

IMPLEMENTACIÓN DE MODELOS DE SERIES TEMPORALES EN CASOS DE COVID-19

Dilan Sneider Atencia Colina
CC 1.001.938.813
Código estudiantil: 201811491239
Correo: dilan.atencia@unisimonbolivar.edu.co

Juan Manuel Molina Londoño
CC 1.192.923.421
Código estudiantil: 201811493587
Juan.Molina2@unisimonbolivar.edu.co

Ronny Rodolfo Ariza Corro
CC 1.001.821.280
Código estudiantil: 201811494192
Correo: rony.ariza@unisimonbolivar.edu.co

Elier Andrés Fince Bertis
CC 1.143.465.645
Código estudiantil: 201611475910
Correo: efince1@unisimonbolivar.edu.co

Nathaly Yohanna Ramírez Pacheco
CC 1.192.816.300
Código estudiantil: 201811493146
Correo: nathaly.ramirez@unisimonbolivar.edu.co

Trabajo de Investigación del Programa **Ingeniería Industrial**

Tutor:
Alexander de Jesús Pulido Rojano

RESUMEN

Antecedentes: Una serie de tiempo se define como una secuencia cronológica de observaciones sobre una característica (univariante o escalar) o varias características (secuencia multivalente o vectorial). Las series temporales se basan en el supuesto de que el valor de la variable que se quiere estudiar es el resultado de tres componentes. Estos componentes son: el proceso tendencial, el componente estacional y el componente aleatorio. El componente tendencial se define como un cambio promedio a largo plazo. El componente estacional se

relaciona en muchas series de tiempo que muestran periodicidad, es decir, cambia un cierto período de tiempo y el componente aleatorio no responden a ningún patrón de comportamiento, solo es el resultado de factores accidentales o aleatorios en una forma aislada en una serie de tiempo. La importancia de las series temporales ha sido demostrada a lo largo del tiempo y muchas han sido sus aplicaciones para la mejora de la predicción. Por ejemplo, se han propuesto para introducir los modelos estadísticos para realizar un control óptimo y eficiente del tráfico de datos de una Red o generar pronósticos para medir la tendencia y comportamiento de datos de producción tomando en cuenta la estacionalidad.

Objetivos: El objetivo principal para este trabajo de investigación es analizar tres métodos de series temporales como: El método de media móvil simple, Media Móvil Ponderada y Suavización Exponencial aplicados en la predicción del número de casos positivos de COVID-19 en Colombia. La idea fue identificar que método arroja un mejor pronóstico para el conjunto de datos analizados. Basado en esto, se realizó un experimento al azar, reuniendo datos COVID reales en Colombia y apoyándonos en el método del cuadrado medio del error.

Materiales y Métodos: En este estudio se aplicaron los métodos de Media móvil simple, Media Móvil Ponderada y Suavización Exponencial aplicados en la predicción del número de casos positivos semanales de COVID-19 en Colombia en el año 2021. Método de la media móvil es una técnica para obtener el concepto general de tendencia de un conjunto de n datos en los últimos n periodos. Se puede calcular para cualquier período de tiempo futuro y cualquier subconjunto de datos. Método de la media móvil ponderado es una técnica de pronósticos que determina un valor más alto a los datos recientemente obtenidos que a los previamente vistos. Suavización exponencial: Este método se aplica cuando hay pocos datos disponibles y para pronosticar datos sin tendencia, se dice que esta técnica puede adaptar los pronósticos en forma contraria a los errores ya cometidos. Por su parte, Cuadrado medio del error (ECM) es un método que utilizamos para la medición de error que hay entre dos agrupaciones de datos, buscamos comparar los datos de nuestro pronóstico con los datos ya observados o conocidos para encontrar el margen de error de nuestro pronóstico.

Resultados: Teniendo en cuenta los resultados de los métodos de media móvil simple con ECM de 1.233.926.961,87; Media móvil ponderado con un ECM de 804.544.122,80 y Método de suavización exponencial con el valor más bajo de ECM de 360.069.095,82 por lo que podemos definir que el método de suavización exponencial es mejor método según el valor del ECM para predecir los valores de casos positivos de COVID-19 en Colombia.

Conclusiones: La aplicación de series temporales en los casos de covid-19 es uno de los mayores retos que ha traído la pandemia, ya sea por la dificultad de calcular los datos, gracias a los inesperados aumentos de contagios o los casos positivos que regularmente son indocumentados. Para esto se implementaron tres métodos de pronósticos, siendo estos: La media móvil simple, Media móvil ponderado y

Suavización exponencial en donde concluimos que el mas preciso fue el método de suavización exponencial con el valor más bajo de ECM de 360.069.095,82.

Palabras clave: Series temporales, suavización exponencial, media móvil, media móvil simple, descomposición temporal, COVID-19.

ABSTRACT

Background: A time series is defined as a chronological sequence of observations on a characteristic (univariate or scalar) or several characteristics (multivalent or vector sequence). Time series are based on the assumption that the value of the variable to be studied is the result of three components. These components are: the trend process, the seasonal component and the random component. The trend component is defined as a long-term average change. The seasonal component is related in many time series that show periodicity, that is, it changes a certain period of time and the random component does not respond to any behavior pattern, it is only the result of accidental or random factors in an isolated way in a Time series. The importance of time series has been demonstrated over time and many applications have been made to improve prediction. For example, they have been proposed to introduce statistical models to perform an optimal and efficient control of the data traffic of a Network or to generate forecasts to measure the trend and behavior of production data taking into account seasonality.

Objective: The main objective for this research is to analyze three time series methods such as: The simple moving average method, Weighted Moving Average and Exponential Smoothing applied in the prediction of the number of positive cases of COVID-19 in Colombia. The idea was to identify which method gives a better prognosis for the analyzed data set. Based on this, a random experiment was conducted, gathering real COVID data in Colombia and relying on the mean square error method.

Materials and Methods: In this study, the Simple Moving Average, Weighted Moving Average and Exponential Smoothing methods were applied in the prediction of the number of weekly positive cases of COVID-19 in Colombia in the year 2021. The moving average method is a technique to obtain the general concept of trend of a set of n data in the last n periods. It can be calculated for any future time period and any subset of data. Weighted moving average method is a forecasting technique that determines a higher value for recently obtained data than previously viewed data. Exponential smoothing: This method is applied when little data is available and to forecast data without trend, it is said that this technique can adapt the forecasts in a way contrary to the errors already made. On the other hand, Mean Square of the error (ECM) is a method that we use to measure the error that exists between two groups of data, we seek to compare the data of our forecast with the data already observed or known to find the margin of error of our forecast.

Results: Taking into account the results of the simple moving average methods with ECM of 1,233,926,961.87; Weighted moving average with an ECM of

804,544,122.80 and Exponential smoothing method with the lowest ECM value of 360,069,095.82, so we can define that the exponential smoothing method is the best method according to the value of the ECM for predict the values of positive cases of COVID-19 in Colombia.

Conclusions: The application of time series in the cases of covid-19 is one of the greatest challenges that the pandemic has brought, either due to the difficulty of calculating the data, thanks to the unexpected increases in infections or the positive cases that are regularly undocumented. For this, three forecasting methods were implemented, these being: The Simple Moving Average, Weighted Moving Average and Exponential Smoothing where we concluded that the most accurate was the exponential smoothing method with the lowest ECM value of 360,069,095.82.

KeyWords: Time series, exponential smoothing, moving average, simple moving average, time decomposition, COVID-19.

REFERENCIAS

1. C. A. MENESES AGUDO, «UNIVERSIDAD DE CANTABRIA,» 2 07 2019. [En línea]. Available: https://repositorio.unican.es/xmlui/bitstream/handle/10902/16901/TFM_Carlos_Meneses_0719.pdf?sequence=1&isAllowed=y. [Último acceso: 15 02 2020].
2. C. E. CATAN, «SERIES TEMPORALES,». [En línea]. Available: http://humanidades.cchs.csic.es/cchs/web_UAE/tutoriales/PDF/SeriesTemporales.pdf. [Último acceso: 20 2 2021].
3. J. VILLACENDO, «Introducción a Series de Tiempo,» 0 2010. [En línea]. Available: https://www.academia.edu/download/38458362/manual_intro_series_tiempo.pdf. [Último acceso: 22 02 2020].
4. M. PEPIO VIÑALS, Series Temporales, vol. 1, Barcelona: Univeritat Politecnica de Catalunya, 2001.
5. C. A. HERNANDEZ S, L. F. PEDRAZA M y A. ESCOBAR DIAZ, «APLICACIONES DE LAS SERIES DE TIEMPO EN MODELOS DE TRÁFICO PARA UNA RED DE DATOS,» Scientia et Technica Año XIV, vol. 1, n° No. 38, pp. pp. 31-36, 06 2008.
6. E. G. MONTES PAEZ, F. E. Calvete González y C. A. Mantilla Duarte, «APLICACIÓN DE SERIES DE TIEMPO EN LA REALIZACIÓN DE PRONÓSTICOS DE PRODUCCIÓN,» Revista Fuentes: El Reventón Energético, vol. Vol. 14, n° No. 1 , pp. pp. 79 - 88, 2016.
7. E. G. MONTES PAEZ, F. E. Calvete González y C. A. Mantilla Duarte, «APLICACIÓN DE SERIES DE TIEMPO EN LA REALIZACIÓN DE PRONÓSTICOS DE PRODUCCIÓN,» Revista Fuentes: El Reventón Energético, vol. 14, n° 1, pp. 79-88, 06 2016.
8. L. R. MONTES DE OCA, D. CHAVEZ ESPONDA y A. GARCIA PEREIRA, «Aplicación de las series temporales para predecir las propiedades de calidad de la piña (Ananas Comosus.), variedad Cayena Lisa, durante el

- proceso de maduración a temperatura ambiente con respecto al tiempo real,» *Revista Ciencias Técnicas Agropecuarias*, vol. 22, n° 4, pp. pp. 32-35, 2013.
9. G. A. CORRES, A. ESTEBAN , J. C. GARCIA y C. ZARATE, «ANÁLISIS DE SERIES TEMPORALES,» *Revista Ingeniería Industria*, vol. 1, n° 1, pp. pp. 21-33, 2009.
 10. A. BARRERAS SERRANO , E. SANCHEZ LOPEZ, F. FIGUEROA SAABEDRA, J. Á. OLIVAS VALDEZ y C. PEREZ LINAREZ, «Uso de un modelo univariado de series de tiempo para la predicción, en el corto plazo, del comportamiento de la producción de carne de bovino en Baja California, México,» *Veterinaria de México*, pp. pp. 1-9, 2014.
 11. J. R. ESCUDERO GARCIA y G. VALLEJO SECO , «Aplicación del diseño de series temporales múltiples a un caso de intervención en dos clases de Enseñanza General Básica,» *Psicothema*, vol. 12, n° 2, pp. pp. 203-210, 2000.
 12. J. S. FERNANDEZ, «SERIESTEMPORALE3.PDF,» [En línea]. Available: http://webpersonal.uma.es/de/J_SANCHEZ/Capitulo4.PDF. [Último acceso: 20 04 2021].
 13. M. P. VIÑALS, «series temporales,» books.google, 0 0 0. [En línea]. Available: https://books.google.es/books?hl=es&lr=&id=zEmCOBABiQ4C&oi=fnd&pg=PA11&dq=que+es+una+serie+temporales&ots=fjl8qymTPo&sig=SIT2SFire6oOEaT1b4K-_A5_vc4#v=onepage&q=que%20es%20una%20serie%20temporales&f=false. [Último acceso: 04 04 2021].
 14. G. A. PEREZ, «Series de Tiempo,» 09 2008. [En línea]. Available: <http://eio.usc.es/eipc1/BASE/BASEMASTER/FORMULARIOS-PHP/MATERIALESMATER/MATERIALESMASTER/Tema1.pdf>. [Último acceso: 19 04 2021].
 15. F. PARRA, «Estadística y Machine Learning con R,» 25 01 2019. [En línea]. Available: <https://bookdown.org/content/2274/portada.html>. [Último acceso: 27 03 2021].
 16. S. BOTERO BOTERO y J. A. CANO CANO, «ANÁLISIS DE SERIES DE TIEMPO PARA LA PREDICCIÓN DE LOS PRECIOS DE LA ENERGÍA EN LA BOLSA DE COLOMBIA,» 0 enero; junio 2008. [En línea]. Available: <redalyc.org/pdf/2821/282122032007.pdf>. [Último acceso: 25 2 2021].
 17. BIBLIOTECA DOUC UC, «Definición,» 0 enero 2018. [En línea]. Available: <http://www.duoc.cl/biblioteca/crai/definicion-y-proposito-de-la-investigacion-aplicada>. [Último acceso: 8 03 2021].
 18. J. PEREZ PORTO y M. MERINO , «metodo-inductivo,» 2012. [En línea]. Available: <https://definicion.de/metodo-inductivo/>. [Último acceso: 9 3 2021].
 19. M. SILVESTRINI RUIZ y J. VARGAS JORGE, «FUENTES DE INFORMACIÓN PRIMARIAS, SECUNDARIAS Y TERCARIAS,» 0 enero 2008. [En línea]. Available: <https://ponce.inter.edu/cai/manuales/FUENTES-PRIMARIA.pdf>. [Último acceso: 13 4 2021].
 20. M. M. RIVERA, «Fuentes de Información,» febrero 2015. [En línea]. Available: <https://repository.uaeh.edu.mx/bitstream/bitstream/handle/123456789/16700/LECT132.pdf>. [Último acceso: 12 04 2021].

21. J. PACHECO , «WEB Y EMPRESA,» 11 AGOSTO 2021. [En línea]. Available: <https://www.webyempresas.com/la-suavizacion-exponencial/>. [Último acceso: 05 septiembre 2021].
22. J. C. GARCIA, «Predicción en el dominio del tiempo,» de Analisis de series temporales, España, 2016.
23. Z. HERNANDEZ MARTIN, «UNIRIOJA,» UNIRIOJA , [En línea]. Available: https://www.unirioja.es/cu/zehernan/docencia/MAD_710/Lib489791.pdf. [Último acceso: 3 septiembre 2021].
24. Z. HERNANDEZ MERTIN, «Unirojas,» universidad de la rioja, [En línea]. Available: https://www.unirioja.es/cu/zehernan/docencia/MAD_710/Lib489791.pdf. [Último acceso: 3 septiembre 2021].
25. P. Sánchez-Sánchez, J.R. García-González, C. H. Fajardo Toro, A. Pulido-Rojano y E. Melamed-Varela, Simulación de sistemas de emergencia en salud. En: A. Pulido-Rojano, P. Sánchez-Sánchez y E. Melamed-Varela. (eds.), Nuevas tendencias en investigación de operaciones y ciencias administrativas: Un enfoque desde estudios iberoamericanos, Barranquilla, Colombia: Ediciones Universidad Simón Bolívar, 2018, pp. 165-210.
26. F. Ceballos, J.P. Betancur Villegas, y J. D. Betancur Villegas, “Simulación Discreta Aplicada a los Modelos de Atención en Salud”, Investigación e Innovación en Ingenierías, Vol. 2, No. 2, 2014, pp. 10-14, <https://doi.org/10.17081/invinno.2.2.2045>
27. D. Heredia Acevedo, Y. Fernando Ceballos y G. Sanchez Torres, “Modelo de simulación de eventos discretos para el análisis y mejora del proceso de atención al cliente”, Investigación e Innovación en Ingenierías, Vol. 8, No. 2, 2020, pp. 44 -61. <https://doi.org/10.17081/invinno.8.2.3639>
28. A.D. Pulido-Rojano, A. Verdeza-Villalobos, G. Salas-Roa, A. Muñoz-Iglesias, A. Echavarría-Collante y M. De La Rosa Romero “Simulación de un Sistema de Colas en un proceso de atención de buques en el puerto de Barranquilla-Colombia”. Vigésima Conferencia Iberoamericana en Sistemas, Cibernética e Informática (CISCI 2021), Congreso Ámbito: Internacional Realizado del 2019-07-18 al 2019-07-21 en Orlando, Florida, pp. 7 – 12.
29. A.D. Pulido-Rojano, A. Calderón-Herrera, A. Gutiérrez-Calderón, L. Mier-Olmos, M. Movilla-Palma y R. Rosales-Rodríguez, “Análisis de Casos Positivos del Covid-19 en Colombia a través de un árbol de Probabilidades”. Vigésima Conferencia Iberoamericana en Sistemas, Cibernética e Informática (CISCI 2021), Congreso Ámbito: Internacional Realizado del 2019-07-18 al 2019-07-21 en Orlando, Florida, pp. 51 – 56.