



Ciencia Latina
Internacional

Ciencia Latina Revista Científica Multidisciplinar, Ciudad de México, México.
ISSN 2707-2207 / ISSN 2707-2215 (en línea), mayo-junio 2024,
Volumen 8, Número 3.

https://doi.org/10.37811/cl_rcm.v8i3

**PRONÓSTICO DEL PRECIO DE BARRIL DE
PETRÓLEO EN MÉXICO PARA EL PERIODO
FEBRERO DE 2023 A ENERO DE 2024 MEDIANTE
UN MODELO ARIMA EN RSTUDIO**

**FORECAST OF THE PRICE OF A BARREL OF OIL IN MEXICO
FOR THE PERIOD FEBRUARY 2023 TO JANUARY 2024 USING
AN ARIMA MODEL IN RSTUDIO**

César Ángel Fierro Torres

Tecnológico Nacional de México, México

Arturo Woocay Prieto

Tecnológico Nacional de México, México

Claudia Irene Torres Saucedo

Centro de Bachillerato Tecnológico Industrial y de Servicios 128, México

Laura Isela Gómez Palma

Centro de Bachillerato Tecnológico Industrial y de Servicios 128, México

Natalia Idaly Barraza Ramírez

Centro de Bachillerato Tecnológico Industrial y de Servicios 128, México

DOI: https://doi.org/10.37811/cl_rcm.v8i3.11326

Pronóstico del Precio de Barril de Petróleo en México para el Periodo Febrero de 2023 a Enero de 2024 Mediante un Modelo ARIMA en RStudio

César Ángel Fierro Torres¹

M21110218@cdjuarez.tecnm.mx
<https://orcid.org/0000-0003-4962-4618>
Tecnológico Nacional de México
Instituto Tecnológico de Ciudad Juárez
México

Arturo Woocay Prieto

arturo.wp@cdjuarez.tecnm.mx
<https://orcid.org/0000-0001-9235-0494>
Tecnológico Nacional de México
Instituto Tecnológico de Ciudad Juárez
México

Claudia Irene Torres Saucedo

claudiarene.torres.cb128@dgeti.sems.gob.mx
<https://orcid.org/0000-0002-9375-5976>
Centro de Bachillerato Tecnológico Industrial
y de Servicios 128
Chihuahua, México

Laura Isela Gómez Palma

laura.gomez@cbtis128.edu.mx
<https://orcid.org/0009-0008-0653-9513>
Centro de Bachillerato Tecnológico Industrial
y de Servicios 128
Chihuahua, México

Natalia Idaly Barraza Ramírez

natalia.barraza@cbtis128.edu.mx
<https://orcid.org/0009-0005-6174-9861>
Centro de Bachillerato Tecnológico Industrial
y de Servicios 128
Chihuahua, México

RESUMEN

El objetivo del presente trabajo de investigación fue desarrollar un caso práctico de pronóstico de series temporales para el precio del barril de petróleo en México, para el periodo febrero de 2023 a enero de 2024. El precio del barril de petróleo se modeló como una serie de tiempo y se pronosticó utilizando un modelo de promedios móviles autorregresivos integrados (ARIMA). Bajo un enfoque estocástico, se define que una serie temporal es la secuencia de valores de una variable a lo largo del tiempo, cuyo comportamiento está dado por la aleatoriedad a partir de un proceso que es totalmente desconocido o difícil de modelar matemáticamente. En el caso del precio del petróleo, este se ve afectado por una gran variedad de variables, factores y sucesos geopolíticos, como es el caso de la guerra en Ucrania en la actualidad, por lo que el enfoque estocástico mediante un modelo ARIMA resulta una opción pertinente. Los resultados comparados con los datos de prueba muestran que el modelo tuvo un error MAPE de 6.33% y un RMSE de 6.106, por lo que se considera que el modelo generó pronósticos precisos, a pesar de que se consideró un horizonte a largo plazo de 12 meses para su evaluación.

Palabras clave: ARIMA, pronósticos, series temporales, proceso estocástico

¹ Autor principal

Correspondencia: M21110218@cdjuarez.tecnm.mx

Forecast of the Price of a Barrel of Oil in Mexico for the Period February 2023 to January 2024 Using an ARIMA Model in RStudio

ABSTRACT

The objective of this research was to develop a case study of forecasting time series for the price of a barrel of oil in Mexico, for the period February 2023 to January 2024. The price of a barrel of oil was modeled as a time series and forecast using a model of self-regressive integrated moving averages (ARIMA). Under a stochastic approach, a time series is defined as the sequence of values of a variable over time, whose behavior is given by randomness from a process that is totally unknown or difficult to model mathematically. In the case of oil prices, oil prices are affected by a wide variety of geopolitical variables, factors and events, such as the war in Ukraine at present, making the stochastic approach using an ARIMA model a relevant option. The results compared with the test data show that the model had a MAPE error of 6.33% and an RMSE of 6.106, so it is considered that the model generated accurate forecasts, although it was considered a long-term horizon of 12 months for its evaluation.

Keywords: ARIMA, forecasts, time series, stochastic process

Artículo recibido 10 abril 2024

Aceptado para publicación: 20 mayo 2024



INTRODUCCIÓN

Una serie de tiempo es un conjunto de datos organizados de forma secuencial en el tiempo en intervalos regulares y bajo una misma escala (días, semanas, meses, etc.) y que sirven para representar cualquier tipo de variable que se busque medir a lo largo del tiempo.

De acuerdo con Render y Heizer (2014), una serie temporal puede descomponerse en sus componentes estacional, tendencia, ciclo y aleatoria, cuyas definiciones son las siguientes:

Estacionalidad

Patrón que siguen los datos que se repite en periodos de tiempo de corto plazo: días, semanas, meses, trimestres, semestres, etc. Por ejemplo, los negocios de entretenimiento presentan estaciones semanales donde el fin de semana es su periodo de mayor venta.

Tendencia

Es el movimiento que sigue la serie de tiempo de forma sucesiva y continua hacia arriba o hacia abajo y que suele representarse como una recta (es la componente de los modelos de regresión lineal).

Ciclos

Son patrones en los datos, mayores de un año, debidos a factores distintos a los generados por la estacionalidad. Los intervalos de tiempo de las variaciones cíclicas no son fijos ya que no siguen ninguna estacionalidad. Los ciclos suelen mostrar periodos de auge, desaceleración o declive de las economías y suelen ser difíciles de predecir, dadas sus múltiples causas que corresponden a cuestiones geopolíticas.

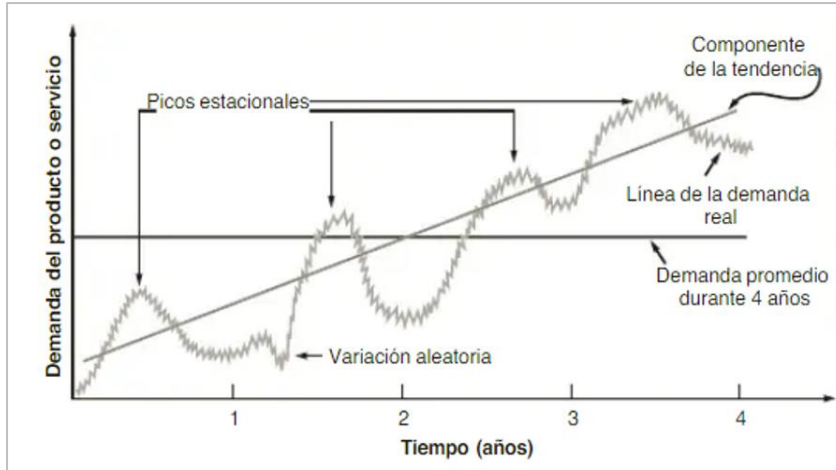
Variaciones aleatorias

Son valores obtenidos por casualidad o situaciones fuera de lo común y que no pueden ser predecibles ya que no siguen ningún patrón de comportamiento.

En la figura 1 se muestra un ejemplo de serie de tiempo de la demanda de un producto. Incluye las componentes de la serie que son: la tendencia, la línea de la demanda, la demanda promedio, los picos estacionales y la variación aleatoria.



Figura 1. Componentes de una serie de tiempo



Fuente: (Render y Heizer, 2014, p. 109)

De acuerdo con García Díaz (2016), una serie temporal puede descomponerse en sus componentes: estacional, tendencia, ciclo y aleatoria, en tres esquemas diferentes: aditivo, multiplicativo y mixto. En el esquema aditivo, la serie de tiempo es el resultado de la suma de sus cuatro componentes, lo cual implica que son independientes entre sí. La serie de tiempo está dada por la ecuación:

$$Z_t = T_t + C_t + S_t + I_t \quad (1)$$

Donde Z_t corresponde a la serie temporal, T_t a la componente de la tendencia (trend), C_t a la componente cíclica (Cycle), S_t a la componente estacional (Seasonal) e I_t a la componente de aleatoriedad (Irregular). Cuando algunas de las componentes no sean independientes y tengan alguna relación, se trata de un esquema multiplicativo de series de tiempo. Cuando todas las componentes son dependientes, la serie de tiempo está dada por la ecuación:

$$Z_t = T_t * C_t * S_t * I_t \quad (2)$$

Para identificar visualmente la interacción aditiva o multiplicativa entre las componentes de tendencia y estacionalidad en el patrón de datos, es de ayuda la clasificación realizada por Carl Pegels en 1969 (Fullana Fuster, 2020). En la figura 2 se presentan los nueve posibles casos del comportamiento de los datos de una serie de tiempo.

Figura 2. Conducta de series de tiempo según la clasificación de Pegels.

	1 No efecto estacional	2 Estacionalidad Aditiva	3 Estacionalidad multiplicativa
A No efecto tendencia			
B Tendencia aditiva			
C Tendencia multiplicativa			

Fuente: (Fullana Fuster, 2020)

De acuerdo con la clasificación de Pegels, se puede observar que la tendencia no afecta la serie de tiempo cuando esta no tiene ninguna inclinación, es decir, tiene un comportamiento horizontal. Cuando se aprecia un comportamiento lineal, se trata de una tendencia aditiva; mientras que cuando es multiplicativa tiene un comportamiento exponencial.

Cuando no hay un patrón en el comportamiento de los datos de la serie de tiempo, se puede inferir que no hay una afectación por la estacionalidad. Cuando hay un patrón en el comportamiento de los datos que se mantiene constante en el tiempo, se trata de una estacionalidad aditiva. Por último, cuando el patrón se expande en el tiempo, se puede considerar como una estacionalidad multiplicativa.

Sin embargo, en series de tiempo más complejas puede darse el caso de que no es posible detectar a simple vista la naturaleza de las componentes de tendencia y estacionalidad, por lo que se recomienda probar modelos aditivos y multiplicativos por separado y elegir el que tenga menor error de pronóstico, también llamado bondad de ajuste.

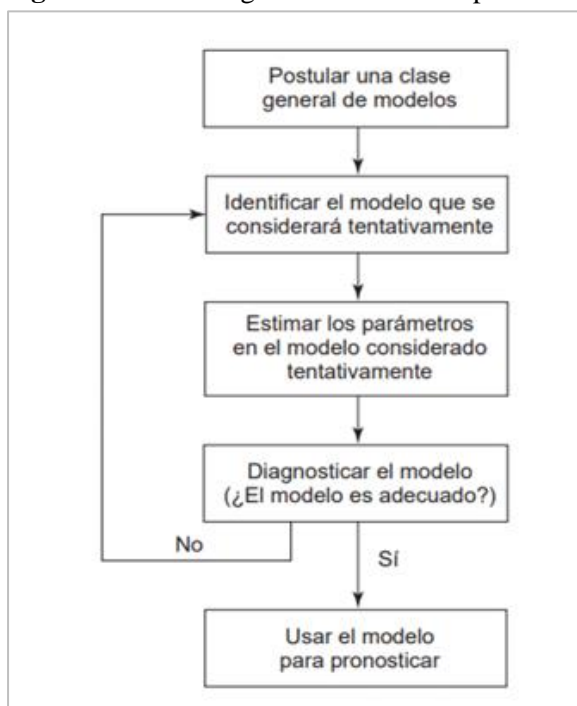
En las últimas décadas del siglo XX, hubo un cambio de paradigma en la modelación de las series de tiempo. El enfoque clásico determinista explica que el valor de la variable dependiente (el pronóstico) Z_t , está dado por las componentes (tendencia, ciclo, estacionalidad y aleatoriedad) que actúan en el tiempo, estas son las variables independientes. Ejemplos de dicha modelación clásica se encuentran en las ecuaciones 1 y 2 explicadas anteriormente. La desventaja del enfoque determinista es que la aleatoriedad o también llamado ruido o residuo es una componente no modelable y por lo tanto se considera como error de pronóstico. Entre más aleatoria sea una serie de tiempo, más difícil es para un

modelo determinista encontrar una función matemática que realice la predicción (Aguirre, 1994). En el caso de estudios de fenómenos sociales, económicos o en el área de la salud, donde la aleatoriedad es una constante que no se puede ajustar como en las ciencias experimentales o la ingeniería, la modelación clásica se vio limitada en sus alcances.

Ante esta problemática, se empezó a proponer un tratamiento estocástico de las series de tiempo, es decir, que la variable de estudio es resultado de la aleatoriedad a partir de un proceso que es totalmente desconocido (o difícil de modelar matemáticamente). Por lo tanto, la modelación estocástica de series de tiempo busca identificar el modelo probabilístico que interprete de mejor manera el comportamiento de la variable a lo largo del tiempo. Una de las ventajas del enfoque estocástico es la flexibilidad para modelar distintos fenómenos bajo una clase general de modelos (Cortés Patiño, 2011).

En la década de los años 1970, hubo un avance notable en la modelación estocástica de series de tiempo, gracias al trabajo realizado por los estadísticos norteamericanos George Edward Pelham Box (1919-2013) y Gwilym Jenkins (1932-1982). Ambos crearon una metodología a partir del modelo de promedio móvil autorregresivo integrado, ARIMA por sus siglas en inglés. También se le conoce como metodología de Box-Jenkins en honor a sus autores (Makridakis, Wheelwright y Hyndman, 2005). metodología de Box Jenkins sigue el procedimiento que se muestra en el diagrama de flujo de la figura 3.

Figura 3. Metodología de Box-Jenkins para modelos ARIMA.



Fuente: (Hanke y Wichern, 2010, p. 400)

1. Selección de un modelo inicial

En la identificación del modelo es necesario verificar la autocorrelación y la autocorrelación parcial de los valores en la serie de tiempo para identificar si se tratan de modelos autorregresivos (AR), modelos de promedio móvil (MA) o modelos de promedio móvil autorregresivos (ARMA). Existen dos tipos de procesos estocásticos: estacionarios, que al tener una distribución de probabilidad constante el tiempo permite realizar pronósticos, y los no estacionarios cuya función de probabilidad no es constante y por lo tanto es impredecible. Cuando se trabaja con series de tiempo no estacionarias es necesario transformarlas, generalmente por diferenciación, generando un modelo ARIMA (p,d,q), siendo p el número de autorregresivos, d el número de diferenciaciones y q el número de medias móviles.

2. Estimación de los coeficientes del modelo

Son los valores de los coeficientes de las ecuaciones del modelo y que pueden ser estimadas empíricamente mediante la observación del ajuste de la serie del tiempo o bien utilizando métodos numéricos para asegurar el mejor ajuste.

3. Verificación del modelo

Se utiliza el modelo planteado con base en un conjunto de datos de entrenamiento o ajuste para realizar pronósticos a un periodo de prueba que pueda ser comparado con el valor real.

4. Diagnóstico del modelo

Se utilizan distintas métricas para evaluar los residuos en los datos de prueba (la diferencia entre el valor histórico y el valor del modelo) y si este es aceptable, se procede a realizar los pronósticos para n periodos de tiempo deseados (Olvera Vázquez, 2020).

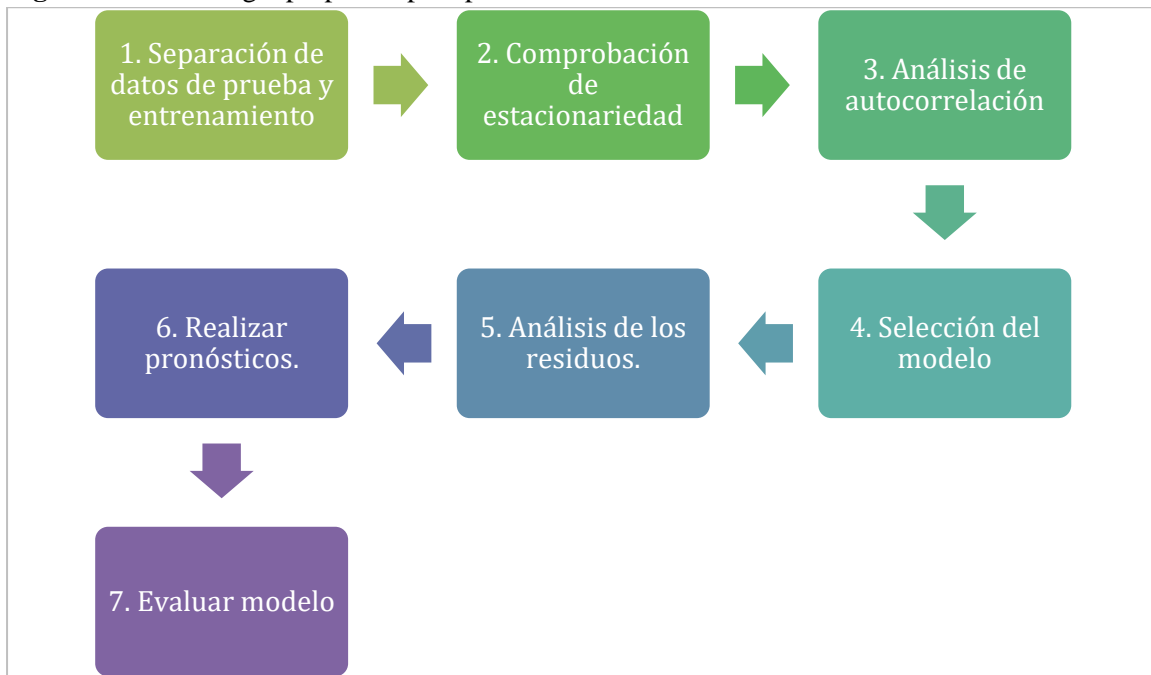
MÉTODO

Hoy en día, con la ayuda de herramientas estadísticas potentes como R, es posible seguir un método más riguroso para la validación y realización de pronósticos de series temporales utilizando los modelos ARIMA con ayuda de pruebas de hipótesis estadísticas más especializadas, en comparación con las tradicionales hojas de cálculo u otros softwares comerciales.

R es un lenguaje de programación creado en 1993 por los profesores Robert Gentleman y Ross Ihaka para enseñar estadística en la Universidad de Auckland, tomando el nombre de las iniciales de sus creadores. Es utilizado para el análisis, procesamiento y visualización de datos. RStudio es un entorno

de desarrollo integrado (IDE), que consiste en una interfaz gráfica para utilizar de forma más amigable y accesible el lenguaje R, que facilita la escritura, depuración y ejecución de código (Vargas y Mesa-Fúquen, 2021). R funciona a partir de librerías con distintas funciones para un propósito en específico. Con base en la literatura consultada se expone el método utilizado para el pronóstico del precio del barril del petróleo en México mediante RStudio en la figura 4.

Figura 4. Metodología propuesta para pronósticos ARIMA.

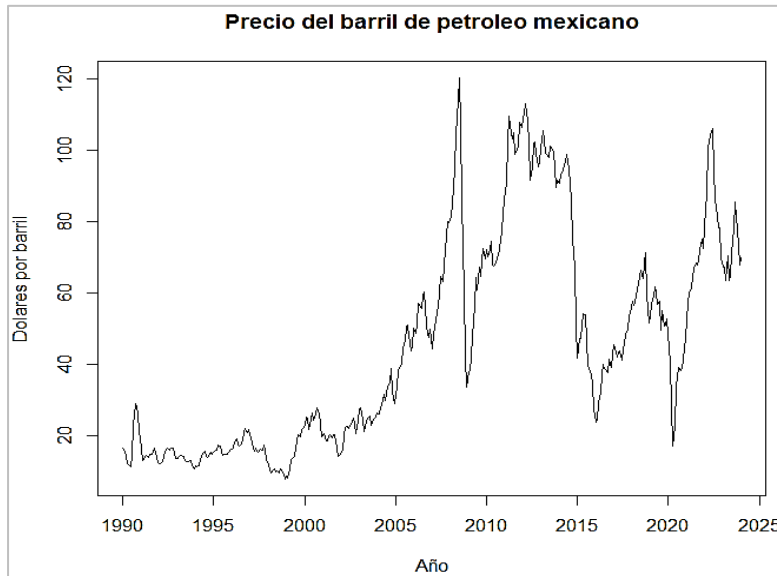


Fuente: Elaboración propia.

1. Separación de datos de prueba y entrenamiento

Para poder realizar y posteriormente validar un modelo de pronósticos es necesario separar los datos en grupos de entrenamiento, que son aquellos con los que se ajusta el modelo, y los de prueba, con los que se realiza la comparación de las estimaciones. Los datos obtenidos a través del Sistema de Información Energética y el Banco de México, permiten elaborar una serie de tiempo mensual desde enero de 1990 a enero de 2024 del precio del barril del petróleo mexicano en dólares, misma que se representa gráficamente en la figura 5.

Figura 5. Precio del barril de petróleo mexicano de enero de 1990 a enero de 2024



Fuente: Elaboración propia con datos del Sistema de Información Energética y el Banco de México.

De esta serie de tiempo se tomará como el grupo de entrenamiento los datos de los meses de enero de 1990 a enero de 2023, mientras que los datos de prueba serán desde febrero de 2023 a enero de 2024, es decir, 12 meses.

2. Comprobación de la estacionariedad

El primer paso para plantear un modelo ARIMA, es determinar si la serie es estacionaria o no, definiendo que el concepto de estacionariedad implica que la serie tiene un comportamiento estable a través del tiempo, lo cual se ve reflejado gráficamente con una forma de “ruido blanco”. Si bien la estacionariedad puede comprobarse de forma visual, detectando si la serie de tiempo tiene tendencia o no, una manera formal de identificarla es mediante la prueba de hipótesis de Dickey-Fuller con la distribución t (Roza, Violita y Aktivani, 2022).

La hipótesis nula es que el modelo de predicción tiene raíz unitaria (tendencia impredecible en un proceso estocástico), la hipótesis alternativa es que la serie de tiempo no tiene raíz unitaria. A la raíz unitaria también se le conoce como la caminata aleatoria, ya que implica que los valores de la variable dependiente no están dados solamente por las mediciones en el tiempo, sino por otras variables aleatorias independientes y con media igual a cero. Por lo tanto, se busca rechazar la hipótesis nula de que la serie de tiempo tiene raíz unitaria, lo que implica que es estacionaria. Los resultados realizados en R, mostrados en la figura 6 muestran un valor P de 0.367, por lo tanto, no se puede rechazar H_0 , lo

que significa que la varianza y la media no son constantes a lo largo del tiempo, sino que son cambiantes y probablemente hay una tendencia.

Figura 6. Test de Dickey-Fuller para la estacionariedad

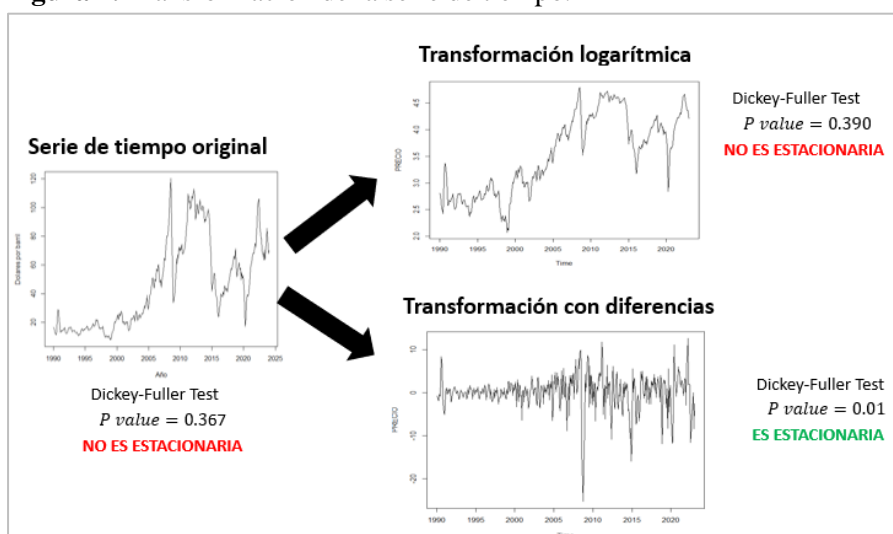
```
Augmented Dickey-Fuller Test
data: ARIMATS
Dickey-Fuller = -2.4979, Lag order = 7, p-value = 0.367
alternative hypothesis: stationary
```

Fuente: Elaboración propia.

Como es de suponerse, es muy común que, dada la naturaleza de ciertas series de tiempo, el requisito de la estacionariedad se incumpla, por lo que para poder trabajar con los modelos ARIMA es necesario transformar la serie de tiempo para volverla estacionaria. Los tres métodos más populares para transformar la serie de tiempo son los logaritmos naturales, la raíz cuadrada y la diferenciación. El primer método consiste en sacar el logaritmo natural de los valores de la serie; en el segundo método se saca la raíz cuadrada de los valores, mientras que en el método de diferenciación se trabaja con las diferencias secuenciadas de los valores.

Es necesario que una vez realizada la transformación se vuelva a realizar la prueba de hipótesis de Dickey-Fuller para confirmar la estacionariedad, si en la primera transformación no se ha podido rechazar H_0 , es necesario repetir el procedimiento hasta lograr la estacionariedad. En la figura 7 se representa el proceso de transformación de la serie de tiempo del precio de petróleo.

Figura 7. Transformación de la serie de tiempo.



Fuente: Elaboración propia.

Se comparan dos métodos, el de la transformación logarítmica y el de transformación con diferencias. Al realizar la primera transformación logarítmica, la raíz unitaria, lo que es igual a la tendencia, se ve incrementada, tal como indica el aumento del valor p; por lo que este método logra lo opuesto a la estacionariedad. Con el método de diferencias, al realizar la primera transformación, visualmente se observa un cambio de la serie temporal hacia la estacionariedad y el valor p de la prueba de hipótesis (0.01), menor a 0.05, por lo que se rechaza la hipótesis nula y se acepta la hipótesis alternativa de que la serie de tiempo es estacionaria. Por lo tanto, para este ejemplo, se trata de un modelo ARIMA con una diferencia (d).

3. Análisis de la autocorrelación

Es común que cuando se mide una variable a lo largo del tiempo, los valores estén correlacionados en diferentes periodos. Esta correlación se mide a través del coeficiente de autocorrelación. “La autocorrelación es la correlación que existe entre una variable retrasada uno o más periodos consigo misma” (Hanke y Wichern, 2010).

Los retardos son el número de veces que han sido desfasados los datos en la escala de tiempo. En la tabla 1 se muestra el ejemplo de un conjunto de datos con retardos $k= 0,1,2,3,4$ del precio del petróleo de enero a octubre de 1990. En el primer retardo se busca encontrar una correlación del dato t (siendo t el periodo de tiempo) con respecto al dato t-1; en el retardo dos se busca encontrar la correlación del dato t con respecto al dato t-2, y así sucesivamente. Por ejemplo, para el dato del precio del mes de junio con retardo igual a 4, significa que se buscara la autocorrelación el precio del mes de febrero (4 meses atrás).

Tabla 1. Retardos a cuatro meses de la serie de tiempo (enero a octubre de 1990)

Mes	Precio	Retardo=0 Yt	Retardo=1 Yt-1	Retardo=2 Yt-2	Retardo=3 Yt-3	Retardo=4 Yt-4
ene-90	16.6	16.6				
feb-90	15.8	15.8	16.6			
mar-90	14.2	14.2	15.8	16.6		
abr-90	12.4	12.4	14.2	15.8	16.6	
may-90	12.1	12.1	12.4	14.2	15.8	16.6
jun-90	11.3	11.3	12.1	12.4	14.2	15.8
jul-90	14.5	14.5	11.3	12.1	12.4	14.2
ago-90	22.8	22.8	14.5	11.3	12.1	12.4
sep-90	29.1	29.1	22.8	14.5	11.3	12.1
oct-90	28.4	28.4	29.1	22.8	14.5	11.3

Fuente: Elaboración propia.

A partir de los k retardos de los datos se pueden calcular el coeficiente de autocorrelación de la variable (Montgomery, Jennings y Kulahci, 2015).

$$\rho_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2} \quad k = 0,1,2 \dots \quad (3)$$

Donde:

r_k = Coeficiente de autocorrelación para retardo en k periodos.

\bar{Y} = Media de los valores de la serie

Y_t = Observación en el periodo t .

Y_{t-k} = Observación k periodos anteriores

Una de las características de la función de autocorrelación es que el coeficiente calculado para un retardo k incluye el efecto de autocorrelación acumulado de los retardos anteriores. Para saber el efecto puro de autocorrelación del retardo k es necesario calcular la función de autocorrelación parcial. Por ejemplo, el coeficiente de correlación para un desfase de $k=6$, es únicamente la correlación que no explican los desfases del 1 al 5. Para poder obtener los coeficientes de autocorrelación parcial es necesario realizar operaciones matriciales dadas por la ecuación 4, que forma parte del conjunto de ecuaciones de Yule-Walker (Montgomery, Jennings y Kulahci, 2015).

$$\phi_k = P_k^{-1} \rho_k \quad k = 0,1,2 \dots \quad (4)$$

Donde:

$$P_k = \begin{bmatrix} 1 & \rho(1) & \rho(2) & \dots & \rho(k-1) \\ \rho(1) & 1 & \rho(3) & \dots & \rho(k-2) \\ \rho(2) & \rho(1) & 1 & \dots & \rho(k-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho(k-1) & \rho(k-2) & \rho(k-3) & \dots & 1 \end{bmatrix}$$

$$\phi_k = \begin{bmatrix} \phi_{1k} \\ \phi_{2k} \\ \phi_{3k} \\ \vdots \\ \phi_{kk} \end{bmatrix} \quad y \quad \rho_k = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \rho(3) \\ \vdots \\ \rho(k) \end{bmatrix}$$

ϕ_k = Vector de autocorrelaciones parciales para k retardos.

P_k = Matriz de autocorrelaciones para k retardos

ρ_k = Vector de autocorrelaciones para k retardos

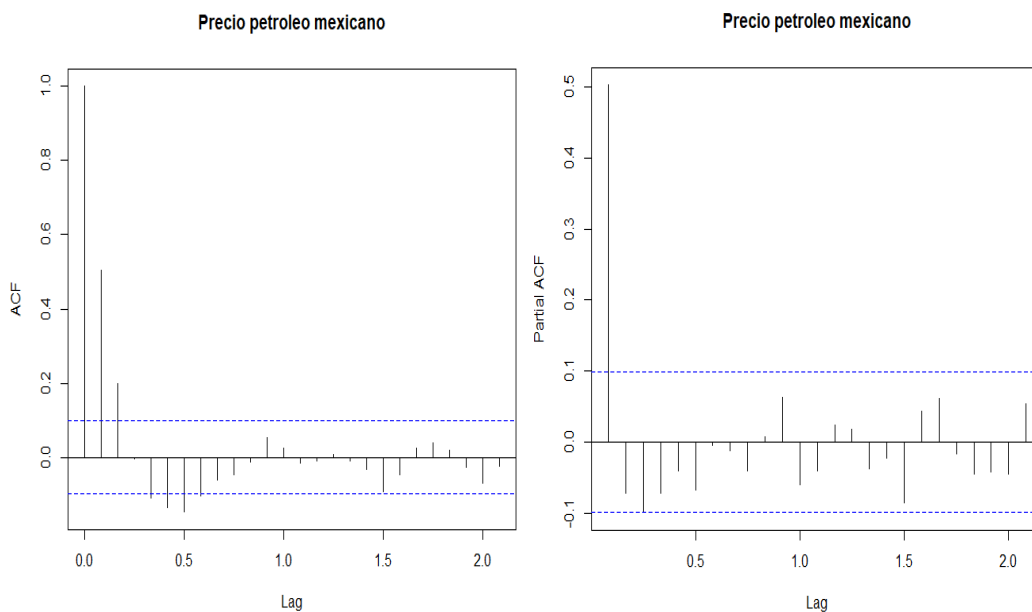
k = Retardos

A través del cálculo de autocorrelaciones, distintos softwares estadísticos como R, permiten crear las gráficas de autocorrelación (ACF) y de autocorrelación parcial (PACF), donde los valores fuera de los límites indican una autocorrelación estadísticamente significativa. Es de esperarse que, en los gráficos de autocorrelación, el comportamiento tienda a cero rápidamente conforme aumenta el número de retardos.

Por ejemplo, el precio del mes de febrero mostrará una correlación estadísticamente significativa con el precio de enero (retardo $k=1$), mientras que el precio de abril comparado con el de enero mostrarán una correlación por debajo de los límites (retardo $k=3$). En los modelos ARIMA los valores del ACF que están fuera de los límites indican el número de términos de promedio móvil del modelo (q), mientras que los valores del PACF fuera de los límites indican el número de términos autorregresivos (p).

En la figura 8 se muestran los gráficos de autocorrelación y autocorrelación parcial del precio del petróleo con los datos de entrenamiento. En el gráfico de autocorrelación ACF se observan seis valores fuera de los límites, lo que indica que el modelo ARIMA tiene $6q$ (medias móviles ponderadas de los errores pasados del pronóstico MA), y en el gráfico de autocorrelación parcial ACP, hay un valor fuera de los límites, por lo tanto, el modelo ARIMA tiene $1p$ (auto regresión de la variable contra valores pasados de ella misma).

Figura 8. Gráficos de autocorrelación (ACF) y autocorrelación parcial (ACP).



Fuente: Elaboración propia.

4. Selección del modelo

Recordando que se realizó una diferencia, el modelo de pronóstico para el precio del petróleo mexicano se trata de un ARIMA (1,1,6). En la práctica, es raro que los valores de p y q excedan a 2, ya que implicarían una fuerte autocorrelación de dos periodos del tiempo que están alejados, por lo que un $q=6$ implica una correlación semestral en el precio del petróleo mexicano. En la figura 9 se muestra el resumen de los resultados arrojados por R donde se arrojan cada uno de los coeficientes de la ecuación del modelo ARIMA. Dentro de los resultados es importante destacar el criterio de información de Akaike (AIC)= 2146.72.

Figura 9. Modelo ARIMA (1,1,6)

```
Call:
arima(x = ARIMAts, order = c(1, 1, 6))

Coefficients:
      ar1      ma1      ma2      ma3      ma4      ma5      ma6
 0.7213 -0.2042 -0.1275 -0.1360 -0.106 -0.0279 -0.0738
s.e.  0.2618  0.2636  0.1436  0.0768  0.050  0.0604  0.0648

sigma^2 estimated as 12.7:  log likelihood = -1065.36,  aic = 2146.72
```

Fuente: Elaboración propia.

El criterio de información de Akaike (AIC) es una medida de la calidad de un modelo en relación con otros modelos. Se utiliza para la selección del modelo. Cuanto menor sea el valor del AIC, mejor será el modelo (Jaramillo y Llamuca, 2022). El AIC es una función del número de parámetros k en un modelo y el valor máximo de la función de verosimilitud (L):

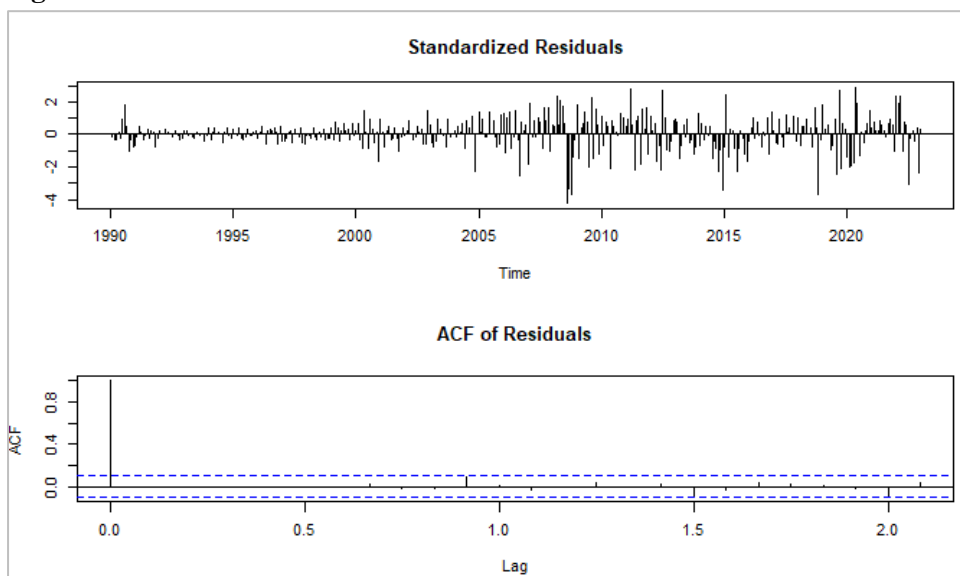
$$AIC = 2k - 2\ln(L) \quad (5)$$

5. Análisis de los residuos

Una vez planteado el modelo es necesario analizar los residuos, es decir, la diferencia entre el pronóstico del dato de prueba realizado por el modelo ARIMA y el valor real. Este análisis debe hacerse con las series de tiempo ajustadas mediante diferencias a fin de evitar alteraciones por la escala original. Se espera que los residuos tengan un comportamiento aleatorio, también conocido como ruido blanco. Otra herramienta de análisis de bondad de ajuste es la gráfica de autocorrelación de los residuos (ACF). La ACF muestra la autocorrelación de los residuos en diferentes rezagos, es decir, compara la autocorrelación de un mes con otro.

En un modelo ARIMA con un buen ajuste de parámetros, los residuos deberían ser ruido blanco, lo que significa que no debería haber ninguna correlación entre ellos. Por lo tanto, en una gráfica ACF para los residuos de un modelo ARIMA bien ajustado, se espera ver picos estadísticamente significativos solo en el rezago 0 (la autocorrelación con sí mismo, que es siempre 1) y cero en los demás rezagos. En la figura 10 se aprecian las gráficas de los residuos estandarizados y de autocorrelación de los residuos, donde se observa el ruido blanco y el pico de autocorrelación solo en el rezago 0, por lo que se asume que el modelo tiene un buen ajuste.

Figura 10. Gráficos de los residuos del modelo ARIMA.



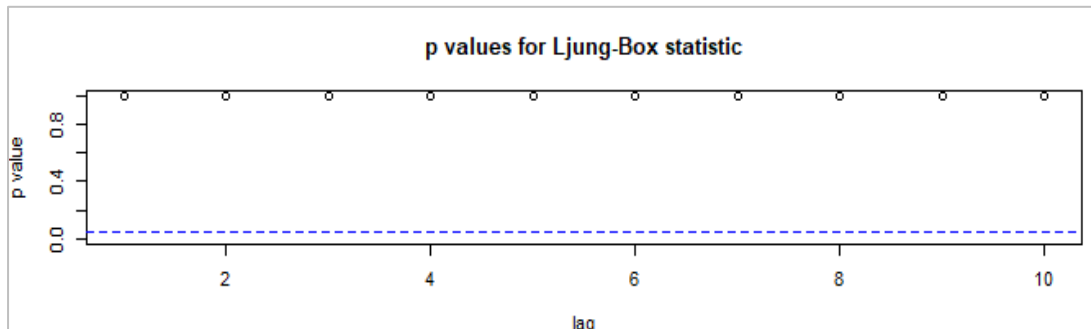
Fuente: Elaboración propia.

Existen pruebas estadísticas específicas para el análisis de los residuos en series de tiempo. Dentro de las más populares se encuentran el test Box-Pierce y su adaptación en el Ljung-Box test, el test de Durbin-Watson y la prueba Breusch–Godfrey (Uyanto 2020). Para este estudio se hace uso de las pruebas Box-Pierce y Ljung-Box test. En ambos casos la hipótesis nula H_0 es que no hay autocorrelación en la serie de tiempo, es decir, que los residuos se distribuyen de forma independiente (tienen ruido blanco). La hipótesis alternativa es que hay una autocorrelación en los residuos.

Cabe recordar que la autocorrelación, que en la práctica no son más que tendencias estacionales, pueden llegar a afectar significativamente los modelos de pronósticos, por lo que se espera que los valores p sean mayor a 0.05 y de esta forma aceptar la hipótesis nula de que los residuos se comportan de manera independiente y por lo tanto no hay autocorrelación. En la figura 11 se muestran los resultados de

ambas pruebas de independencia donde el valor p es mayor a 0.05 y por lo tanto no se puede rechazar H_0 . El test de Ljung-Box se muestra de forma gráfica y el de Box-Pierce de forma numérica.

Figura 11. Test de Ljung-Box y Box-Pierce de los residuos del modelo ARIMA



```
Box-Pierce test
data: residuals(modelo1)
X-squared = 2.0895e-05, df = 1, p-value = 0.9964
```

Fuente: Elaboración propia.

RESULTADOS

6. Realización de pronósticos

En la figura 12, se muestran los pronósticos tal cual fueron arrojados por RStudio. Una de las ventajas de R, es que calcula automáticamente intervalos de confianza para los pronósticos a un nivel de confiabilidad del 80% y 95%.

Figura 12. Pronósticos de febrero de 2023 a enero 2024 arrojados por RStudio.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Feb 2023	66.86814	62.27497	71.46132	59.84349	73.89280
Mar 2023	67.50566	59.11840	75.89291	54.67845	80.33286
Apr 2023	68.59041	56.98054	80.20027	50.83465	86.34617
May 2023	69.74448	55.53374	83.95523	48.01103	91.47794
Jun 2023	70.77092	54.50159	87.04024	45.88913	95.65270
Jul 2023	71.59005	53.68881	89.49129	44.21246	98.96764
Aug 2023	72.19150	52.97457	91.40843	42.80175	101.58126
Sep 2023	72.60070	52.29375	92.90764	41.54390	103.65749
Oct 2023	72.85720	51.61735	94.09706	40.37364	105.34076
Nov 2023	73.00206	50.93694	95.06718	39.25637	106.74775
Dec 2023	73.07128	50.25403	95.88852	38.17532	107.96724
Jan 2024	73.09330	49.57366	96.61294	37.12311	109.06348

Fuente: Elaboración propia.

7. Evaluación del modelo

Para validar el modelo es necesario comprobarlo con los datos de prueba, en este caso se harán los pronósticos del precio del barril del petróleo de febrero de 2023 a enero de 2024. En la tabla 2 se muestran los valores reales y los pronósticos para dichos meses, también se muestra el error porcentual en cada mes y el error porcentual absoluto medio (MAPE) de todos los meses.

Tabla 2. Evaluación de los pronósticos del precio de barril de petróleo.

Pronósticos precio del barril de petróleo mexicano			
MODELO ARIMA (1,1,6)			
(Febrero 2023 a Enero 2024)			
Fecha	Real	Pronóstico	Error
feb-23	67.10	66.41	1.0%
mar-23	63.32	66.64	5.2%
abr-23	70.51	67.63	4.1%
may-23	63.55	68.39	7.6%
jun-23	64.61	69.52	7.6%
jul-23	70.91	70.27	0.9%
ago-23	78.03	70.81	9.3%
sep-23	85.58	71.20	16.8%
oct-23	81.34	71.48	12.1%
nov-23	73.58	71.68	2.6%
dic-23	67.93	71.83	5.7%
ene-24	69.74	71.93	3.1%
		MAPE	6.339%

Fuente: Elaboración propia.

Un método popular para evaluar los pronósticos es el error porcentual absoluto medio (MAPE), representado en la ecuación 6. El MAPE tiene la ventaja de evaluar los errores en términos de porcentajes. Se calcula primero dividiendo cada valor absoluto del error de pronóstico entre el valor real de la serie de tiempo en cada periodo, posteriormente se promedian todos los errores porcentuales absolutos y se multiplican por 100. La desventaja de la métrica MAPE es que, dada la fórmula, no se pueden tener valores de la serie de tiempo iguales a cero.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|Y_t - \hat{Y}_t|}{|Y_t|} * 100 \quad (6)$$

Por su parte, Lewis (1982) propuso una tabla para la evaluación de los valores MAPE típicos en datos industriales. Los intervalos y su interpretación se muestran en la tabla 2.4, misma que se ha vuelto una referencia popular hasta la fecha para la evaluación de pronósticos.

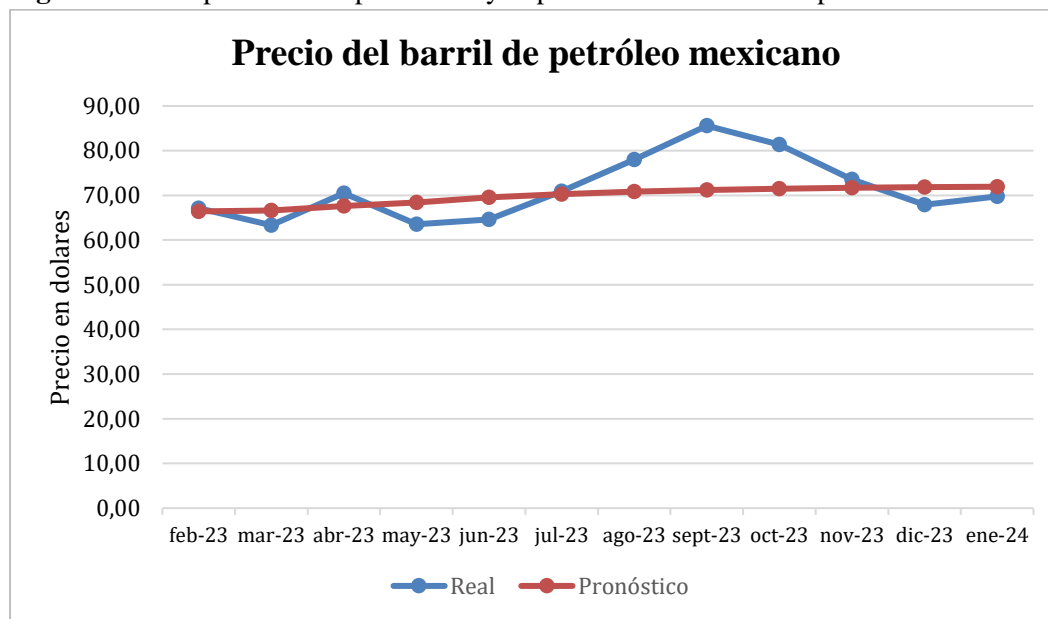
Tabla 3. Interpretación de los valores típicos del error MAPE en datos industriales.

MAPE	Interpretación
< 10	Pronóstico preciso
10-20	Buen pronóstico
20-50	Pronóstico razonable
> 50	Pronóstico inexacto

Fuente: Lewis (1982)

Tomando como referencia la tabla de Lewis, se evalúa el modelo y se considera que genera pronósticos precisos al presentar un error MAPE de 6.339%. En la figura 13 se muestran las series de tiempo de los valores reales y los pronósticos del precio del petróleo.

Figura 12. Comparación del pronóstico y el precio real del barril de petróleo.



El modelo pronosticó un comportamiento estable durante todo el año, sin embargo, al observar el comportamiento real, se aprecia un aumento lineal de los meses junio a septiembre, mismo mes en que se obtuvo el mayor error MAPE (16.8%). Al consultar las fuentes especializadas que expliquen dicho fenómeno, se tiene como una de las explicaciones que “los precios internacionales del petróleo repuntaron hasta 34% en el tercer trimestre de este año (2023), debido principalmente a los recortes de Arabia Saudita y Rusia, dos de los principales productores y exportadores de la materia prima en el

mundo, anunciados en junio y que provocaron un déficit en la oferta” (Díaz Mora, 2023). De manera directa la mezcla mexicana repuntó de igual forma un 34.54%. Al estabilizarse los precios de petróleo para fines de 2024, Rusia y su aliado Arabia Saudita decidieron continuar con los recortes de oferta de producción de petróleo hasta mediados de 2024 a fin de impulsar los precios nuevamente, meta que se ha propuesto Rusia cuya economía está orientada a apoyar la ofensiva militar en Ucrania (Banca y Negocios, 2024). Lo anterior es un ejemplo de que los modelos de pronósticos de series de tiempo en el área de la economía se ven afectados por sucesos aleatorios imprevistos que responden, más que a variables cuantitativas externas, a decisiones y factores geopolíticos que influyen directamente en el comportamiento de la serie.

Una de las funciones disponibles dentro de R, y que sirven para comprobar la pertinencia de la metodología propuesta, es la función “autoarima”, dentro de la librería “forecast”, la cual realiza de forma automática una serie de iteraciones para encontrar el modelo que mejor se ajuste al criterio a especificar, como puede ser el criterio de información de Akaike (Awan y Aslam, 2020). En la figura 13 se muestran las iteraciones realizadas por el algoritmo automático, que arroja el modelo con el menor AIC. El resultado es un modelo ARIMA (2,1,1).

Figura 13. Elección del modelo mediante la función autoarima.

```
Fitting models using approximations to speed things up...
ARIMA(2,1,2) (1,0,1) [12] with drift : 2151.942
ARIMA(0,1,0) with drift : 2258.283
ARIMA(1,1,0) (1,0,0) [12] with drift : 2151.448
ARIMA(0,1,1) (0,0,1) [12] with drift : 2164.272
ARIMA(0,1,0) : 2256.634
ARIMA(1,1,0) with drift : 2145.501
ARIMA(1,1,0) (0,0,1) [12] with drift : 2147.42
ARIMA(1,1,0) (1,0,1) [12] with drift : 2153.434
ARIMA(2,1,0) with drift : 2146.443
ARIMA(1,1,1) with drift : 2146.133
ARIMA(0,1,1) with drift : 2162.272
ARIMA(2,1,1) with drift : 2141
ARIMA(2,1,1) (1,0,0) [12] with drift : 2148.331
ARIMA(2,1,1) (0,0,1) [12] with drift : 2142.994
ARIMA(2,1,1) (1,0,1) [12] with drift : 2150.145
ARIMA(3,1,1) with drift : 2143.521
ARIMA(2,1,2) with drift : 2142.464
ARIMA(1,1,2) with drift : 2145.97
ARIMA(3,1,0) with drift : 2145.299
ARIMA(3,1,2) with drift : 2145.593
ARIMA(2,1,1) with drift : 2139.419
ARIMA(2,1,1) (1,0,0) [12] : 2146.611
ARIMA(2,1,1) (0,0,1) [12] : 2141.408
ARIMA(2,1,1) (1,0,1) [12] : 2148.445
ARIMA(1,1,1) : 2144.224
ARIMA(2,1,0) : 2144.559
ARIMA(3,1,1) : 2141.843
ARIMA(2,1,2) : 2140.788
ARIMA(1,1,0) : 2143.588
ARIMA(1,1,2) : 2144.052
ARIMA(3,1,0) : 2143.458
ARIMA(3,1,2) : 2143.942

Now re-fitting the best model(s) without approximations...
ARIMA(2,1,1) : 2140.269
Best model: ARIMA(2,1,1)
```

Fuente: Elaboración propia.



Una vez elegido el modelo de forma automática, se proceden a realizar los pronósticos, los cuales se muestran en la figura 14, con sus respectivos intervalos de confianza al 80% y 95%.

Figura 14. Pronósticos realizados por el modelo autoarima.

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
Feb 2023	66.86814	62.27497	71.46132	59.84349	73.89280
Mar 2023	67.50566	59.11840	75.89291	54.67845	80.33286
Apr 2023	68.59041	56.98054	80.20027	50.83465	86.34617
May 2023	69.74448	55.53374	83.95523	48.01103	91.47794
Jun 2023	70.77092	54.50159	87.04024	45.88913	95.65270
Jul 2023	71.59005	53.68881	89.49129	44.21246	98.96764
Aug 2023	72.19150	52.97457	91.40843	42.80175	101.58126
Sep 2023	72.60070	52.29375	92.90764	41.54390	103.65749
Oct 2023	72.85720	51.61735	94.09706	40.37364	105.34076
Nov 2023	73.00206	50.93694	95.06718	39.25637	106.74775
Dec 2023	73.07128	50.25403	95.88852	38.17532	107.96724
Jan 2024	73.09330	49.57366	96.61294	37.12311	109.06348

Fuente: Elaboración propia.

Posteriormente, se procede a hacer la evaluación con los datos de prueba, detalla en la tabla 4, a fin de comparar la metodología propuesta con el modelo automático. Los resultados muestran que el error MAPE obtenido en el modelo automático (6.345%) es prácticamente igual que el modelo propuesto (6.339%), por lo que se confirma la consistencia metodológica seguida para el pronóstico del precio del barril de petróleo mexicano.

Tabla 4. Evaluación de pronósticos del modelo automático ARIMA (2,1,1).

Pronósticos precio del barril de petróleo mexicano			
MODELO ARIMA (2,1,1)			
(Febrero 2023 a Enero 2024)			
Fecha	Real	Pronóstico	Error
feb-23	67.10	66.87	0.3%
mar-23	63.32	67.51	6.6%
abr-23	70.51	68.59	2.7%
may-23	63.55	69.74	9.7%
jun-23	64.61	70.77	9.5%
jul-23	70.91	71.59	1.0%
ago-23	78.03	72.19	7.5%
sep-23	85.58	72.60	15.2%
oct-23	81.34	72.86	10.4%
nov-23	73.58	73.00	0.8%
dic-23	67.93	73.07	7.6%
ene-24	69.74	73.09	4.8%
		MAPE	6.345%

DISCUSIÓN

Con más de cincuenta años desde su creación, el modelo ARIMA se ha mantenido como una alternativa confiable de realización de pronósticos. En el caso de la industria petrolera se ha utilizado con buenos resultados en distintas investigaciones como la de [Sokkalingam y otros \(2022\)](#), donde se propone el uso del modelo ARIMA para pronosticar el precio volátil del petróleo en Malasia, tras el cambio de un mecanismo automático de precios a un sistema de flotación administrada en 2016.

Gasper y Mbwambob (2023) hicieron lo propio para pronosticar los precios del petróleo en Tanzania. Al igual que el presente estudio, dictaminaron que la guerra de Ucrania fue uno de los factores que tuvo un fuerte impacto en el aumento de los precios del petróleo, sin embargo, concluyen que los modelos ARIMA, con sus limitaciones, pueden continuar pronosticando con efectividad el comportamiento del precio del combustible.

De acuerdo con Purohit y Panigrahi (2024), una de las deficiencias de los modelos ARIMA es que estos asumen la serie temporal como el resultado de un fenómeno lineal, lo que proporciona pronósticos deficientes para las series temporales complejas, no lineales y no gaussianas. Lo anterior se comprueba con los resultados de los pronósticos realizados del precio del petróleo mexicano, donde el modelo ARIMA tuvo un comportamiento lineal, lo que no ajustó correctamente el comportamiento de los datos reales.

Si bien, en la presente investigación se ha utilizado como parámetro de evaluación el MAPE, existen otros parámetros para la evaluación de pronósticos, como el error cuadrático medio (MSE). En este método, los errores se elevan al cuadrado y luego se promedian con el total de observaciones. Esta técnica “castiga” a los errores grandes, ya que los eleva al cuadrado y los maximiza, sin embargo, esto es útil para evaluar si un modelo de pronósticos no es homogéneo en sus residuos (es preferible un modelo con errores moderados, a un modelo con errores bajos y altos, aunque en promedio sean iguales).

$$MSE = \frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2 \quad (7)$$

Nótese que la ecuación 7 es similar a la fórmula de la varianza poblacional, por lo que en pocas palabras el MSE es la varianza del error de pronóstico. A partir de la MSE se puede calcular la raíz

cuadrada del error cuadrático medio (RMSE), representada en la ecuación 8. El RMSE equivale a la desviación estándar del error de pronóstico. Tanto MSE y RMSE sancionan a los errores de pronóstico grandes, sin embargo, el RMSE tiene la ventaja de que esta en las mismas unidades que la variable de la serie de tiempo, lo que facilita su interpretación, por lo que es una de las métricas más populares para la evaluación del error de pronóstico.

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (Y_t - \hat{Y}_t)^2} \quad (8)$$

En la tabla 5 se muestran los cálculos del MSE y RMSE de los pronósticos realizados durante el periodo febrero del año 2023 a enero de 2024 de la mezcla de petróleo mexicana.

Tabla 5. Evaluación de pronósticos con las métricas MSE y RMSE.

Fecha	Valor Real Y_t	Pronóstico \hat{y}_t	Error	$(Y_t - \hat{y}_t)^2$
feb-23	67.10	66.41	1.0%	0.47242018
mar-23	63.32	66.64	5.2%	10.9733144
abr-23	70.51	67.63	4.1%	8.24159737
may-23	63.55	68.39	7.6%	23.3492996
jun-23	64.61	69.52	7.6%	24.147656
jul-23	70.91	70.27	0.9%	0.40683732
ago-23	78.03	70.81	9.3%	52.0928878
sep-23	85.58	71.20	16.8%	206.775869
oct-23	81.34	71.48	12.1%	97.2998246
nov-23	73.58	71.68	2.6%	3.59995506
dic-23	67.93	71.83	5.7%	15.1956302
ene-24	69.74	71.93	3.1%	4.79398191
			MSE	37.279
			RMSE	6.106

Fuente: Elaboración propia.

Chen, Twycross y Garibaldi (2017) sugieren que, aunque las métricas MAPE y RMSE son las más utilizadas dentro del campo de las series temporales, estas muestran deficiencias y sesgos particulares para evaluar los modelos de pronósticos, sobre todo cuando se presentan valores atípicos que afectan la evaluación dependiendo de la escala de la variable o si se ven implicados valores cercanos a cero. En su estudio, los autores analizan toda una variedad de métricas para evaluar modelos de predicción y

concluyen proponiendo una nueva medida de evaluación denominada Error Absoluto Relativo Limitado Medio No Escalado (UMBRAE), que combina las mejores características de varias medidas alternativas, para abordar los problemas comunes de las métricas ya existentes. A la fecha, no ha habido un consenso académico sobre el uso de la métrica UMBRAE y los propios autores reconocen que sus propiedades estadísticas no han sido bien estudiadas.

A partir de la realización de competencias de modelos de pronósticos, se ha llegado a la aseveración empírica de que la combinación de componentes de distintos modelos, tanto estadísticos tradicionales, como los que utilizan modernas técnicas de inteligencia artificial, genera mejores resultados que utilizar los modelos de forma separada. El futuro de la investigación de pronósticos se encuentra en la creación de modelos híbridos, mismos que contienen características obtenidas de los modelos estadísticos, como es el caso de ARIMA, por lo que su aplicación se mantiene vigente (Fierro, Castillo y Torres, 2022).

CONCLUSIONES

ARIMA son las siglas en inglés del modelo autorregresivo integrado de media móvil. Es un método popular de pronóstico de series de tiempo utilizado en estadística para la evaluación de variables económicas con fluctuación en el tiempo. Los modelos ARIMA son capaces de capturar un conjunto de diferentes estructuras temporales en datos de series temporales, incluidas tendencias, estacionalidad y autocorrelación. Una de las mayores limitantes es que debido a su modelación sencilla, los modelos ARIMA asumen la linealidad en la autoregresión (valores pasados). A pesar de que las series temporales pueden ser generalmente no lineales, en muchos casos, los patrones pueden ser capturados de manera efectiva mediante ARIMA sobre todo cuando se presentan tendencias marcadas. En el presente estudio, el modelo generó pronósticos precisos del precio del petróleo mexicano, lo que se ve reflejado en un error MAPE de 6.339%, y que fue comprobado con la función `autoarima` en R, al no encontrar mejor resultado de forma iterativa.

A pesar de esto último, hubo un mayor error de pronóstico de julio a noviembre de 2023, debido a las tensiones geopolíticas provocadas por la guerra de Ucrania que ocasionaron un aumento inesperado en el precio del petróleo. Lo anterior refleja la importancia de la capacidad del analista de pronósticos para realizar los ajustes en las estimaciones, con base en su experiencia e información disponible del contexto en el que actúa la variable y que no puede ser representada matemáticamente mediante un modelo. Esta

forma de previsión se le conoce como “pronósticos cualitativos”, basados generalmente en la opinión de un grupo de expertos. Trabajar a la par ambos enfoques: cualitativo y cuantitativo, garantizaran el éxito de un sistema de pronósticos.

Contribuciones a futuras líneas de investigación

El presente trabajo de investigación constituye un caso práctico del uso de la metodología de pronósticos con un enfoque didáctico e informativo del análisis de series de tiempo y la previsión estadística mediante un modelo ARIMA. Dicha propuesta metodológica puede ser replicada en el análisis de cualquier variable que se analice en un periodo de tiempo determinado. La gratuidad y acceso a documentación sobre R y Rstudio, permite el uso de dicha herramienta para el análisis potente de datos, en este caso, series de tiempo. Si bien, el software cuenta con una función automática para crear e implementar distintos modelos, como ARIMA, se recomienda en un principio diseñar el algoritmo a cuenta propia para comprender el funcionamiento de un modelo de pronósticos. Por último, se alienta a que una vez utilizado y dominado el modelo ARIMA, el analista o investigador se adentre al uso de otras alternativas, como lo son las redes neuronales y posteriormente a la hibridación de dichos modelos, a fin de buscar aquellas metodologías que generen mayor precisión en las previsiones a futuro de las variables de estudio.

REFERENCIAS BIBLIOGRAFICAS

- Aguirre Jaime, A. (1994). *Introducción al tratamiento de series temporales: Aplicación a las ciencias de la salud*. Ediciones Diaz de Santos.
- Awan, T. M., & Aslam, F. (2020). Prediction of daily covid-19 cases in european countries using automatic arima model. *Journal of Public Health Research*, 9(3), jphr.2020.1765. <https://doi.org/10.4081/jphr.2020.1765>
- Banca y Negocios. (2024, marzo 3). Arabia Saudita y Rusia prolongan los recortes de su producción de crudo para impulsar los precios. *Banca y Negocios*. <https://www.bancaynegocios.com/arabia-saudita-y-rusia-prolongan-los-recortes-de-su-produccion-de-crudo-para-impulsar-los-precios/>
- Chen, C., Twycross, J., & Garibaldi, J. M. (2017). A new accuracy measure based on bounded relative error for time series forecasting. *PLOS ONE*, 12(3), e0174202. <https://doi.org/10.1371/journal.pone.0174202>



- Cortés Patiño, J. M. (2011). Aplicación de series de tiempo en el monitoreo estructural [Tesis de licenciatura, Universidad Nacional Autónoma de México].
<https://ru.dgb.unam.mx/handle/20.500.14330/TES01000667008>
- Díaz Mora, S. (2023, octubre 1). Petróleo sube hasta 34% en el tercer trimestre de 2023. *El Economista*.
<https://www.economista.com.mx/mercados/Petroleo-sube-hasta-34-en-el-tercer-trimestre-de-2023-20231001-0048.html>
- Fierro Torres, C. Á., Castillo Pérez, V. H., & Torres Saucedo, C. I. (2022). Análisis comparativo de modelos tradicionales y modernos para pronóstico de la demanda: Enfoques y características. *RIDE Revista Iberoamericana para la Investigación y el Desarrollo Educativo*, 12(24).
<https://doi.org/10.23913/ride.v12i24.1203>
- Fullana Fuster, P. (2020). *Diseño de la red de la cadena de suministro de una marca de coches eléctricos para su entrada en Europa* [Tesis de maestría, Universitat politècnica de València].
<https://acortar.link/Q1dnQN>
- García Díaz, J.C. (2016). Predicción en el dominio del tiempo: Análisis de series temporales para ingenieros. Universitat politècnica de València.
- Gasper, L., & Mbwambo, H. (2023). Forecasting crude oil prices by using arima model: Evidence from tanzania. *Journal of Accounting Finance and Auditing Studies (JAFAS)*, 2.
<https://doi.org/10.32602/jafas.2023.017>
- Hanke, J. E., & Wichern, D. W. (2010). *Pronósticos en los negocios* (9a ed.). Pearson Educación.
- Jaramillo, M., & Llamuca, S. (2022). A proposed model for electricity demand forecasting in Ecuador considering Akaike Criterion. En Á. Rocha, P. C. López-López, & J. P. Salgado-Guerrero (Eds.), *Communication, Smart Technologies and Innovation for Society* (Vol. 252, pp. 345-355). Springer Singapore. https://doi.org/10.1007/978-981-16-4126-8_32
- Lewis, C. D. (1982). *Industrial and business forecasting methods: A practical guide to exponential smoothing and curve fitting*. Butterworth Scientific.
- Makridakis, S. G., Wheelwright, S. C., & Hyndman, R. J. (2005). *Forecasting: Methods and applications* (Third edition, [Wiley student edition]). Wiley.



- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting* (Second edition). Wiley.
- Olvera Vázquez, J. (2020). *Aplicación de métodos de pronósticos híbridos en base a métodos clásicos y de optimización inteligente en problemas emergentes tipo COVID 19* [Tesis de maestría, Instituto Tecnológico de Ciudad Madero]. <https://rinacional.tecnm.mx/handle/TecNM/941>
- Purohit, S. K., & Panigrahi, S. (2024). Novel deterministic and probabilistic forecasting methods for crude oil price employing optimized deep learning, statistical and hybrid models. *Information Sciences*, 658, 120021. <https://doi.org/10.1016/j.ins.2023.120021>
- Render, B., & Heizer, J. H. (2014). *Principios de administración de operaciones* (9a ed.). Pearson.
- Roza, A., Violita, E. S., & Aktivani, S. (2022). Study of inflation using stationary test with augmented dickey fuller & phillips-peron unit root test (Case in bukittinggi city inflation for 2014-2019). *EKSAKTA: Berkala Ilmiah Bidang MIPA*, 23(02), 106-116. <https://doi.org/10.24036/eksakta/vol23-iss02/303>
- Sokkalingam, R., Sarpong-Streeter, R. M. N. Y., Othman, M., Daud, H., & Owusu, D. A. (2021). *Forecasting petroleum fuel price in malaysia by arima model*. En S. A. Abdul Karim, M. F. Abd Shukur, C. Fai Kait, H. Soleimani, & H. Sakidin (Eds.), *Proceedings of the 6th International Conference on Fundamental and Applied Sciences* (pp. 671-678). Springer Nature Singapore. https://doi.org/10.1007/978-981-16-4513-6_58
- Uyanto, S. S. (2020). Power comparisons of five most commonly used autocorrelation tests. *Pakistan Journal of Statistics and Operation Research*, 16(1), 119-130. <https://doi.org/10.18187/pjsor.v16i1.2691>
- Vargas, L y Mesa-Fúquen, E. (2021). *Introducción al análisis de datos con RStudio*. Cenipalma.