

Tasa de mortalidad por causa externa en Costa Rica: Un análisis de la serie temporal 2000 - 2017

César Gamboa Sanabria^{*,a}

^a*Escuela de Estadística, Universidad de Costa Rica*

RESUMEN

La tasa de mortalidad por causa externa es un indicador demográfico utilizado para analizar las defunciones que tuvieron lugar debido a accidentes, homicidios, suicidios o condiciones semejantes. El presente artículo analiza la tasa de mortalidad por causa externa (TMCE) desde Enero del año 2000 a Diciembre del año 2017 mediante métodos de series temporales. Se utilizó una descomposición aditiva de la serie debido a los cambios en la poca variancia que presenta la misma a lo largo del tiempo. Se ajustaron distintos modelos de regresión, suavizamiento exponencial, un $ARIMA(2, 0, 1)(3, 0, 3)_{12}$ identificado manualmente y un $ARIMA(3, 1, 1)(0, 0, 1)_{12}$ de forma automática respectivamente; los cuales fueron seleccionados para su posterior comparación. El modelo final para realizar el pronóstico es el $ARIMA(2, 0, 1)(3, 0, 3)_{12}$ debido a que presenta mejores medidas de rendimiento (MAE, MAPE, RMSE y MASE) y también de las mejores estimaciones en cuanto a los criterios de información se refiere (AIC, AICc, BIC); dicho pronóstico para los siguientes 12 del 2018 periodos con un 95 % de confianza abarca valores mínimos de 3.92, mientras que para el límite superior hasta 4.44.

Palabras clave: mortalidad, causa externa, R, ARIMA, RStudio, homicidios, suicidios, accidentes, tránsito

1. INTRODUCCIÓN

La violencia es un acto tan antiguo como el mundo, sin embargo, la evolución de esta en conjunto con el crecimiento de su relación con las defunciones registradas en una población la vuelven un problema de salud pública. En base a la clasificación Internacional de Enfermedades (OPS (2016)) de la Organización Mundial de la Salud¹, las defunciones pueden clasificarse en cuatro grandes grupos, siendo el más importante el de las causas naturales, el cual incluye enfermedades congénitas, cardiopatías u otras relacionadas con la vejez. En menor cuantía se encuentran las causas de muerte ignoradas, las cuales se dan cuando la causa de muerte es desconocida y de intención indeterminada; y de forma similar se encuentran las causas de muerte que se mantienen en estudio, bien sea por parte de la morgue o de algún otro organismo, esta última tiene pocos registros conforme más se retrocede en el tiempo.

El otro gran grupo, aunque considerablemente menor que las causas naturales, son las causas externas, las cuales son objeto de análisis en la presente investigación. En este grupo puede a su vez ser clasificado en cuatro categorías, que son los homicidios, los suicidios y las muertes accidentales, esta última comprende los accidentes de tránsito, las muertes por caídas, personas ahogadas, víctimas de incendios, terraplenes u otros similares. Aunado a estas categorías se encuentran también las causas indeterminadas, las cuales se diferencian a las ignoradas en que se sabe que se debe a una causa externa pero no se conoce con certeza a cuál categoría pertenece o aún está en investigación, tal es el caso de una persona que fallece debido a una alta ingesta de drogas o estupefacientes; bien pudo haber consumido intencionalmente hasta morir, lo cual sería un suicidio, o bien el consumo excesivo se debió a un accidente.

*email: cesar.gamboasanabria@ucr.ac.cr | LinkedIn: <https://linkedin.com/in/cgamboasanabria> | Github: <https://github.com/cgamboasanabria>

¹<https://www.paho.org/salud-en-las-americas-2017/?lang=es>

En Costa Rica para el año 2011, las muertes por causas externas ocuparon el tercer lugar, siendo solo superadas por las enfermedades del sistema circulatorio, en particular las enfermedades cardiovasculares, y los tumores, ambos casos mostraron una tendencia ascendente (Nación 2013). Es debido a los elevados costos económicos y sociales (Cardona 2013) que se aborda la imperiosa necesidad comprender el comportamiento de las defunciones debido a las causas externas con el fin de contar con un punto de partida para la elaboración de políticas públicas que busquen reducir al mínimo este tipo de eventos.

2. METODOLOGÍA

Los datos utilizados para esta investigación fueron obtenidos en la Unidad de Estadísticas Demográficas del INEC², y corresponde a información mensual de la tasa de mortalidad por causas externas registradas anualmente desde Enero del año 2000 hasta Diciembre del año 2017. Estos datos serán procesados mediante el lenguaje R³ y su interfaz RStudio⁴. Para la manipulación de los datos se utilizarán los paquetes de *R readxl* (Wickham and Bryan 2018), *tidyr* (Wickham and Henry 2018), y *dplyr* (Wickham et al. 2018), mientras que para el manejo de datos temporales se utilizó el paquete *lubridate* (Grolemund and Wickham 2011) y *forecast* (Hyndman R 2018) para realizar los pronósticos. En cuanto a la visualización, los paquetes *ggplot2* (Wickham 2016), *ggpmisc* (Aphalo 2016) y *ggseas* (Ellis 2018) serán empleados para generar distintos tipos de gráficos.

El método de análisis consiste en analizar el comportamiento de las defunciones registradas debido a alguna causa externa, para lo cual se realizará un análisis descriptivo de la serie y, posteriormente, una descomposición de la misma para estudiar sus componentes. La serie completa será particionada en dos conjuntos, uno servirá de entrenamiento para la construcción de los modelos para ajustar la serie y otro servirá de validación una vez que el modelo sea generado, esto con el fin de conocer qué tan bien se ajustan las estimaciones a los datos reales; siendo los datos correspondientes a los periodos de Enero 2000 - Diciembre 2016 el conjunto de entrenamiento y de Enero 2017 - Diciembre 2017 para la validación. Se estimarán distintas variantes de los modelos de regresión, de suavizamiento exponencial y ARIMA.

En el caso de los modelos de regresión, se estimará uno solo con la tendencia, otro con la tendencia y la estacionalidad, y un tercero que consiste en una variante exponencial, todos ellos estimados mediante la máxima verosimilitud maximizando los coeficientes respecto a los parámetros de la función expresada en (1).

$$\ln(L) = -\frac{T}{2}\ln(2\pi) - \frac{T}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(Y - X\beta)'(Y - X\beta) \quad (1)$$

Para los modelos de suavizamiento exponencial, se estimarán distintas variantes considerando errores aditivos y multiplicativos, sin tendencia y con tendencia aditiva y multiplicativa, y sin estacionalidad y con estacionalidad aditiva y multiplicativa, todo esto además de considerar el amortiguamiento.

La estimación de los modelos ARIMA se abordará de dos maneras: 1) identificando los modelos de forma manual tanto en su parte estacional como en la no estacional y 2) de forma automática mediante la función `auto.arima()`.

Una vez estimados estos modelos, se seleccionarán los mejores de cada clase, es decir, el mejor modelo de regresión, el mejor de suavizamiento exponencial, el mejor ARIMA identificado y el mejor ARIMA obtenido de forma automática. Dicha selección se hará analizando los criterios de información indicados en las ecuaciones (2), (3) y (4) donde los valores más bajos indican cuál es el mejor modelo, mientras que las medidas de rendimientos utilizadas se especifican en las ecuaciones (5), (6), (7) y (8) donde las más cercanas a cero sirven para indicar cuál modelo es el más adecuado,

$$AIC = -2\log L(\hat{\theta}) + 2k \quad (2)$$

²<http://www.inec.go.cr/>

³<https://cran.r-project.org/>

⁴<https://www.rstudio.com/>

$$AICc = -2\log L(\hat{\theta}) + 2k + \frac{2k+1}{n-k-1} \quad (3)$$

$$BIC = -2\log L(\hat{\theta}) + k \cdot \log(n) \quad (4)$$

$$MAE = n^{-1} \sum_{t=1}^n |y_t - f_t| \quad (5)$$

$$RMSE = \sqrt{n^{-1} \sum_{t=1}^n (y_t - f_t)^2} \quad (6)$$

$$MAPE = 100n^{-1} \sum_{t=1}^n \frac{|y_t - f_t|}{|y_t|} \quad (7)$$

$$MASE = n^{-1} \sum_{t=1}^n \frac{|y_t - f_t|}{q}; q = (n-m)^{-1} \sum_{t=2}^n |y_t - y_{t-1}| \quad (8)$$

donde k es el número de parámetros y n es el número de datos.

Finalmente se escogerá el modelo más adecuado en relación a tasa de mortalidad por causa externa (TMCE) para posteriormente realizar un pronóstico de los posibles valores que alcanzará este indicador en el futuro, así como un contextualización del fenómeno con respecto a la actualidad.

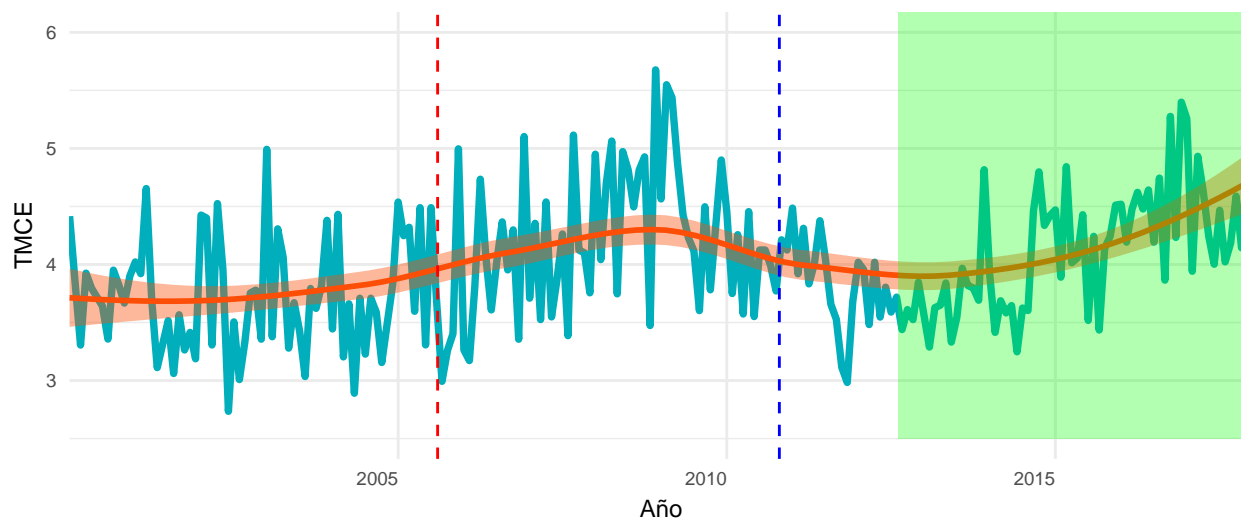
3. RESULTADOS

Como fue mencionado en la introducción, un alto porcentaje de las defunciones registradas se deben, como es de esperar, a las causas naturales; sin embargo la mortalidad debido a causas externas representa un problema de salud pública que debe ser solventado. Conforme la sociedad crece, también se vuelve más complicado manejar situaciones de seguridad ciudadana por parte del Estado que a la postre terminan en la disminución de la población debido a los crímenes asociados a la delincuencia, defunciones que presentan un alto porcentaje desde hace más de dos décadas (Mata 1995). A lo largo de los años se han planteado diversas reformas para reducir la delincuencia, los accidentes en carretera y otros aspectos relacionados, y es en base a estas condiciones que se analizan este tipo de muertes durante los últimos 18 años.

Dado que los registros de defunciones por causa externa se realizan diariamente, conviene analizar su comportamiento de manera mensual desde inicios del milenio de una manera más general, dicho comportamiento puede observarse en el gráfico 1.

La serie muestra picos y valles pronunciados a lo largo de todo el periodo. A modo de visualización, se ajustó un suavizamiento de Loess, o regresión local, para buscar señales de tendencia y concavidad en los datos temporales. La línea roja punteada se ubica aproximadamente en el mes de Junio del año 2008, pues a partir de ese punto la regresión local muestra un ligero cambio en la concavidad, lo cual sugiere el inicio de un aumento en las defunciones que culmina con una caída en Noviembre del año 2011 (línea azul punteada) para posteriormente iniciar otro cambio en la concavidad, que significa un rápido ascenso en la cantidad de muertes debido a causas externas.

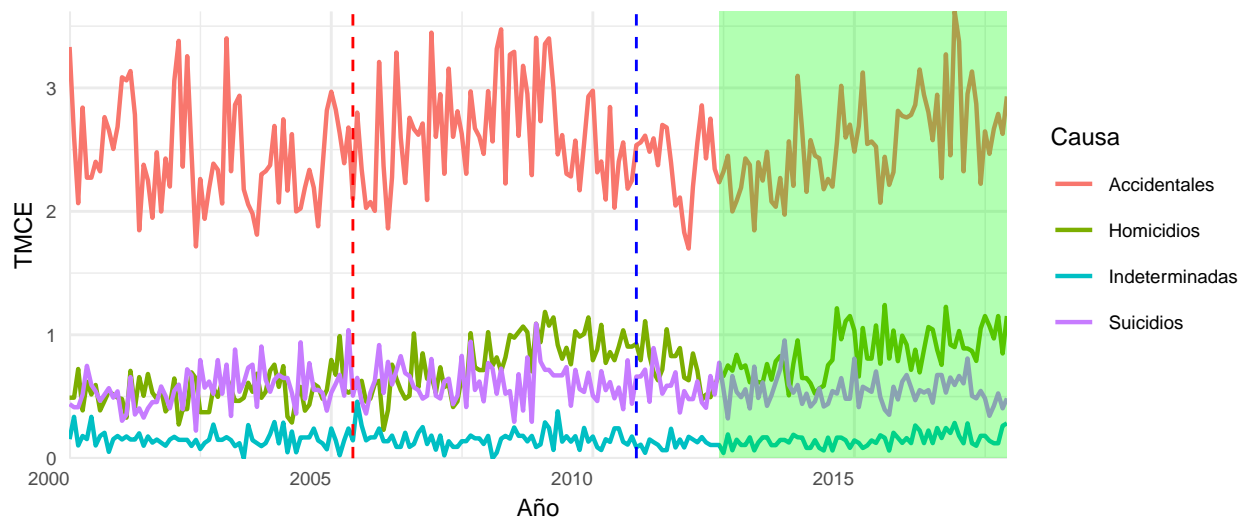
Gráfico 1: Mortalidad por causa externa 2000 – 2017



Fuente: UED-INEC

Es importante recalcar que, entre Junio del año 2012 y Diciembre del año 2017 (área verde), el aumento en la tasa de cambio de la cantidad de defunciones debido a causas externas coincide con el aumento de la flotilla de motocicletas, pues en un período de cinco años esta cifra creció en un 189 % (Vázquez 2017). Conviene entonces verificar el comportamiento a lo interno de la serie en referencias a las categorías de las causas externas.

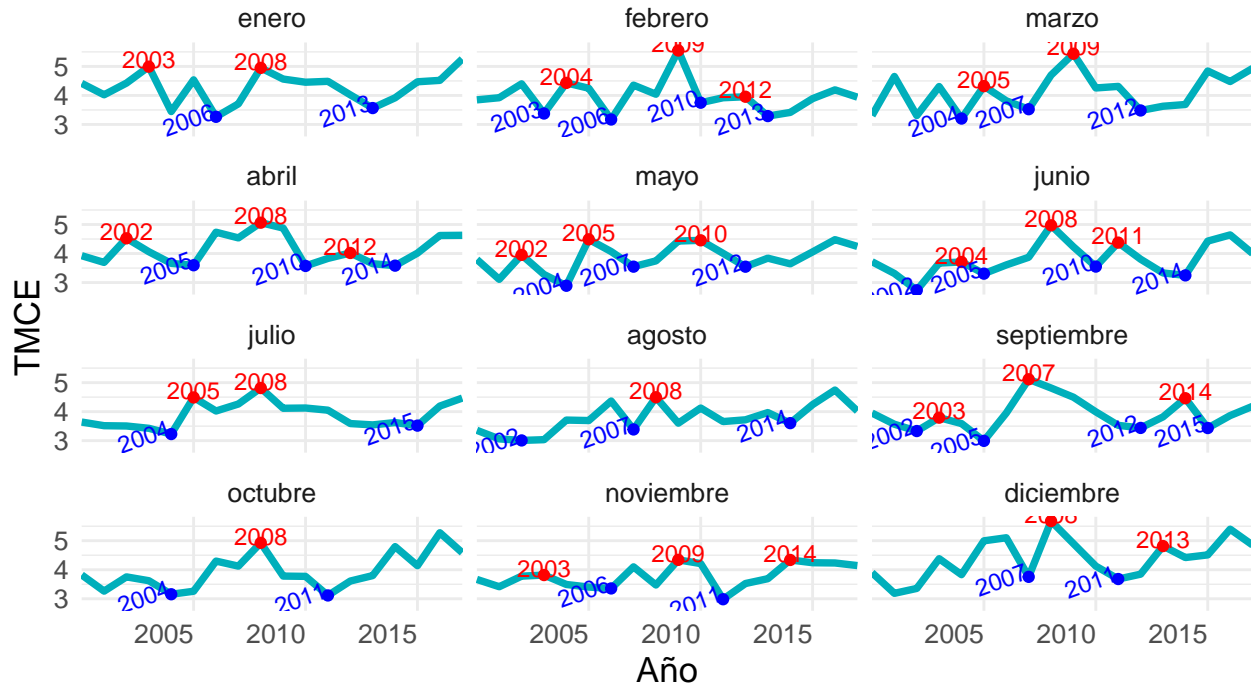
Gráfico 2: Mortalidad por causa externa según categoría 2000 – 2017



Fuente: UED-INEC

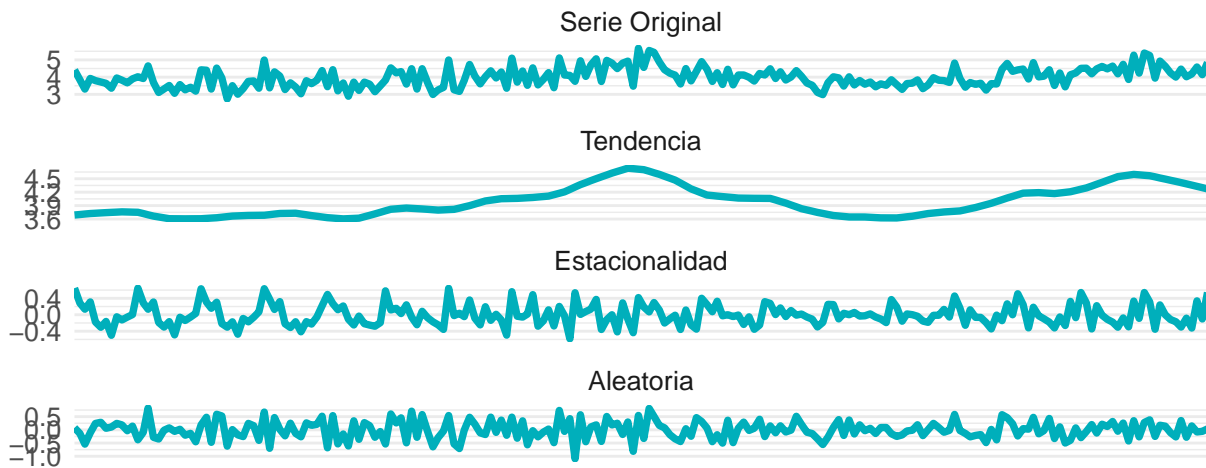
El cambio descrito en base a lo mostrado en el gráfico 1 puede explicarse mediante el gráfico 2, pues el aumento mostrado en el área verde correspondiente al periodo Junio 2012 - Diciembre 2012 coincide en ambos gráficos, por lo que dicho descenso se debe a un fuerte cambio en las muertes por causas accidentales, posiblemente ocasionado por el aumento de la cantidad de motocicletas en carretera. Conociendo este hecho, resulta importante estudiar el comportamiento de la serie para cada uno de los meses durante todo el periodo.

Gráfico 3: Mortalidad por causa externa 2000 – 2017 según mes



Fuente: UED-INEC

Gráfico 4: Descomposición de las defunciones por causa externa en el periodo 2000 –



Fuente: UED-INEC

Del gráfico 3 puede notarse que cada mes tiene sus picos y valles durante cada mes a lo largo del periodo, siendo los meses de Enero, Abril y Diciembre los que presentaron valores ligeramente más altos entre los años 2000 y 2017.

La descomposición de la serie se hará de forma aditiva debido a que en el gráfico 1 no se observan grandes cambios en la variabilidad a lo largo del tiempo. El gráfico 3 muestra que la tendencia se mantiene casi constante a lo

largo del tiempo, mientras que parece haber estacionalidad en ciertos lapsos de la segunda mitad del año. Además, el componente aleatorio muestra como los errores no son constantes a lo largo de todo el período.

Tras tener un panorama más claro del comportamiento de la serie mediante el análisis descriptivo anterior, es importante realizar las estimaciones de los modelos. Para las estimación de los mismo y todos los aspectos relacionados, se dividió la base de datos en dos subconjuntos: 1) la base de entrenamiento son los datos desde el año 2000 y hasta el 2016 inclusive, y 2) la base de validación, que corresponde a los datos del año 2017.

La primera estimación hecha es la que corresponde a los modelos de regresión. Las tres variantes descritas en el apartado metodológico de este artículo calculadas para luego contrastarlos mediante las medidas de rendimiento, los resultados pueden verse en cuadro 3 de la sección de anexos. El mejor modelo en este caso es el de variante exponencial, expresado mediante la ecuación (9), pues es el que cuenta con mejores medidas en el conjunto de entrenamiento y es muy competente en el conjunto de validación.

$$\begin{aligned} \log(\hat{y}_t) = & 1,3699 + 0,0007t - 0,0597s_2 - 0,0411s_3 - 0,0269s_4 - 0,0968s_5 - 0,1132s_6 \\ & - 0,0913s_7 - 0,1242s_8 - 0,0888s_9 - 0,0861s_{10} - 0,1184s_{11} + 0,0161s_{12} \end{aligned} \quad (9)$$

El segundo gran grupo de estimaciones corresponde a los modelos de suavizamiento exponencial. Para todas las variantes mencionadas en el apartado metodológico, se calcularon las medidas de rendimiento y los criterios de información para cada uno de los modelos, dichos resultados pueden verse en el cuadro 4 presente en la sección de anexos. Tras esta comparación, el mejor modelo de suavizamiento exponencial es que tiene errores multiplicativos, sin tendencia y con estacionalidad aditiva descrito en la ecuación (10), pues presenta mejores medidas de rendimiento en el set de validación y medidas cercanas al mínimo en el set de entrenamiento.

$$\hat{y}_{t+h|t} = l_t + s_{t+h-m(k+1)} \quad (10)$$

Donde $l_t = 0,1742(y_t - s_{t-m}) + (1 - 0,1742)l_{t-1}$; $s_t = 0,0003(y_t - l_{t-1}) + (1 - 0,0003)s_{t-m}$.

Antes de estimar los modelos ARIMA es necesario realizar una inspección en base a los autocorrelogramas, y para esto es necesario verificar la estacionaridad de la serie, es decir, que el promedio y la varianza de la serie se mantengan a lo largo del tiempo. Con este fin es que se analiza el gráfico 5, en el cual se puede ver que para volver la serie estacionaria es necesario aplicar el logaritmo y hacer una doble diferenciación cada 12 meses. El gráfico 6 muestra los autocorrelogramas parcial y total para la serie estacionaria, el cual parece sugerir que se tiene un MA(1) en la parte estacional, sin embargo ante la poca claridad de este criterio conviene realizar una sobreparametrización para encontrar el modelo más adecuado.

Para el proceso de sobreparametrización, se hace uso de la función `sarima()`, en la cual ese especifican los valores tanto en la parte estacional como en la no estacional, y a partir de esto se obtienen las probabilidades asociadas a los coeficientes y se descartan aquellos modelos cuyos coeficientes no sean significativamente distintos de cero. Al ser este un proceso iterativo, se llegaron a estimar 177 modelos ARIMA, de los cuales el más adecuado es el $ARIMA(2, 0, 1)(3, 0, 3)_{12}$ mostrado en la ecuación (11). Los resultados de los mejores 5 modelos ARIMA pueden encontrarse en el cuadro 5 en la sección de anexos.

$$\begin{aligned} y_t = & 4,0043 + 0,5421y_{t-1} + 0,3193y_{t-2} - 0,8346y_{t-12} + 0,9826y_{t-24} + 0,8413y_{t-36} \\ & - 0,5033y_{t-1} + 1,0054y_{t-12} - 0,9267y_{t-24} - 0,9733y_{t-36} + u_t \end{aligned} \quad (11)$$

La última estimación de los modelos corresponde ejecutado mediante la función `auto.arima()`. En este caso, el modelo sugerido es un $ARIMA(3, 1, 1)(0, 0, 1)_{12}$, el cual se muestra en la ecuación (12).

$$y_t = -0,0457y_{t-1} + 0,2462y_{t-2} + 0,0390y_{t-3} - 0,8904y_{t-1} + 0,1164y_{t-12} + u_t \quad (12)$$

Los diagnósticos del modelo identificado como $ARIMA(2, 0, 1)(3, 0, 3)$ muestran que tanto el gráfico de residuos estandarizados como el ACF sugieren que, debido a su comportamiento, los errores son de ruido blanco. Además, el gráfico QQ parece indicar que estos errores siguen una distribución Normal. El gráfico de la prueba Ljung-Box, sin embargo, sugiere que existe algo que no se está considerando en el modelo, pues hasta el rezago 21 los valores p caen debajo de la línea azul punteada, luego de este se mantienen siempre por encima. A pesar de esto, se seguirá utilizando este modelo como el mejor ARIMA identificado.

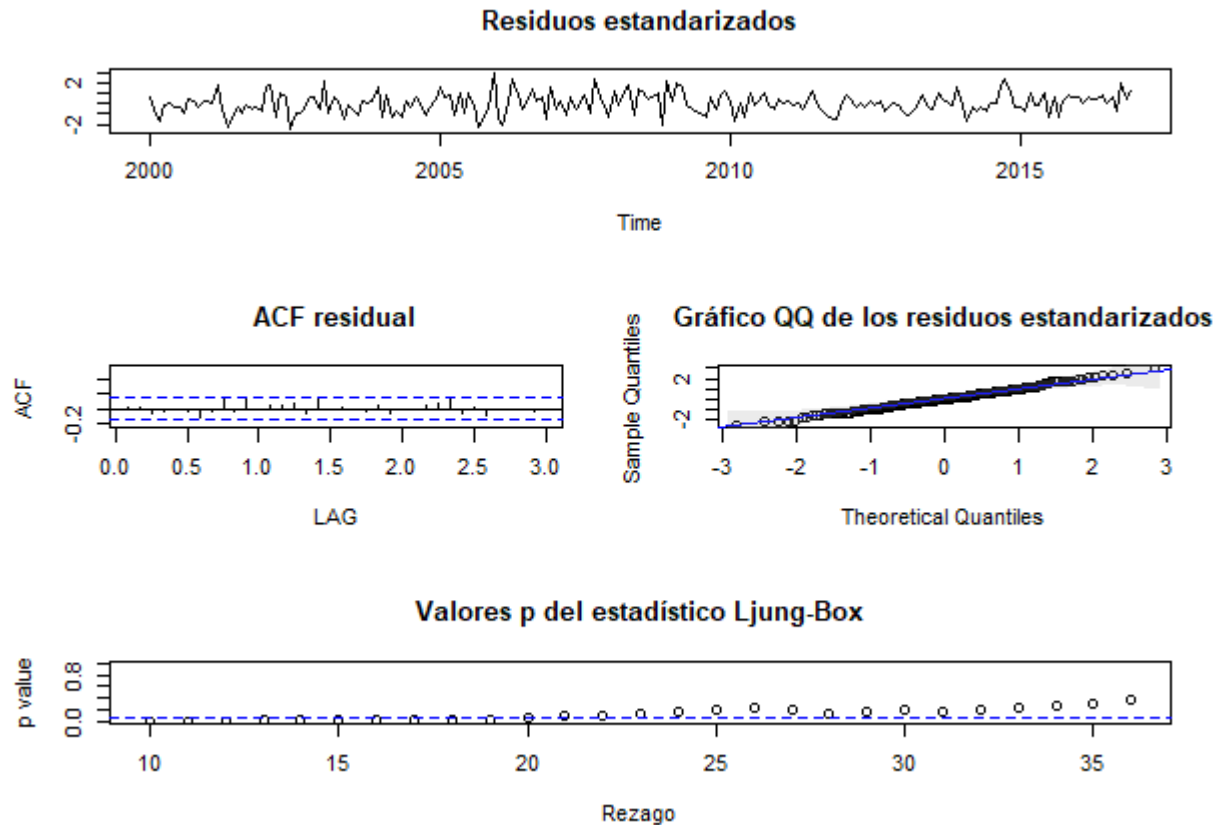


Gráfico 5: Estacionarización de la serie de Muertes por causa externa 2000 – 2017

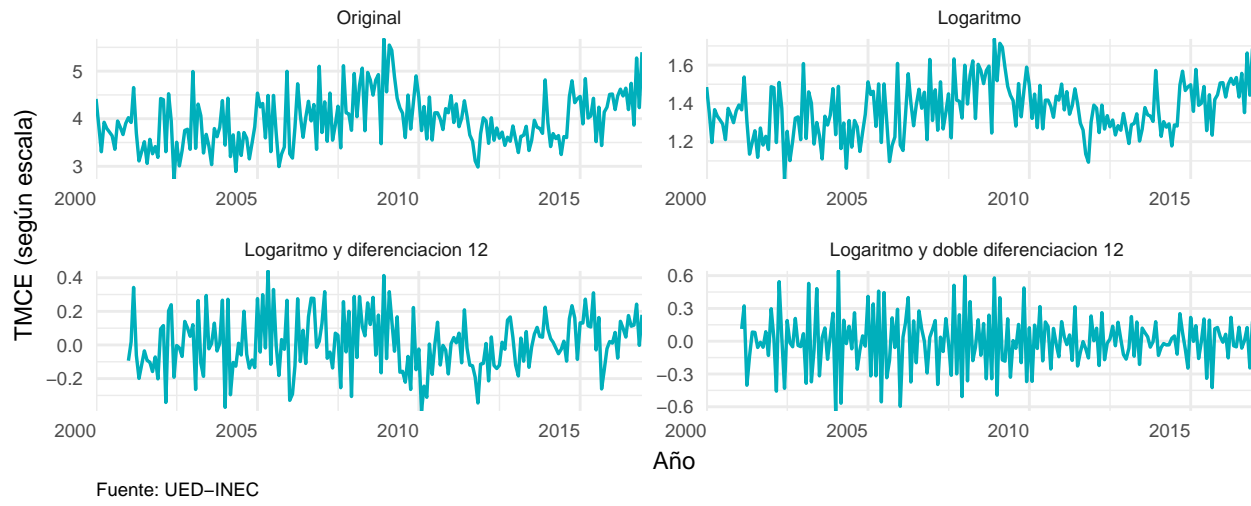
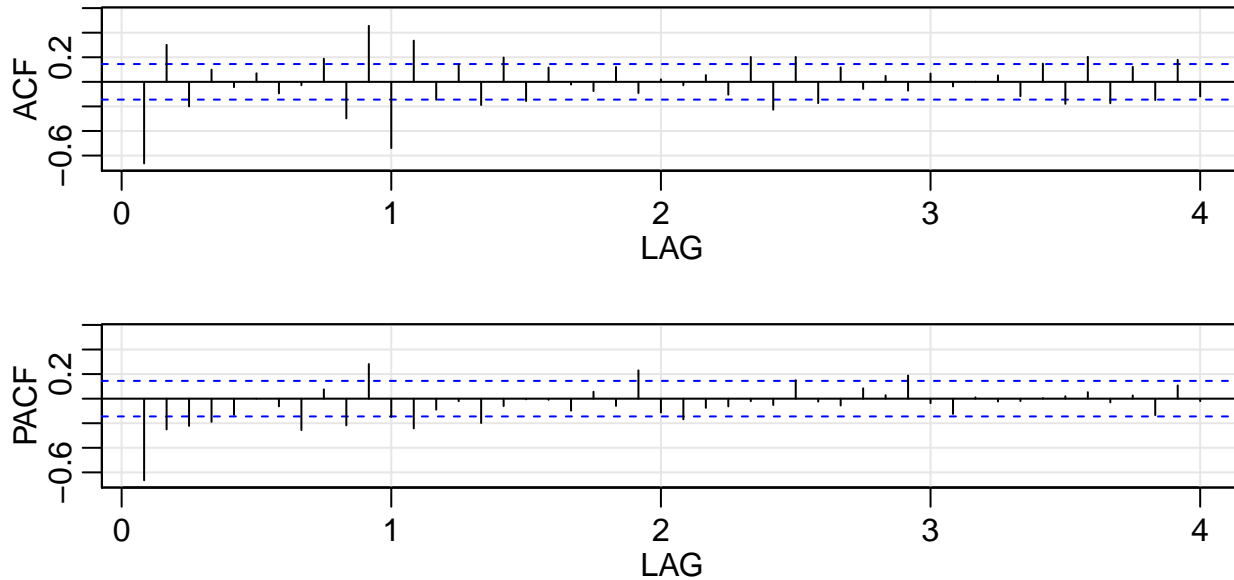


Gráfico 6: Autocorrelación y autocorrelación parcial



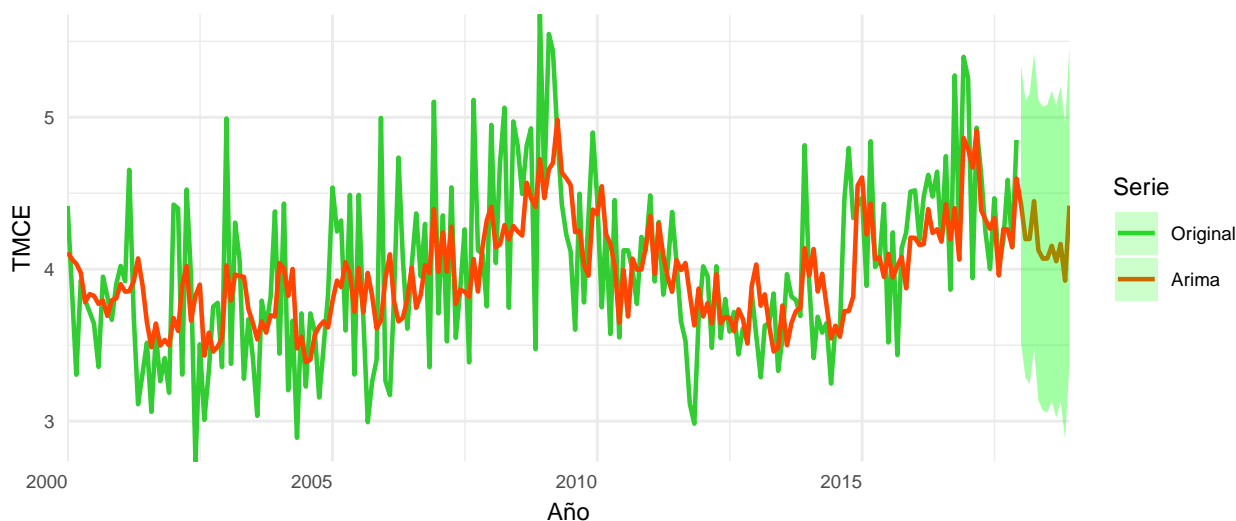
Con los 4 mejores modelos seleccionados en cada categoría, se procede a comparar sus medidas de rendimiento y, posteriormente, el pronóstico para cada uno de ellos y verificar, mediante sus pronósticos, cual de todos es el más adecuado. El cuadro 1 muestra los criterios de información y las medidas de rendimiento para los datos de entrenamiento (rojo) y de validación (azul), siendo el modelo ARIMA estimado el mejor en el conjunto de entrenamiento y el segundo mejor en el de validación, pues queda ligeramente superado por el modelo de suavizamiento exponencial con errores multiplicativos, sin tendencia y con estacionalidad multiplicativa; una comparación gráfica de estos modelos en su pronóstico para el año 2014 puede verse en los gráfico 8-11 en la sección de anexos. Para culminar el análisis, se ofrece el pronóstico con el mejor modelo, que es el $ARIMA(2, 0, 1)(3, 0, 4)_{12}$ para los siguientes 12 periodos correspondientes al año 2018, así como el detalle de dicho pronóstico para cada uno de los meses en el cuadro 1 y el gráfico 7 respectivamente.

Cuadro 1: Medidas de rendimiento y criterios de información para los 4 modelos comparados

Modelo	AIC	AICc	BIC	MAE	MAPE	RMSE	MASE
Reg. tendencia exponencial	-813.33	-811.1	-766.87	0.42	10.6	0.51	0.78
Reg. tendencia exponencial Validacion				0.28	5.94	0.36	0.51
(M,N,A)	796.24	798.79	846.01	0.38	9.54	0.46	0.69
(M,N,A) Validacion				0.27	6.26	0.32	0.49
ARIMA(2,0,1)(3,0,3)[12]	298.32	299.69	334.82	0.35	8.97	0.45	0.65
ARIMA(2,0,1)(3,0,3)[12] Validacion				0.24	5.36	0.32	0.44
ARIMA(3,1,1)(0,0,1)[12]	297.37	297.8	317.25	0.38	9.58	0.49	0.7
ARIMA(3,1,1)(0,0,1)[12] Validacion				0.37	8.59	0.45	0.69

Fuente: UED-INEC

Gráfico 7: Pronóstico de la TMCE para el 2018



Fuente: UED-INEC

Cuadro 2: Pronóstico de la TMCE para 2018 e intervalos de confianza de 95 por ciento

Periodo	Pronóstico	Límite inferior	Límite superior
Ene 2018	4.429742	3.518728	5.340755
Feb 2018	4.198171	3.286861	5.109481
Mar 2018	4.198892	3.244204	5.153581
Abr 2018	4.447787	3.477401	5.418173
May 2018	4.125172	3.137033	5.113311
Jun 2018	4.069298	3.068145	5.070451
Jul 2018	4.070296	3.058014	5.082578
Ago 2018	4.154106	3.132793	5.175419
Sep 2018	4.052150	3.023351	5.080949
Oct 2018	4.166288	3.131317	5.201259
Nov 2018	3.924440	2.884366	4.964514
Dic 2018	4.418334	3.374038	5.462631

Fuente: UED-INEC

4. CONCLUSIONES Y RECOMENDACIONES

Con todo lo descrito en las secciones previas, se han cubierto algunos de los aspectos más fundamentales en un análisis de series temporales, como lo son los análisis descriptivos, la estimación de modelos junto con sus medidas de rendimiento y, lo más importante de todo, el pronóstico. Los resultados obtenidos ofrecen una guía para futuros estudios relacionados con la tasa de mortalidad infantil por causa externa.

La tasa de mortalidad por causas externas es, por su naturaleza, una serie muy volátil, y aunque se busque llegar a los valores más bajos posibles en este indicador, el crecimiento de la población y las condiciones propias de un país en vías de desarrollo hacen que la criminalidad aumente y por ende aumenten también los homicidios. Algo similar ocurre con las muertes accidentales, particularmente en los accidentes de tránsito, pues la alta densidad de vehículos genera mayores dificultades para desplazarse, lo cual a su vez genera irritabilidad en los conductores, facilitando así la aparición de eventos que terminan con el fallecimiento de las personas.

Los análisis descriptivos ofrecen una visualización de rápido crecimiento de las muertes accidentales, lo cual explica en buena medida el comportamiento del total de la TMCE y que además coincide con el aumento de la flotilla de motocicletas. De todos los modelos que se estimaron los mejores son los modelos ARIMA, pues logran capturar de buena manera el comportamiento global de la serie.

El pronóstico realizado ofrece un intervalo de confianza que es congruente con el comportamiento de la serie en general, pues alcanza como valores máximos los que se han venido mostrando desde 2012, mientras que para los valores mínimos es posible que se llegue a niveles similares los presentados desde 2017 debido al aumento de operativos en carretera y el mejoramiento de la infraestructura vial.

A pesar de estos resultados conviene tener presente el peso que tienen las defunciones por causas naturales en la mortalidad general, pues este grupo representa un alto porcentaje. Este artículo plantea un punto de partida para estudiar la mortalidad en forma global para a su vez incluir las técnicas aquí tratadas en las estimaciones y proyecciones de población, pues la mortalidad es un importante componente de estas.

5. ANEXOS

Cuadro 3: Medidas de rendimiento y criterios de informacion para modelos de regresion

Modelo	AIC	AICc	BIC	MAE	MAPE	RMSE	MASE
Regresion T	-242.31	-242.19	-232.36	0.44	11.19	0.54	0.81
Regresion T Validacion				0.35	7.53	0.45	0.65
Regresion T+E	-244.73	-242.51	-198.28	0.42	10.68	0.51	0.78
Regresion T+E Validacion				0.27	5.7	0.35	0.49
Regresión exponencial	-813.33	-811.1	-766.87	0.42	10.6	0.51	0.78
Regresión exponencial Validacion				0.28	5.94	0.36	0.51

Fuente: UED-INEC

Cuadro 4: Medidas de rendimiento y criterios de informacion para modelos de suavizamiento exponencial

Modelo	AIC	AICc	BIC	MAE	MAPE	MASE	RMSE
(A,N,N)	2352.65	2352.77	2362.61	17.15	9.86	0.72	22.03
(A,N,N) Validacion				18.06	8.39	0.76	20.85
(A,A,N)	2354.98	2355.28	2371.57	17.22	10.03	0.72	21.94
(A,A,N) Validacion				19.31	9.06	0.81	22.34
(A,Ad,N)	2357.67	2358.1	2377.58	17.02	9.8	0.71	21.98
(A,Ad,N) Validacion				20.57	9.68	0.86	23.78
(A,N,A)	2343.26	2345.81	2393.03	16.44	9.49	0.69	20.3
(A,N,A) Validacion				13.47	6.39	0.57	16.47
(A,A,A)	2344.67	2347.96	2401.07	16.28	9.5	0.68	20.17
(A,A,A) Validacion				15.61	7.41	0.66	18.37
(A,Ad,A)	2347.18	2350.87	2406.9	16.26	9.4	0.68	20.2
(A,Ad,A) Validacion				17.6	8.33	0.74	19.94
(M,N,N)	2355.14	2355.26	2365.1	17.13	9.83	0.72	22.04
(M,N,N) Validacion				17.86	8.26	0.75	20.53
(M,M,N)	2357	2357.3	2373.59	17.24	10.04	0.72	21.95
(M,M,N) Validacion				19.67	9.24	0.83	22.74
(M,Md,N)	2358.42	2358.85	2378.33	17.13	9.86	0.72	22.03
(M,Md,N) Validacion				21.36	10.06	0.9	24.74
(M,N,M)	2343.37	2345.92	2393.14	16.3	9.39	0.68	20.43
(M,N,M) Validacion				11.45	5.45	0.48	14.22
(M,Md,M)	2345.69	2349.39	2405.41	16.36	9.48	0.69	20.24
(M,Md,M) Validacion				17.41	8.31	0.73	20.51
(M,N,A)	2339.09	2341.64	2388.86	16.42	9.45	0.69	20.3
(M,N,A) Validacion				13.34	6.3	0.56	16.28

Fuente: UED-INEC

Cuadro 5: Medidas de rendimiento y criterios de informacion para modelos ARIMA

Modelo	AIC	AICc	BIC	MAE	MAPE	RMSE	MASE
ARIMA(2,0,1)(3,0,3)[12]	298.32	299.69	334.82	0.35	8.97	0.45	0.65
ARIMA(2,0,1)(3,0,3)[12] Validacion				0.24	5.36	0.32	0.44
ARIMA(2,0,4)(3,0,3)[12]	300.86	303.08	347.31	0.34	8.75	0.44	0.63
ARIMA(2,0,4)(3,0,3)[12] Validacion				0.25	5.67	0.31	0.46
ARIMA(3,0,1)(3,0,3)[12]	296.93	298.56	336.74	0.34	8.76	0.44	0.63
ARIMA(3,0,1)(3,0,3)[12] Validacion				0.25	5.69	0.32	0.46
ARIMA(1,0,1)(1,0,1)[12]	299.56	299.99	319.47	0.39	9.88	0.48	0.72
ARIMA(1,0,1)(1,0,1)[12] Validacion				0.27	6.13	0.32	0.5
ARIMA(4,0,1)(3,0,3)[12]	302.82	304.74	345.96	0.37	9.29	0.47	0.68
ARIMA(4,0,1)(3,0,3)[12] Validacion				0.27	6.09	0.36	0.51

Fuente: UED-INEC

Gráfico 8: Serie original y pronóstico para el 2017

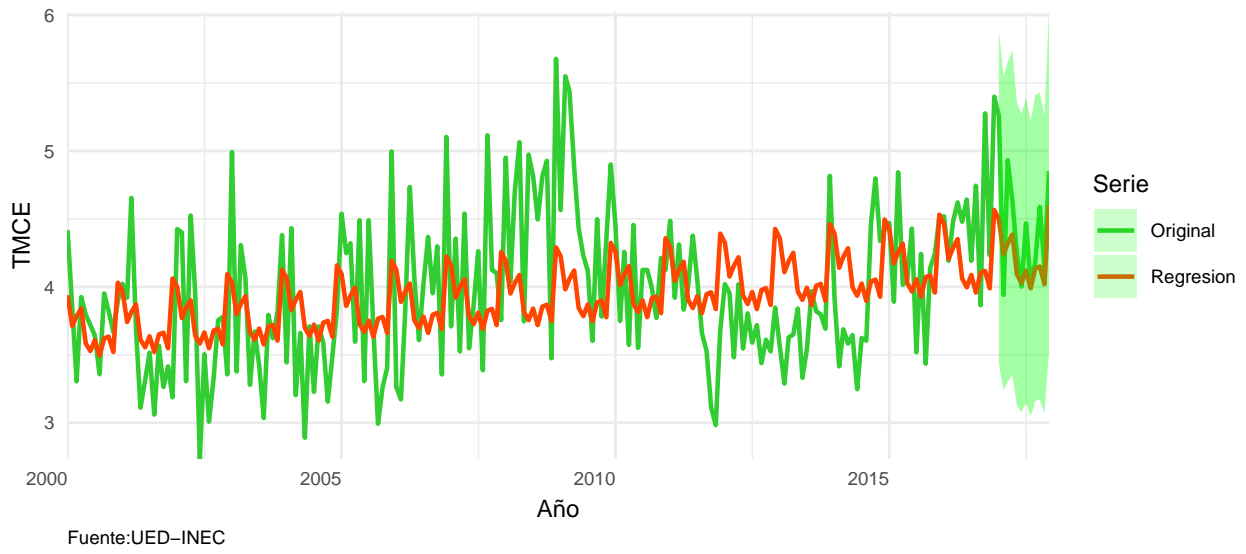


Gráfico 9: Serie original y pronóstico para el 2017

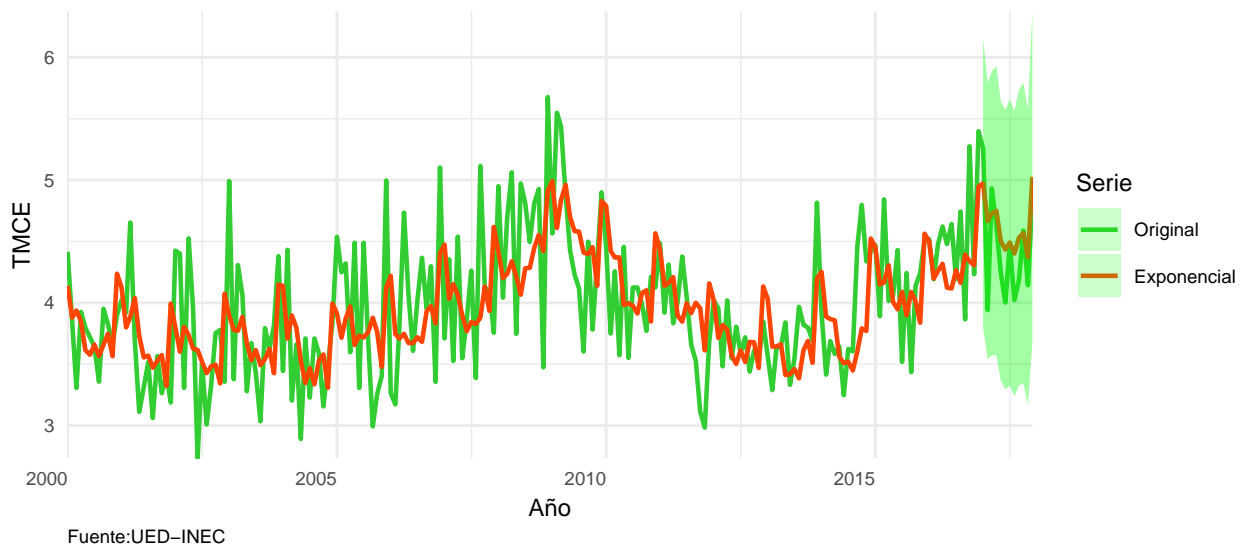


Gráfico 10: Serie original y pronóstico para el 2017

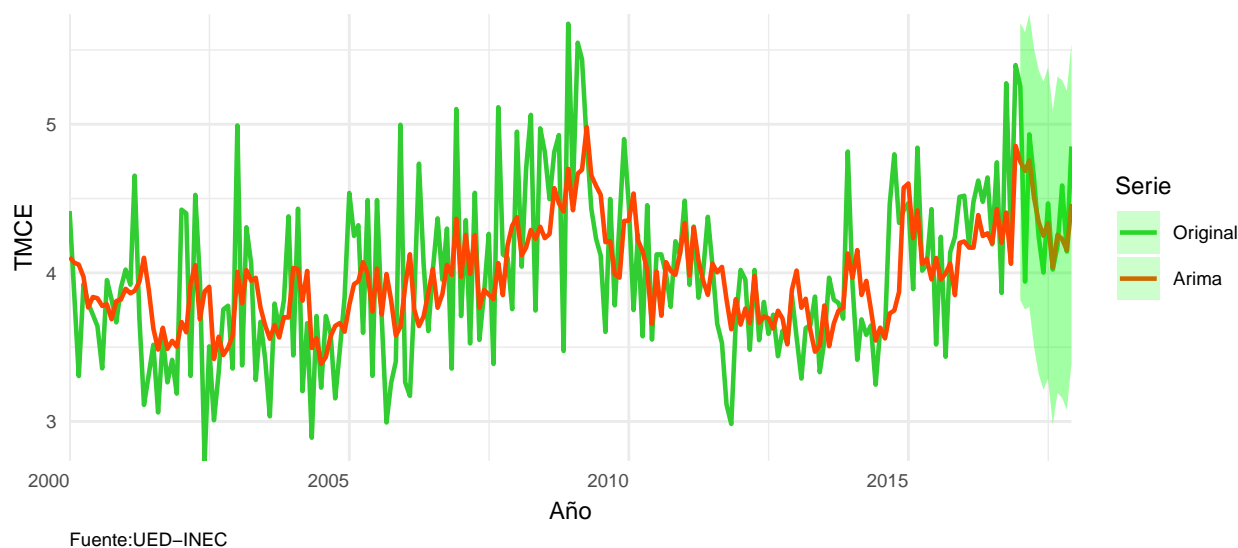
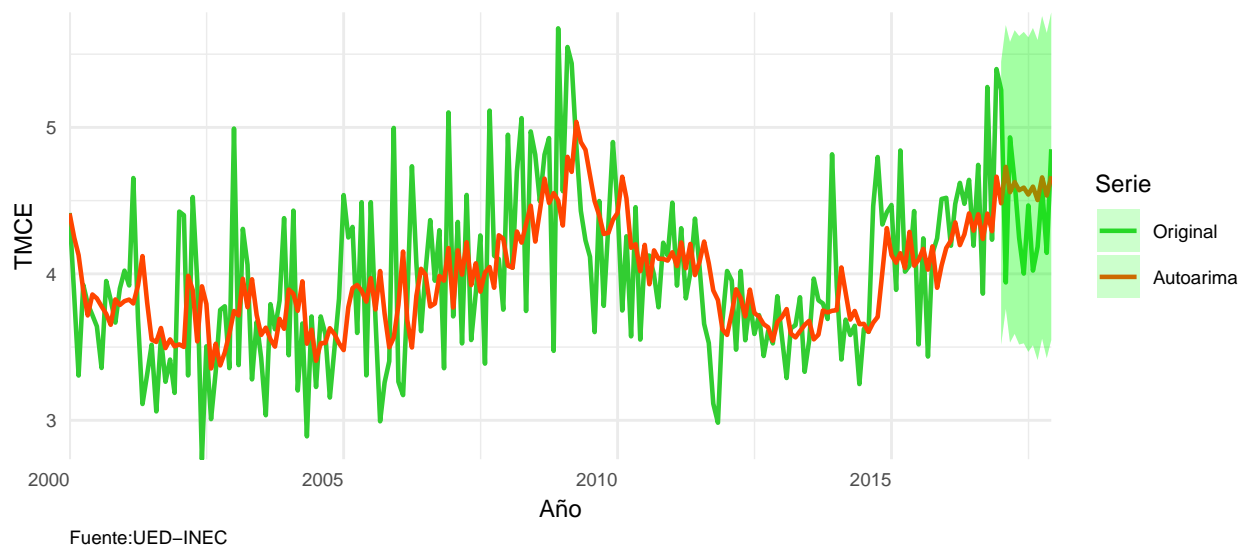


Gráfico 11: Serie original y pronóstico para el 2017



6. REFERENCIAS

- Aphalo, Pedro J. 2016. *Learn R ...as You Learnt Your Mother Tongue*. Leanpub. <https://leanpub.com/learnr>.
- Cardona, G.; Fantín, D.; Escané. 2013. "Mortalidad Por Causas Externas: Un Problema de Salud Pública. Argentina, Chile Y Colombia. 2000-2008." *Revista Electrónica Semestral* 10 (2). https://www.researchgate.net/publication/274885475_Mortalidad_por_causa_externas_un_problema_de_salud_publica_Argentina_Chile_y_Colombia_2000-2008.
- Ellis, Peter. 2018. *Ggseas: 'Stats' for Seasonal Adjustment on the Fly with 'Ggplot2'*. <https://CRAN.R-project.org/package=ggseas>.
- Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of*

Statistical Software 40 (3): 1–25. <http://www.jstatsoft.org/v40/i03/>.

Hyndman R, Bergmeir C, Athanasopoulos G. 2018. *Forecast: Forecasting Functions for Time Series and Linear Models*. <http://www.jstatsoft.org/article/view/v027i03>.

Mata, L. 1995. “Homicidios Y Suicidios En Costa Rica.” *La Nacion*, December. <https://bit.ly/2P579gR>.

Nación. 2013. “Morbilidad Y Mortalidad En Costa Rica.” *La Nacion*, January. <https://bit.ly/2xWUeXU>.

OPS. 2016. “Clasificación Estadística Internacional de Enfermedades Y Problemas Relacionados Con La Salud.” In, 2015th ed. Vol. 2. OMS.

Vázquez, J. 2017. “En 5 Años Flotilla de Motos Se Disparó En Un 189 Por Ciento.” *CR Hoy*, June. <https://bit.ly/2QmQQfE>.

Wickham, Hadley. 2016. *Ggplot2: Elegant Graphics for Data Analysis*. Springer-Verlag New York. <http://ggplot2.org>.

Wickham, Hadley, and Jennifer Bryan. 2018. *Readxl: Read Excel Files*. <https://CRAN.R-project.org/package=readxl>.

Wickham, Hadley, and Lionel Henry. 2018. *Tidyr: Easily Tidy Data with 'Spread()' and 'Gather()' Functions*. <https://CRAN.R-project.org/package=tidyr>.

Wickham, Hadley, Romain François, Lionel Henry, and Kirill Müller. 2018. *Dplyr: A Grammar of Data Manipulation*. <https://CRAN.R-project.org/package=dplyr>.