-Contreras AJ. Criticar la ciencia y ciencia de la crítica. *Rev Hab Cienc Méd* [en línea]. 2007 Nov [citado 11 Ago 2020]; 6(4):[aprox. 5 p.] Disponible en: <http://scielo.sld.cu/pdf/rhcm/v6n4/rhcm01407.pdf>
6. Holton G. On the Integrity of Science: The Issues Since Bronowski. *Leonardo* [en línea]. 1985 [citado 11 Ago 2020]; 18(4):229-232. Disponible en: <https://www.jstor.org/stable/1578071?seq=1>
7. Singhal A, Singh P, Lall B, Joshi SD. Modeling and prediction of COVID-19 pandemic using Gaussian mixture model. *Chaos Solitons Fractals* [en línea]. 2020 [citado 11 Ago 2020]; 138:110023. DOI: <https://doi.org/10.1016/j.chaos.2020.110023>
8. Chaolin Huang YW, Xingwang Li, Lili Ren, Jianping Zhao, Yi Hu, Li Zhang, Guohui Fan JX, *et al.* Clinical features of patients infected with 2019 novel coronavirus in wuhan, china. *Lancet* [en línea]. 2020 [citado 11 Ago 2020]; 395(10223):497-506. DOI: [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5)
9. Cooper BS, Pitman R, Edmunds WJ, Gay NJ. Delaying the international spread of pandemic influenza. *PLoS Med* [en línea]. 2006 [citado 11 Ago 2020]; 3(e212). DOI: <https://doi.org/10.1371/journal.pmed.0030212>
10. Covid19-Dashboard Cuba. [Covid19CubaData](https://covid19cubadata.github.io/#cuba) 2019. [citado 15 Sep 2020]. Disponible en: <https://covid19cubadata.github.io/#cuba>



**Conflicto de intereses:**

Los autores declaran que no tienen conflicto de intereses.

**Contribución de los autores:**

JPC: participó en la concepción y diseño del estudio, realizó el análisis e interpretación de datos. Creó el método matemático que se propone, y lo aplicó como se observa en Figura 2. Realizó la elaboración del borrador del artículo, así como su revisión crítica, con aportes importantes a su contenido intelectual. Participó en la aprobación de la versión final a ser publicada.

YCO: participó en la concepción y diseño del estudio, y construyó la base de datos a partir del tablero utilizado. Aplicó el método propuesto resultando las Figuras 3 y 4. Realizó una revisión crítica, con aportes importantes a su contenido intelectual, así como elementos destacados de edición. Participó en la aprobación de la versión final a ser publicada.

AJPC: participó en la recogida de datos, conformación y cálculos de la base de datos. Colaboró en la aplicación del método propuesto resultando las Figuras 3 y 4. Realizó una revisión crítica de su contenido intelectual. Participó en la aprobación de la versión final a ser publicada.

