

PRONÓSTICO DE SINIESTROS VIALES EN EL ECUADOR BASADO EN DESCOMPOSICIÓN DE VALORES SINGULARES MULTINIVEL - MSVD

FORECASTING OF TRAFFIC ACCIDENTS IN ECUADOR BASED ON MULTILEVEL SINGULAR VALUE DECOMPOSITION - MSVD

Ana Elizabeth Congacha

Profesora, Universidad Nacional de Chimborazo, (Ecuador).

E-mail: acongacha@unach.edu.ec ORCID: <https://orcid.org/0000-0002-2488-9110>

Lady Espinoza

Profesora, Universidad Nacional de Chimborazo, (Ecuador).

E-mail: lespinoza@unach.edu.ec ORCID: <https://orcid.org/0000-0001-6569-3686>

Lida Barba

Profesora, Universidad Nacional de Chimborazo, (Ecuador).

E-mail: lbarba@unach.edu.ec ORCID: <https://orcid.org/0000-0003-4191-9497>

Cristian Hugo Morales

Investigador Externo Universidad Nacional de Chimborazo, (Ecuador).

E-mail: cristianmorales18m@gmail.com ORCID: <https://orcid.org/0000-0002-0197-0581>

Recepción: 04/08/2020 **Aceptación:** 24/09/2020 **Publicación:** 14/12/2020

Citación sugerida:

Congacha, A.E., Espinoza, L., Barba, L., y Morales, C.H. (2020). Pronóstico de siniestros viales en el Ecuador basado en descomposición de valores singulares multinivel - MSVD. *3C Tecnología. Glosas de innovación aplicadas a la pyme*, 9(4), 45-63. <https://doi.org/10.17993/3ctecno/2020.v9n4e36.45-63>

RESUMEN

Las instituciones y organizaciones necesitan apoyar sus tareas de planificación para mejorar la toma de decisiones, requieren de información conocida y del conocimiento no visible que contienen los datos. Uno de los principales desafíos para la mayoría de los países es disminuir la tasa de siniestros viales implementando acciones estratégicas como estudios científicos centrados en apoyar planes de prevención. En este artículo se aplica el método basado en la descomposición de valores singulares multinivel (MSVD) para mejorar la exactitud del pronóstico multihorizonte de un modelo lineal MIMO-AR (múltiple entrada-múltiple salida autorregresivo). El modelo propuesto se evalúa por medio de las series de tiempo de siniestros de tránsito en Ecuador, recolectados semanalmente entre los años 2015 a 2018, el número de personas lesionadas, ilesas, fallecidas, vehículos retenidos y conductores identificados. Los resultados muestran que el modelo alcanza una eficiencia de Nash-Sutcliffe promedio de 89.2% para un horizonte de 12 semanas, y un coeficiente de determinación R^2 promedio de 98.9% para el mismo horizonte.

PALABRAS CLAVE

MSVD, MIMO-AR, Series de tiempo, Pronóstico, Siniestros de tránsito.

ABSTRACT

Institutions and organizations need to support their planning tasks to improve decision-making, they require known information and the non-visible knowledge contained in the data. One of the main challenges for most countries is to reduce the rate of road accidents by implementing strategic actions such as scientific studies focused on supporting prevention plans. In this article, the method based on multilevel singular value decomposition (MSVD) is applied to improve the accuracy of the multi-horizon forecast of a linear MIMO-AR (multiple input-multiple output autoregressive) model. The proposed model is evaluated by means of the time series of traffic accidents in Ecuador, collected weekly between the years 2015 to 2018, the number of injured, uninjured, deceased, detained vehicles and identified drivers. The results show that the model reaches an average Nash-Sutcliffe efficiency of 89.2% for a 12-week horizon, and an average Coefficient of Determination R^2 of 98.9% for the same horizon.

KEYWORDS

MSVD, MIMO-AR, Time series, Forecast, Traffic accidents.

1. INTRODUCCIÓN

Los accidentes de tráfico provocan la muerte de aproximadamente 1,35 millones de personas en todo el mundo cada año y dejan entre 20 y 50 millones de personas con lesiones no mortales (World Health Organization, 2019). Las muertes y lesiones causadas en el siniestro vial involucran a usuarios vulnerables de la vía, como peatones, ciclistas, motociclistas y sus pasajeros. En Ecuador, según la Agencia Nacional de Tránsito (ANT) en 2018 el número de siniestros de tránsito a nivel nacional fue de 25.530 con una reducción del 12% en relación al año 2017 con 28.967 siniestros (Agencia Nacional de Tránsito, 2019). Son múltiples los factores que intervienen en los siniestros en la vía, se identifican 28 causas en el Ecuador, la causa predominante es la conducción con falta de atención a las condiciones del tránsito, con una incidencia del 56.8% (Congacha *et al.*, 2019).

En general, estos datos denotan un problema de atención pública prioritaria de interés global, la Agenda 2030 para el Desarrollo Sostenible, ha definido metas ambiciosas para conseguir una reducción del 50% de las muertes y traumatismos por accidentes de tránsito (Organización Mundial de la Salud, 2015). Por lo tanto, los pronósticos son una herramienta que proporciona un estimado cuantitativo de la probabilidad de eventos futuros. La relevancia de incorporar pronósticos en siniestros viales deriva de la gran carga social, sanitaria y económica que suponen, donde la idea principal es que los tomadores de decisiones determinen estrategias a través de la recuperación y análisis de datos históricos conocidos como series de tiempo, estableciendo una coordinación con los pronósticos (Contreras *et al.*, 2016).

Una serie de tiempo es una colección de observaciones hechas cronológicamente (Fu, 2011). En casi todos los campos científicos, las mediciones se realizan en el tiempo, estas observaciones conducen a una recopilación de datos organizados llamados series temporales (Esling y Agon, 2012). La previsión de series de tiempo ha alcanzado gran importancia en la planificación y gestión de las instituciones gubernamentales, industrias y empresas. Los algoritmos predictivos se pueden emplear con éxito en la estimación de las variables de rendimiento y la predicción de eventos críticos (Vilalta *et al.*, 2002). Las implementaciones de pronóstico son limitadas debido a la complejidad de los datos, condiciones

ambientales, variables económicas, situaciones de riesgo, entre otras, se originan en sistemas altamente dinámicos; en consecuencia, su análisis se vuelve complejo y se han obtenido resultados inexactos.

El modelado de datos permite la obtención de información válida, en este contexto los datos históricos de accidentes de tránsito constituyen valiosas fuentes de información para extraer patrones de comportamiento de alto impacto socioeconómico. El método de predicción basado en la descomposición de Valores Singulares Multinivel (MSVD) es un método inspirado en el proceso piramidal implementado en el análisis multiresolución del algoritmo Mallat que se definió para la representación de wavelets. En este método se propone la descomposición multinivel de una matriz de Hankel. (Barba y Rodríguez, 2017). El modelo AR se implementa para pronosticar las series de tiempo utilizando la estrategia múltiple entradas - múltiples salidas MIMO que se utiliza para superar el problema de acumulación de errores que se observa en la estrategia recursiva y la estrategia directa y para preservar las relaciones aleatorias entre los valores pronosticados (Barba y Rodríguez, 2017).

En esta investigación se aplica el método MSVD para la extracción de componentes de baja y alta frecuencia de una serie temporal no estacionaria. A partir de las componentes extraídas se implementan los modelos de pronóstico de horizonte amplio. MSVD ha demostrado ser eficiente en otros contextos tales como el pronóstico del CO₂ (Barba *et al.*, 2017) y pesca en Chile (Barba, 2017).

En general, aplicar modelos de pronóstico aporta con información para la toma de decisiones; buenas prácticas en materia de prevención que puedan ser implementadas por las instituciones públicas y privadas para disminuir la tasa de siniestros viales. El artículo se encuentra organizado en cuatro secciones. La sección dos presenta la metodología del pronóstico y métricas de eficiencia. La sección tres muestra los resultados obtenidos al aplicar MSVD+MIMO-AR. Finalmente, la sección cuatro concluye el documento.

2. METODOLOGÍA

La presente investigación se desarrolló utilizando un enfoque mixto, los datos proporcionados por la Dirección Nacional de Control de Tránsito y Seguridad Vial (DNT) fueron entregados en una hoja de cálculo, los cuales poseen variables de tipo cualitativo que se utilizaron para realizar un análisis descriptivo, y de tipo cuantitativo que hacen relación a los siniestros de tránsito, correspondientes a número de conductores identificados, personas fallecidas, ilesos, lesionados y vehículos retenidos.

La metodología de pronóstico se describe en dos etapas; la primera es la descomposición de valores singulares multinivel MSVD para descomponer una serie temporal en dos componentes de baja y alta frecuencia, y en la segunda etapa se realiza pronóstico a través del modelo MIMO-AR.

Descomposición de valores singulares multinivel

El método de Descomposición de Valores Singulares Multinivel (MSVD) fue planteado por Barba y Rodríguez (2017), para descomponer series de tiempo no estacionarias en componentes de baja y alta frecuencia.

MSVD implementa un proceso iterativo que consiste en el mapeo de una serie de tiempo x de longitud N en una matriz de Hankel H de orden $m \times n$, con $m=2$, y $n=N-1$.

La matriz es descompuesta mediante la descomposición de valores singulares SVD, por medio de la cual se obtienen valores singulares λ_1 y λ_2 , una matriz U y una matriz V de vectores singulares izquierdo y derecho respectivamente.

Por medio de los valores y vectores singulares se representan matrices elementales H_1 y H_2 , utilizadas para extraer las componentes de baja y alta frecuencia C_L y C_H respectivamente.

El proceso es controlado por medio del cálculo de la función tasa de espectro singular $f(\lambda R)$, en cada iteración; el algoritmo se detiene cuando la función alcanza su valor asintótico.

Modelo Autorregresivo basado en MIMO

El modelo AR se implementa para pronosticar la serie de tiempo usando la estrategia MIMO. Esta estrategia calcula el pronóstico de horizonte múltiple en un solo paso, preserva las relaciones aleatorias entre los valores históricos utilizados como predictores (Wang *et al.*, 2016). El modelo MIMO-AR se expresa de la siguiente manera:

$$[\hat{x}(n+1), \dots, \hat{x}(n+h)] = f[z(n), z(n-1), \dots, z(n-P+1)]$$

donde n es el instante de tiempo actual, h es el tamaño del horizonte de pronóstico, z es el vector regresor, y P es el tamaño del vector regresor. La siguiente ecuación define MIMO en forma de matriz:

$$\hat{X} = \beta Z^T$$

donde \hat{X} es la matriz de valores estimados, β es la matriz de coeficientes de la regresión de dimensión $h \times 2P$. La matriz Z está conformada por los valores autorregresivos las componentes de baja frecuencia y los valores autorregresivos de la componente de alta frecuencia. La matriz Z tiene dimensión $N_t \times 2P$, donde N_t es el número de muestras de entrenamiento. Los coeficientes de la regresión son estimados por medio del método de los mínimos cuadrados (LS):

$$\beta = X \times Z^\tau$$

donde τ es la matriz pseudoinversa de Moore-Penrose (Serre, 2002).

2.1. METRICAS DE EFICIENCIA

En este estudio el desempeño de los modelos MSVD + MIMO-AR se evalúa con los siguientes criterios de eficiencia: error de porcentaje absoluto medio (MAPE), error cuadrático medio (RMSE), validación cruzada generalizada (GCV), coeficiente de determinación R^2 , Eficiencia de Nash-Sutcliffe (E). En la Tabla 1 se describen las métricas.

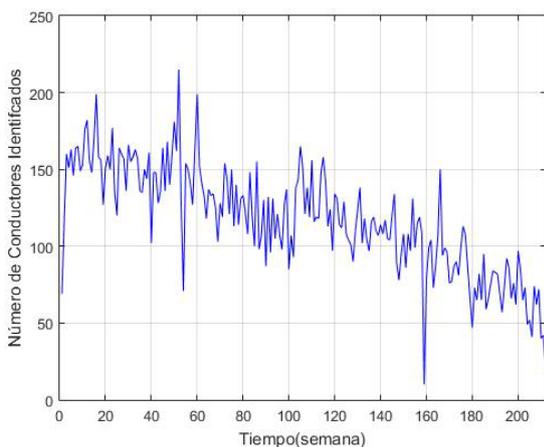
Tabla 1. Métricas de eficiencia.

MÉTRICA	FÓRMULA	DESCRIPCIÓN
<i>MAPE</i> : error de porcentaje absoluto medio	$MAPE = \left[\frac{1}{N_v} \sum_{i=1}^{N_v} \left \frac{x_i - \hat{x}_i}{x_i} \right \right] \times 100$	<p>Donde:</p> <p>N_v es el tamaño de la muestra de validación</p> <p>x_i es el ith valor observado</p> <p>\hat{x}_i es el ith valor estimado y</p> <p>K es la longitud del vector regresor de entrada (Barba, Rodríguez y Montt, 2014).</p>
<i>RMSE</i> : error cuadrático medio	$RMSE = \sqrt{\frac{1}{N_v} \sum_{i=1}^{N_v} (x_i - \hat{x}_i)^2}$	
<i>GCV</i> : validación cruzada generalizada	$GCV = \frac{RMSE}{(1 - K / N_v)^2}$	
R^2 : coeficiente de determinación	$R^2 = \left[1 - \frac{var(x - \hat{x})}{var(x)} \right] \times 100$	Donde x es la señal sobsejada, \hat{x} es la señal predicha y var es la varianza (Barba y Rodríguez, 2016).
<i>E</i> : Eficiencia de Nash-Sutcliffe	$E = 1 - \frac{\sum_{i=1}^n (O_i - P_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}$	Donde O_i es el valor observado, P_i es el valor pronosticado (Krause, Boyle, y Base, 2005).

Fuente: elaboración propia.

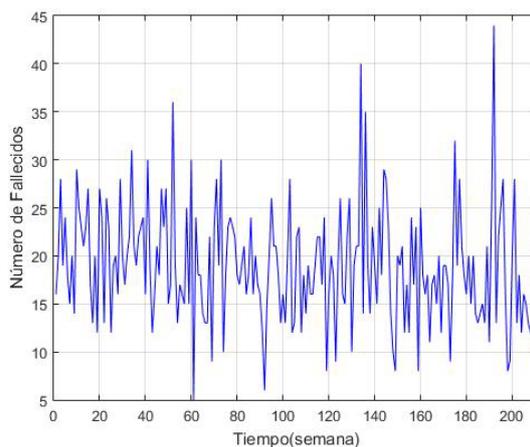
2.2. CASO DE ESTUDIO

DNT es la institución del gobierno del Ecuador, responsable de planificar, regular y controlar la gestión del transporte terrestre, tránsito y seguridad vial, a fin de garantizar la libre y segura movilidad terrestre (Dirección Nacional de Control de Tránsito y Seguridad Vial, 2019). Las zonas de su competencia corresponden a 22 de las 24 provincias del Ecuador, por tanto, las provincias que no han sido consideradas son Guayas y Santa Elena, correspondientes a la zona 8. Los datos se recopilan continuamente, para este estudio se tomaron muestras con un intervalo fijo de 7 días (una semana) entre el año 2015 y el 2018. Se utilizan cinco series de tiempo discretas de conductores identificados, fallecidos, ilesos, lesionados y vehículos retenidos en siniestros viables.



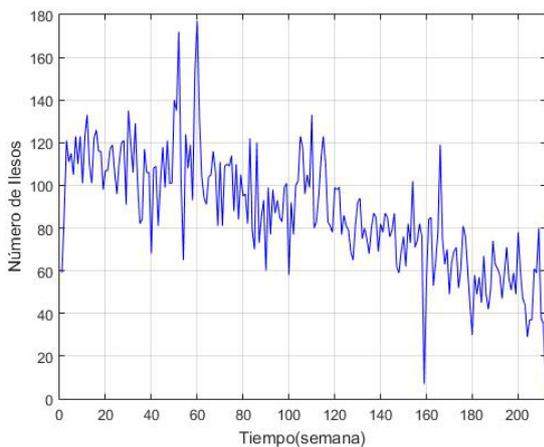
(a)

Número de conductores identificados por siniestros viales en el Ecuador, entre el año 2015 al 2018.



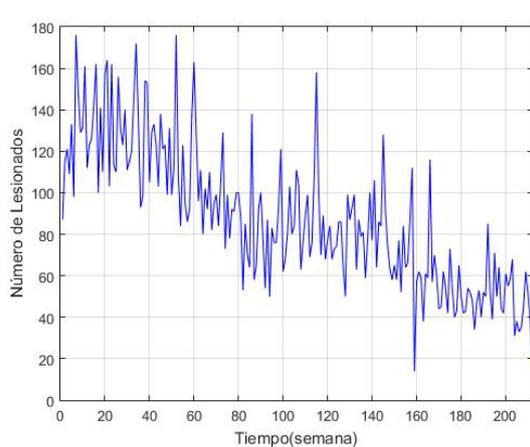
(b)

Número de personas fallecidas por siniestros viales en el Ecuador, entre el año 2015 al 2018



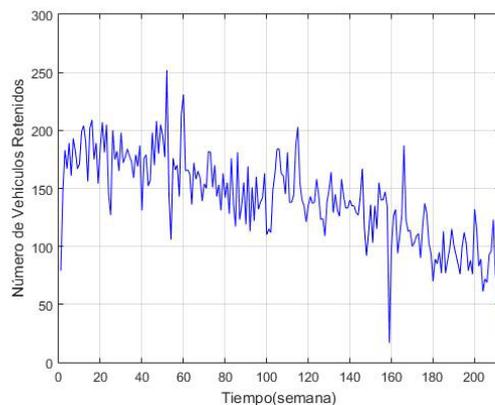
(c)

Número de personas ilesas por siniestros viales en el Ecuador, entre el año 2015 al 2018



(d)

Número de personas lesionadas por siniestros viales en el Ecuador, entre el año 2015 al 2018



(e)

Número de vehículos retenidos por siniestros viales en el Ecuador, entre el año 2015 al 2018

Figura 1. Series de tiempo observadas: (a) Conductores identificados, (b) Fallecidos, (c) Ilesos, (d) Lesionados, (e) Vehículos retenidos.

Fuente: elaboración propia.

La Figura 1 muestra una variabilidad de datos durante el período analizado en las series de tiempo observadas; por ejemplo, (a) Conductores identificados presenta una tendencia al alza entre las semanas 1 y 52, seguida de una tendencia a la baja hasta la semana 212. Por otro lado, (b) Fallecidos presenta una tendencia ascendente desde la semana 1 hasta la 52 seguida de una tendencia a la baja hasta la semana 130 seguida de una tendencia al alza hasta la 192 seguida de una tendencia a la baja hasta el final, (c) Ilesos presenta una tendencia descendente desde la semana 61 hasta el final, (d) Lesionados presenta una tendencia descendente desde la semana 53 hasta el final y (e) Vehículos retenidos presenta una tendencia a la baja a partir de la semana 53 hasta el final.

3. RESULTADOS

En esta investigación se utilizan cinco métricas para evaluar los modelos de pronóstico, en este caso, el error de porcentaje absoluto medio (MAPE), error cuadrático medio (RMSE), la validación cruzada generalizada (GCV), el coeficiente de determinación R^2 y la Eficiencia de Nash-Sutcliffe (E).

3.1. APLICACIÓN DE MÉTRICAS DE EFICIENCIA

Se observa un buen ajuste entre cifras reales y estimadas. El cálculo de métricas MAPE, RMSE, GCV, R2, E, nos permite conocer que el modelo implementado en horizonte 12 en las series observadas es mayor al 95% en la métrica R2.

Tabla 2. Resultados del pronóstico MSVD + MIMO-AR: conductores identificados, fallecidos e ilesos.

h (semana)	CONDUCTORES IDENTIFICADOS					FALLECIDOS					ILESOS				
	MAPE (%)	RMSE (%)	GCV (%)	R2 (%)	E (%)	MAPE (%)	RMSE (%)	GCV (%)	R2 (%)	E (%)	MAPE (%)	RMSE (%)	GCV (%)	R2 (%)	E (%)
1	0,001	0,001	0,001	100,000	99,996	0,002	0,000	0,001	100,000	99,993	0,005	0,003	0,006	100,000	99,979
2	0,002	0,002	0,004	100,000	99,989	0,001	0,000	0,000	100,000	99,997	0,006	0,003	0,007	100,000	99,973
3	0,011	0,008	0,018	100,000	99,949	0,006	0,001	0,002	100,000	99,981	0,008	0,004	0,008	100,000	99,969
4	0,001	0,001	0,001	100,000	99,997	0,003	0,000	0,001	100,000	99,990	0,008	0,004	0,010	100,000	99,966
5	0,015	0,011	0,023	100,000	99,936	0,004	0,001	0,002	100,000	99,984	0,012	0,007	0,014	100,000	99,952
6	0,013	0,011	0,024	100,000	99,947	0,017	0,003	0,007	100,000	99,942	0,025	0,014	0,030	100,000	99,908
7	0,027	0,025	0,053	100,000	99,888	0,053	0,011	0,023	100,000	99,818	0,033	0,024	0,051	100,000	99,866
8	0,102	0,074	0,158	99,999	99,665	0,147	0,031	0,067	99,998	99,477	0,116	0,061	0,131	99,999	99,628
9	0,253	0,186	0,396	99,993	99,129	0,391	0,081	0,171	99,984	98,678	0,284	0,160	0,340	99,991	98,980
10	0,568	0,435	0,926	99,963	97,939	0,892	0,188	0,400	99,916	96,936	0,687	0,373	0,795	99,954	97,636
11	1,066	0,928	1,977	99,833	95,646	2,052	0,398	0,848	99,627	93,514	1,334	0,798	1,699	99,791	94,997
12	2,170	1,810	3,855	99,337	91,225	3,969	0,771	1,642	98,693	88,038	2,654	1,555	3,312	99,210	90,278
Min	0,001	0,001	0,001	99,337	91,225	0,001	0,000	0,000	98,693	88,038	0,005	0,003	0,006	99,210	90,278
Max	2,170	1,810	3,855	100,000	99,997	3,969	0,771	1,642	100,000	99,997	2,654	1,555	3,312	100,000	99,979
Media	0,352	0,291	0,620	99,927	98,609	0,628	0,124	0,264	99,851	98,029	0,431	0,251	0,534	99,912	98,428

Fuente: elaboración propia.

Tabla 3. Resultados del pronóstico MSVD + MIMO-AR: lesionados y vehículos retenidos.

h (semana)	LESIONADOS					VEHICULOS RETENIDOS				
	MAPE (%)	RMSE (%)	GCV (%)	R2 (%)	E (%)	MAPE (%)	RMSE (%)	GCV (%)	R2 (%)	E (%)
1	0,009	0,005	0,010	100,000	99,964	0,006	0,006	0,013	100,000	99,970
2	0,013	0,007	0,014	100,000	99,946	0,015	0,015	0,031	100,000	99,925
3	0,017	0,009	0,019	100,000	99,926	0,001	0,001	0,002	100,000	99,995
4	0,015	0,008	0,017	100,000	99,937	0,006	0,006	0,013	100,000	99,970
5	0,017	0,009	0,020	100,000	99,931	0,017	0,016	0,035	100,000	99,921
6	0,013	0,009	0,019	100,000	99,947	0,023	0,023	0,050	100,000	99,897
7	0,043	0,029	0,061	100,000	99,828	0,025	0,031	0,067	100,000	99,882
8	0,146	0,083	0,176	99,998	99,507	0,085	0,090	0,193	99,999	99,659
9	0,352	0,214	0,455	99,985	98,686	0,229	0,233	0,496	99,992	99,079
10	0,795	0,501	1,067	99,919	96,892	0,520	0,547	1,166	99,956	97,806
11	1,697	1,071	2,281	99,627	93,177	0,938	1,164	2,480	99,798	95,264
12	3,213	2,093	4,459	98,441	86,043	2,166	2,254	4,800	99,230	90,642

h (semana)	LESIONADOS					VEHICULOS RETENIDOS				
	MAPE (%)	RMSE (%)	GCV (%)	R2 (%)	E (%)	MAPE (%)	RMSE (%)	GCV (%)	R2 (%)	E (%)
Min	0,009	0,005	0,010	98,441	86,043	0,001	0,001	0,002	99,230	90,642
Max	3,213	2,093	4,459	100,000	99,964	2,166	2,254	4,800	100,000	99,995
Media	0,528	0,336	0,716	99,831	97,815	0,336	0,366	0,779	99,915	98,501

Fuente: elaboración propia.

Tabla 4. Resumen de métricas.

SERIES	METRICAS				
	MAPE	RMSE	GCV	R2	E
CONDUCTORES IDENTIFICADOS	0,352	0,291	0,62	99,93	98,61
FALLECIDOS	0,628	0,124	0,264	99,85	98,03
ILESOS	0,431	0,251	0,534	99,91	98,43
LESIONADOS	0,528	0,336	0,716	99,83	97,82
VEHICULOS RETENIDOS	0,336	0,366	0,779	99,92	98,50

Fuente: elaboración propia.

Los resultados obtenidos al aplicar MSVD+MIMO-AR se presentan divididos en dos etapas: descomposición y predicción.

Descomposición. La descomposición basada en MSVD aplica un proceso iterativo que concluye cuando la función ΔR , correspondiente a la tasa de espectro singular, calculada en cada descomposición. La función alcanza el valor asintótico, $\Delta R \approx I$, para este caso, se alcanza en la repetición 20, por lo tanto, el proceso iterativo termina con la iteración 20.

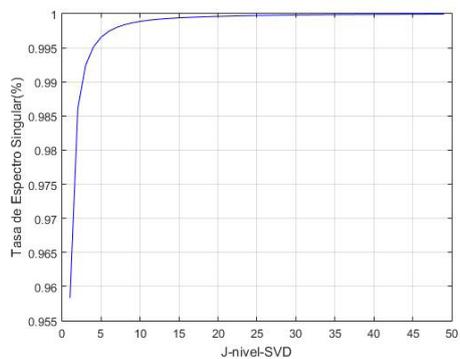
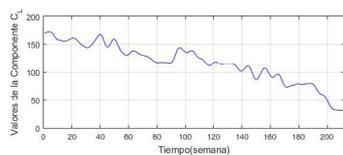
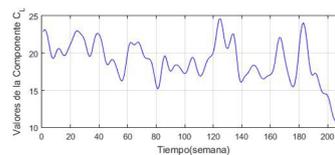


Figura 2. Niveles de descomposición versus Tasa de Espectro Singular.

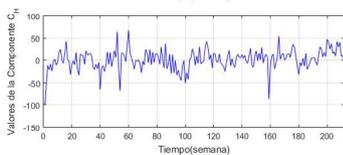
Fuente: elaboración propia.



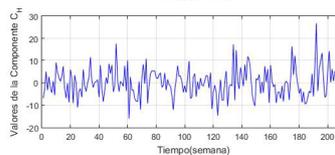
(a)



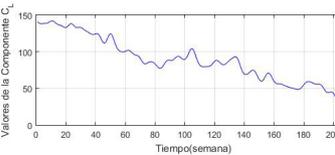
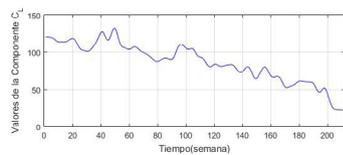
(b)

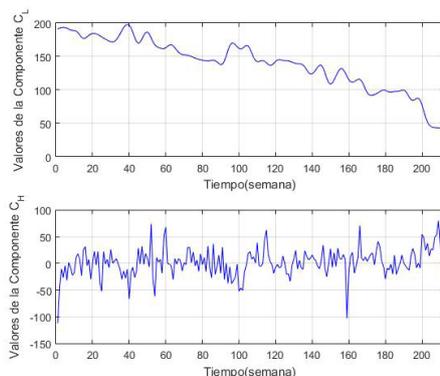


(c)



(d)





(e)

Figure 3. Componente de baja y alta frecuencia: (a) Conductores identificados, (b) Fallecidos, (c) Ilesos, (d) Lesionados, (e) Vehículos retenidos.

Fuente: elaboración propia.

La Figura 3 muestra los componentes de baja y alta frecuencia obtenidos con MSVD para las cinco series de tiempo de conductores identificados, fallecidos, ilesos, lesionados y vehículos retenidos en siniestros viales, respectivamente. Los componentes C_L extraídos muestran fluctuaciones de larga duración, mientras que los componentes C_H muestran fluctuaciones de corta duración.

Predicción. La estrategia MIMO se implementa para predecir las cinco series de estudio, con horizonte múltiple mediante el modelo autorregresivo. Las entradas son los P valores correspondientes a la ventana de tiempo elegida para el modelo AR, a través de un proceso de selección por medio de pruebas de repetición evaluadas con la métrica GCV. Los componentes C_L y C_H que fueron extraídas previamente con MSVD son las entradas del modelo AR. Cada conjunto de datos de baja y alta frecuencia se han dividido en dos subconjuntos, experimentación o entrenamiento y validación. El subconjunto de entrenamiento (N_{tr}) involucra el 70% de las muestras, el subconjunto de validación (N_v) implica 30% restante. La exactitud del pronóstico se evalúa con las métricas MAPE, RMSE, GCV, R^2 y E, presentadas en las Tablas 2 y 3. Las curvas observadas y pronosticadas para el máximo horizonte de 12 semanas, se presentan en la Figura 4.

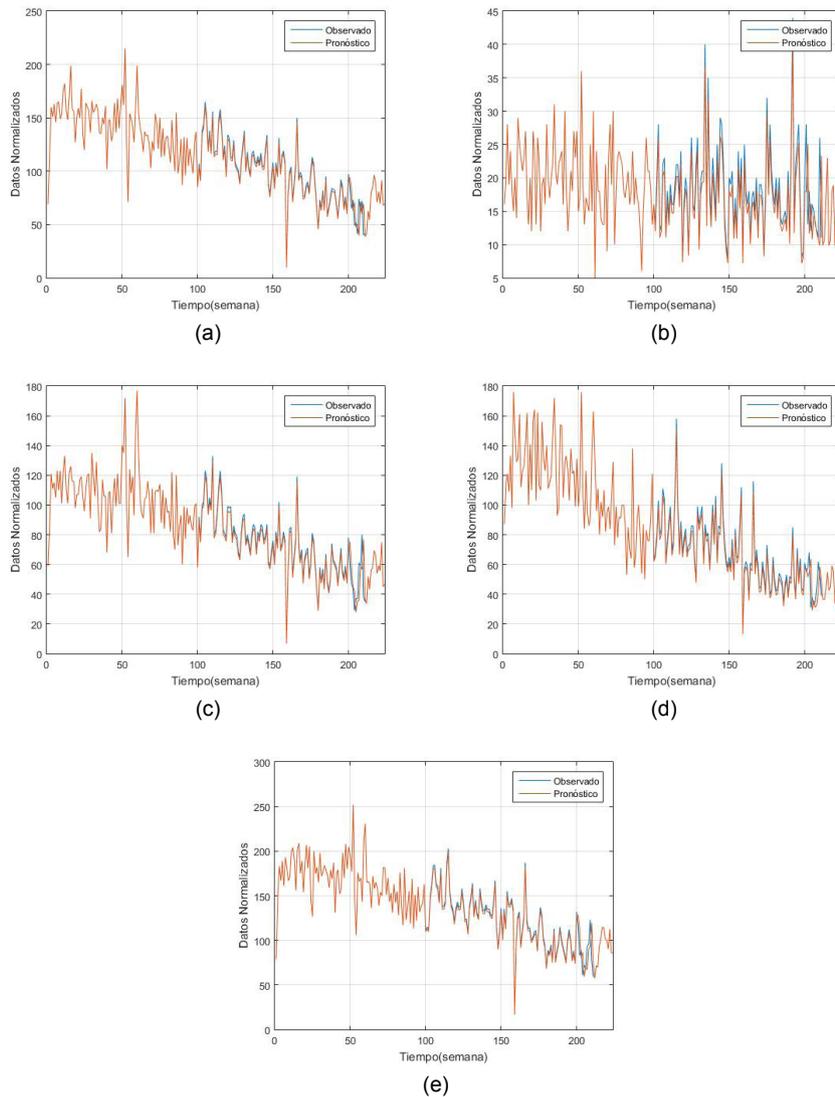


Figure 4. Series de tiempo proyectadas (a) Conductores identificados, (b) Fallecidos, (c) Ilesos, (d) Lesionados, (e) Vehículos retenidos.

Fuente: elaboración propia.

La Figura 4 muestra las curvas de valores observados versus valores pronosticados de cada una de las series de tiempo analizadas. A partir de las cifras se observa un buen ajuste entre los valores reales y estimados. En todos los casos se ha realizado una proyección de 12 semanas, es decir desde la semana 213 a la 224. Las proyecciones muestran un decrecimiento en las figuras 4(a), 4(c), 4(d) y 4(e). Por otro lado, las proyecciones muestran un comportamiento constante en la Figura 4(b) Fallecidos.

4. CONCLUSIONES

En el presente artículo se evidencia la importancia de realizar pronóstico para un horizonte amplio, logrado por medio de un modelo convencional, lineal, AR, mismo que a diferencia de los métodos no lineales presenta baja complejidad computacional. En etapas tempranas y antes de aplicar el modelo de pronóstico se aplica un método de descomposición de las series de tiempo, basado en la descomposición de valores singulares en múltiples pasos, es decir es de tipo iterativa, denotada como MSVD.

Por medio del método MSVD se logran extraer componentes de alta y baja frecuencia, mismas que representan las fluctuaciones del fenómeno de siniestros de tránsito, aquellas de mayor duración representadas en la componente de baja frecuencia, mientras que aquellas que tienen duración corta están representadas por la componente de alta frecuencia.

El modelo MIMO-AR fue implementado para cada serie de tiempo: conductores identificados, fallecidos, ilesos, lesionados y vehículos retenidos encontrando que el modelo es eficiente para un horizonte de 12 semanas, alcanzando una eficiencia promedio E de 89,2%, un coeficiente de determinación promedio R^2 de 98,9%, un error cuadrático medio RMSE de 1,7%, un error de porcentaje absoluto medio MAPE de 2,8%, y una validación cruzada generalizada promedio GCV de 3,6%.

Los resultados obtenidos permitirán a las instituciones de gobierno, responsables de la planificación de estrategias minimizar el número de siniestros viales en el Ecuador y contar con información a priori del fenómeno con implicaciones sociales y económicas. En el futuro, dada la efectividad del método, se

realizarán nuevas simulaciones de pronóstico con series de tiempo recolectadas en otros periodos y otras áreas del conocimiento.

REFERENCIAS BIBLIOGRÁFICAS

- Agencia Nacional de Tránsito.** (2019). *Siniestros diciembre 2018*. <https://www.ant.gob.ec/index.php/ley-de-transparencia/ley-de-transparencia-2020/file/6096-siniestros-diciembre-2018>
- Armstrong, J., y Collopy, F.** (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, 8(1), 69–80. [https://doi.org/10.1016/0169-2070\(92\)90008-w](https://doi.org/10.1016/0169-2070(92)90008-w)
- Barba, L.** (2017). Multiscale Forecasting Models Based on Singular Values for Nonstationary Time Series. En *III Concurso Latinoamericano de Tesis de Doctorado (CLTD-CLEI)- JAIIO 46*, Córdoba. <http://sedici.unlp.edu.ar/handle/10915/64919>
- Barba, L., y Rodríguez, N.** (2016). Hybrid Models Based on Singular Values and Autoregressive Methods for Multistep Ahead Forecasting of Traffic Accidents. *Mathematical Problems in Engineering*. Article ID 2030647. <http://dx.doi.org/10.1155/2016/2030647>
- Barba, L., Rodríguez, N., y Montt, C.** (2014). Smoothing Strategies Combined with ARIMA and Neural Networks to Improve the Forecasting of Traffic Accidents. *The Scientific World Journal*. Article ID 152375. <https://doi.org/10.1155/2014/152375>
- Barba, L., y Rodriguez, N.** (2017). A Novel Multilevel-SVD Method to Improve Multistep Ahead Forecasting in Traffic Accidents Domain. *Computational Intelligence and Neuroscience*. Article ID 7951395. <https://doi.org/10.1155/2017/7951395>
- Barba, L., Machado, G., Molina, L., Congacha A., Delgado, J., y Espinoza, L.** (2017). *Forecasting of CO2 emissions based on Preprocessing Techniques*. ITISE.

- Chatfield, C., y Yar, M.** (1988). *Holt-Winters Forecasting: Some Practical Issues*. *The Statistician*, 37(2), 129. <https://doi.org/10.2307/2348687>
- Congacha, A., Barba, J., Palacios, L., y Delgado, J.** (2019). Caracterización de los Siniestros Viales en el Ecuador. *NOVASINERGLA*, 2(2), 17-29. <https://doi.org/10.37135/unach.001.04.02>
- Contreras, A., Zuñiga, C., Martínez, J., y Sánchez, D.** (2016). Análisis de series de tiempo en el pronóstico de la demanda de almacenamiento de productos perecederos. *Estudios Gerenciales*, 32(141), 387-396. <https://doi.org/10.1016/j.estger.2016.11.002>
- Dirección Nacional de Control de Tránsito y Seguridad Vial.** (2019). *Policía Nacional del Ecuador*. <https://www.policiaecuador.gob.ec/quienes-somos-direccion-nacional-de-control-de-transito-y-seguridad-vial/>
- Esling, P., y Agon, C.** (2012). Time-Series Data Mining. *ACM Computing Surveys*, 45(1), 1-34. <https://doi.org/10.1145/2379776.2379788>
- Fu, T.** (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24(1), 164-181. <https://doi.org/10.1016/J.ENGAPPAL.2010.09.007>
- Krause, P., Boyle, D. P., y Base, F.** (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 5, 89–97. <https://doi.org/10.5194/adgeo-5-89-2005>
- Makridakis, S., Andersen, A., Carbone, R., Fildes, R., Hibon, M., Lewandowski, R., Newton, J., Parzen, E., y Winkler, R.** (1982). The accuracy of extrapolation (time series) methods: Results of a forecasting competition. *Journal of forecasting*, 1(2), 111-153. <https://doi.org/10.1002/for.3980010202>
- Organización Mundial de la Salud.** (2015). *Informe sobre la situación mundial de la seguridad vial 2015*. https://www.who.int/violence_injury_prevention/road_safety_status/2015/Summary_GSRRS2015_SPA.pdf?ua=1

- Organización Mundial de la Salud.** (2017). *Seguridad de los vehículos de motor de dos y tres ruedas: Manual de seguridad vial para decisores y profesionales.* <https://apps.who.int/iris/bitstream/handle/10665/272757/9789243511924-spa.pdf?ua=1>
- Serre, D.** (2002). *Matrices: Theory and Applications.* Springer, New York, 218-220.
- Vilalta, R., Apte, C., Hellerstein, J., Ma, S., y Weiss.** (2002). Predictive algorithms in the management of computer systems. *IBM Systems Journal*, 41(3), 461 – 474. <https://doi.org/10.1147/sj.413.0461>
- Wang, J., Song, Y., Liu, F., y Hou, R.** (2016). Analysis and application of forecasting models in wind power integration: a review of multi-step-ahead wind speed forecasting models. *Renewable and Sustainable Energy Reviews*, 60, 960–981. <https://ideas.repec.org/a/eee/rensus/v60y2016icp960-981.html>
- World Health Organization.** (2015). *Global status report on road safety 2015.* Geneva. https://www.who.int/violence_injury_prevention/road_safety_status/2015/es/
- World Health Organization.** (2019, 8 de agosto). Road traffic injuries. https://www.who.int/health-topics/road-safety#tab=tab_1