

Ajuste de un modelo ARIMA para los precios de exportación del aguacate calibre 60

Belem A. Aguilar ¹, Roberto S. Acosta Abreu².

¹Licenciatura en Física y Matemáticas, ESFM-Instituto Politécnico Nacional, Ciudad de México., México., Teléfono (55) 3494-9649

²Departamento de Matemáticas, ESFM-Instituto Politécnico Nacional, Ciudad de México., México., Teléfono (55) 1307-1412



RESUMEN/ABSTRACT

El propósito de este artículo es analizar los precios de exportación a EE.UU. Del aguacate calibre 60, desarrollar de manera resumida la teoría de los modelos ARIMA a la vez que se va obteniendo un modelo resultante de la metodología Box-Jenkins no estacional. El documento consiste en las justificaciones necesarias para el uso de la teoría elegida, el análisis de sus elementos; comparación de modelos AR, MA, ARMA y finalmente ARIMA. Obteniendo un modelo ARIMA(2,1,0) cuya confiabilidad se limita a un par de pasos por delante.

INTRODUCCIÓN

Para este documento se utilizó R para procesar, visualizar y analizar los precios de exportación a EE.UU. del aguacate calibre 60. Los datos fueron obtenidos de la revista virtual MODULA del Consejo Nacional de Productores de Aguacate A.C. (CONAPA) [1]. Se escogieron este conjunto de precios debido a una enorme cantidad de factores que los vuelven estocásticos entre temporadas, además de que estos varían de acuerdo al peso del aguacate. Se eligió un calibre estándar con la esperanza de que este modelo guíe a un modelo para cada uno de los demás calibres. Los datos comprendieron un periodo de tres años, comenzando en enero del año 2019 y concluyendo a finales de diciembre del año 2021.

Con la hipótesis de varianza constante, se siguió la metodología Box-Jenkins no estacional para la obtención del ajuste ARIMA(2,1,0) a la serie de precios, con el objetivo de obtener pronósticos de un par de pasos adelante confiables.

METODOLOGÍA

La metodología a usar es Box-Jenkins (Box y J.M. Jenkins, 1970's), la cual consiste en cinco etapas, descritas a continuación.

ESTACIONARIEDAD

Las series de tiempo se clasifican en estacionarias y no estacionarias. De forma intuitiva, una serie es estacionaria si sus propiedades estadísticas son esencialmente constantes a través del tiempo. Así, el primer paso de la metodología es determinar si la serie es estacionaria, y en caso de no serlo, aplicar una transformación que induzca ello. Las Fig.1 y 2 muestran la serie de precios y su respectivo autocorrelograma (ACF), este segundo siendo el gráfico con el que se determina la estacionariedad.

Analizando los gráficos se concluyó que se requeriría de una transformación de datos, siendo las primeras diferencias la transformación elegida, cuyo ACF de serie diferenciada se muestra en la Fig.3

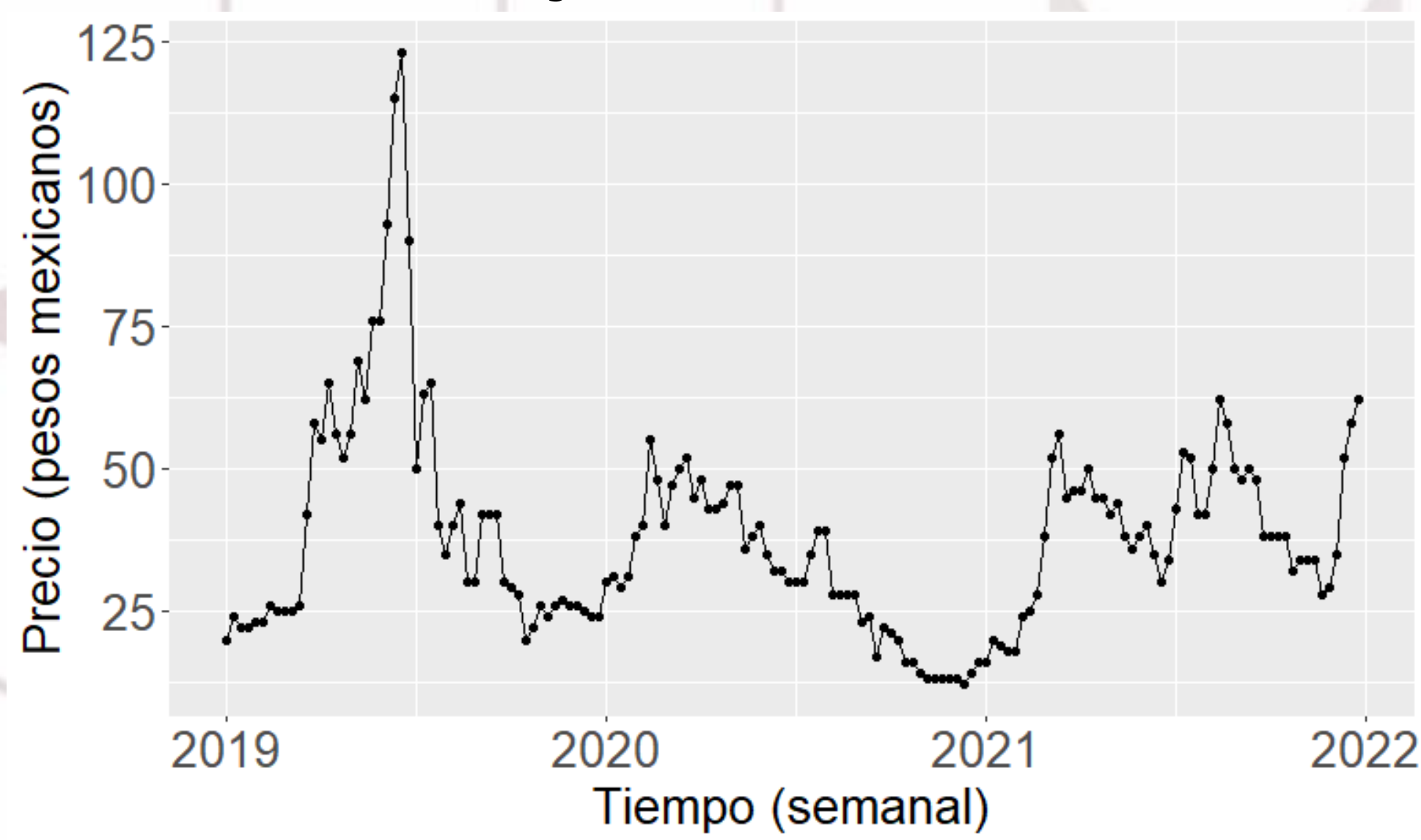


Figura 1. Serie de precios de exportación del aguacate calibre 60.

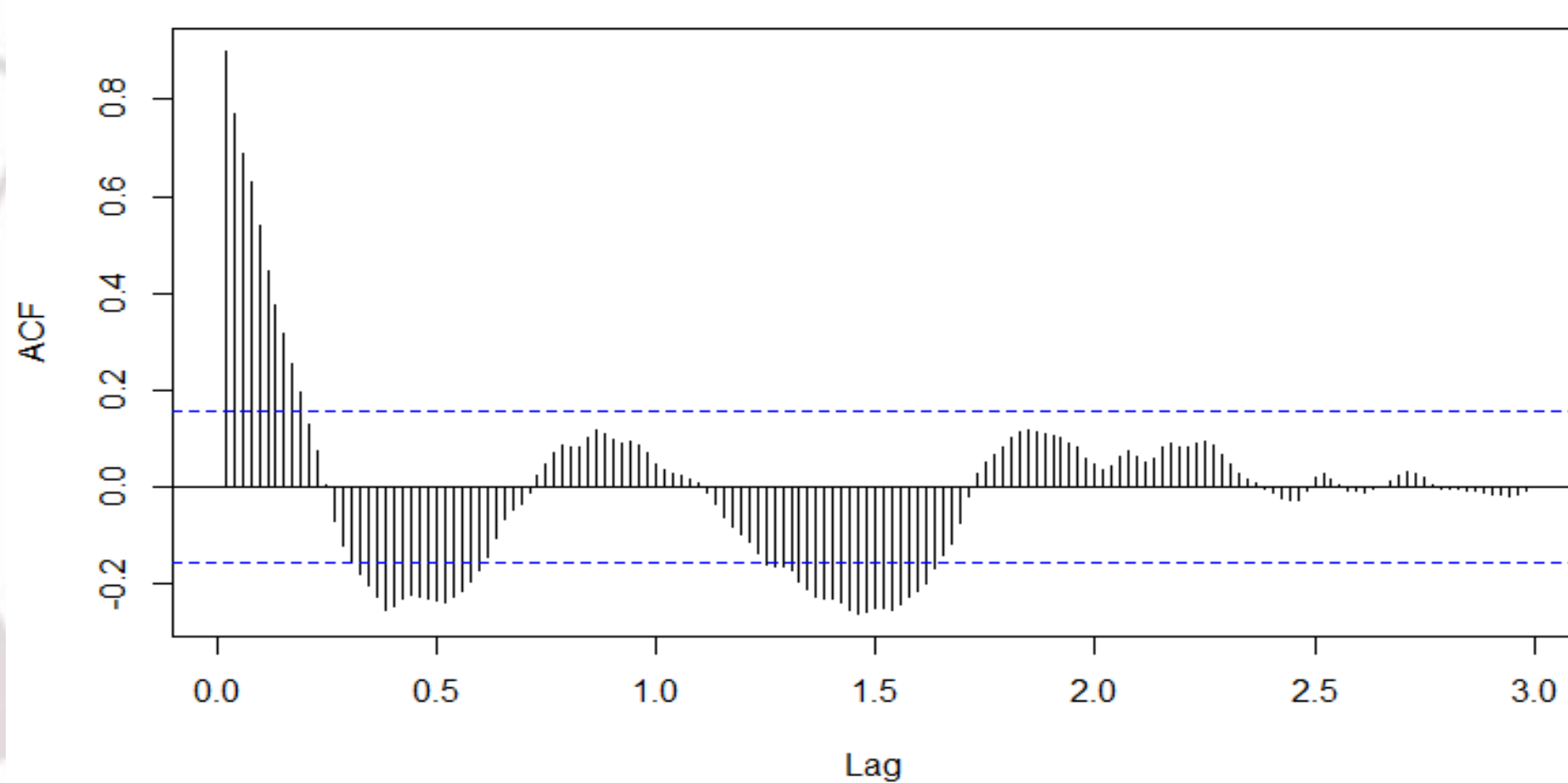


Figura 2. ACF de la serie de precios

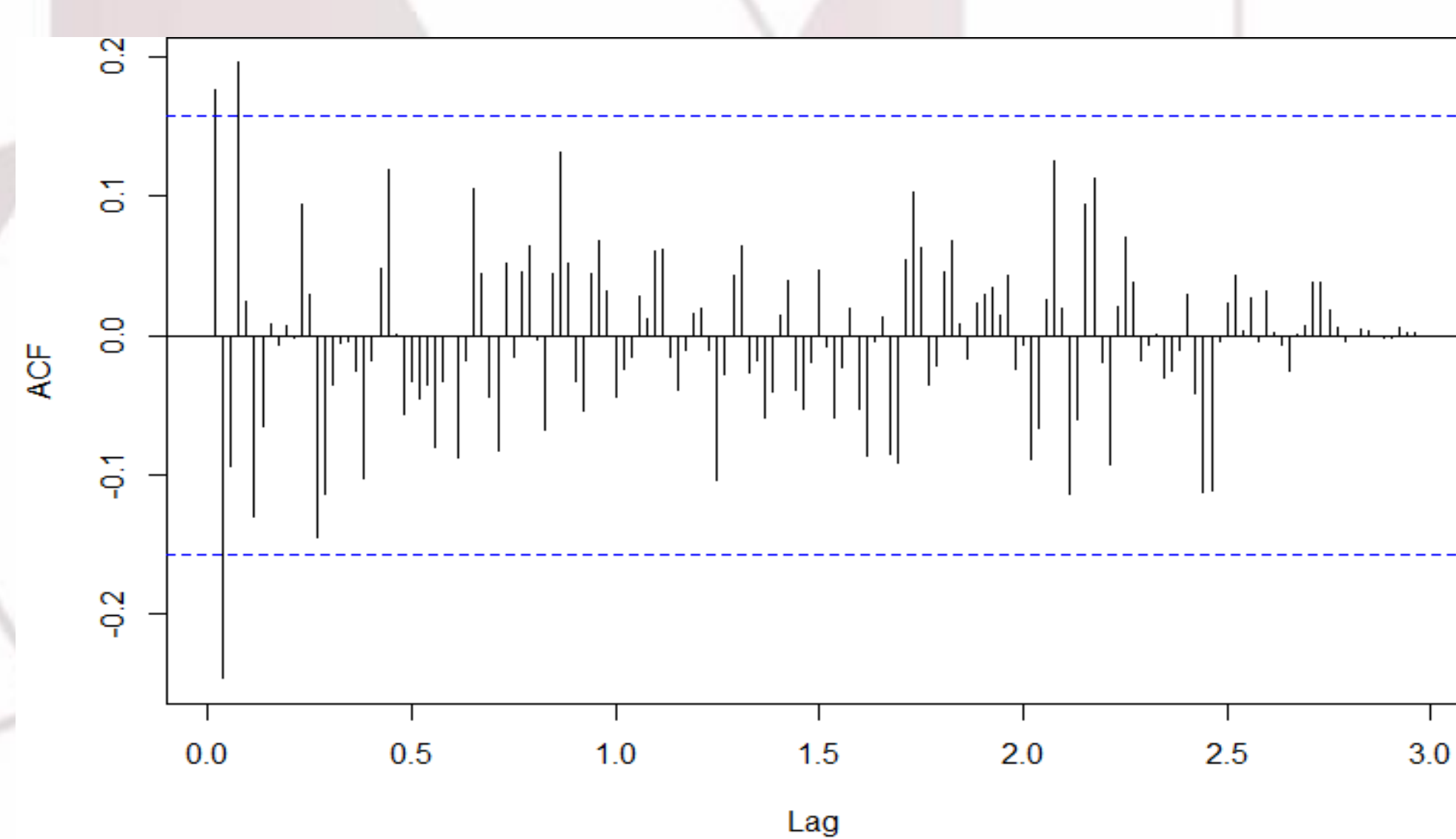


Figura 3. ACF de la serie diferenciada

CONCLUSIONES

- En general, el modelo obtenido simula de manera suficiente el comportamiento de la serie de precios para un par de pasos adelante, con intervalos de confianza bastante confiables.
- Se recomienda seguir ajustando conforme surjan nuevos datos debido a que la información estudiada contiene eventos externos únicos, como una temporada de escasas (mostrada a mediados del año 2019) y los efectos tardíos de una pandemia (mediados del 2020).
- El ajuste de un modelo ARIMA(2,1,2) resulta bastante similar al estudiado aquí, por lo que se optó por el modelo ARIMA(2,1,0).

IDENTIFICACIÓN

Partiendo de las ideas de que las ACF y PACF obtenidos se asemejan a una o varias ACF y PACF teóricas y estas a su vez están asociadas a un modelo ARMA, la Fig. 4 muestra la PACF de la serie diferenciada. Y a partir de ambos gráficos se proponen los modelos ARIMA(4,1,0), ARIMA(0,1,4) y ARIMA(2,1,2).

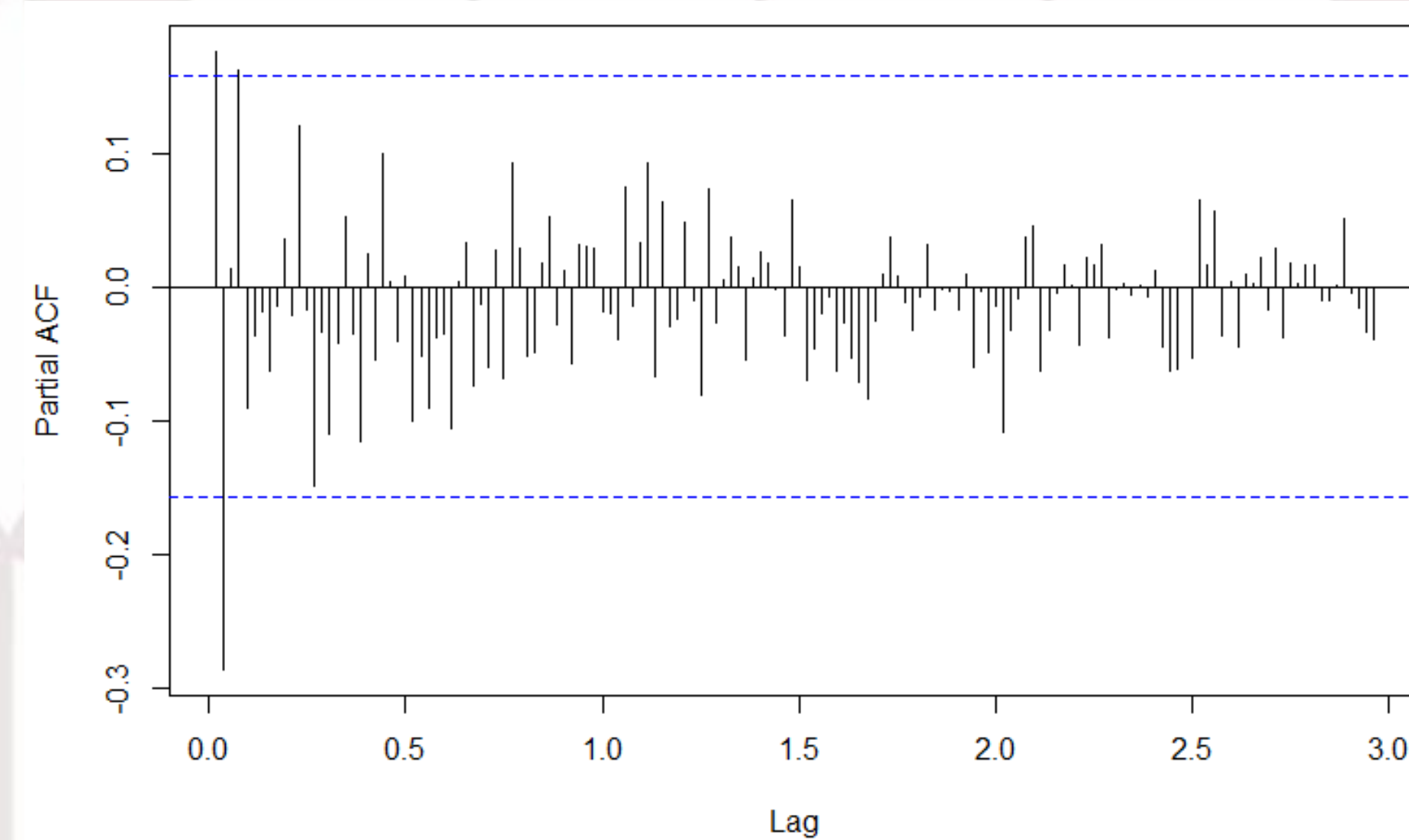


Figura 4. PACF de la serie diferenciada.

ESTIMACIÓN

La Tabla I muestra las estimaciones de los coeficientes de los posibles modelos. El criterio principal que se utiliza es el llamado "Estimación de máxima verosimilitud", el cual se utiliza en "Criterio de Información de Akaike (AIC)" para la selección del modelo. Los demás criterios que se consideran para la elección del modelo son la "parsimonia de parámetros", que los coeficientes del modelo sean "estadísticamente significativos" y aquel con grado de ajuste RMSE de menor valor.

Siguiendo todos estos puntos, se redujeron los modelos de la Tabla I a los modelos de la Tabla I. Del cual se seleccionó el ARIMA(2,1,0) por tener menor AIC.

Modelo	Estimaciones	Error estándar	Valor t	Valor p
ARIMA(2,1,0)				
AIC = 1057.4	$\alpha_1 = 0.2274$	0.0769	2.9570	0.0031
RMSE = 7.1625	$\alpha_2 = -0.2844$	0.0767	3.7079	0.0002
ARIMA(0,1,1)				
AIC = 1064.3	$\beta_1 = 0.2971$	0.0839	3.5411	0.0003
RMSE = 7.373377				

Tabla I. Tabla de modelos reducidos.

EVALUACIÓN

La siguiente etapa es la evaluación. En este punto se verifica la eficiencia del modelo y se decide si es estadísticamente adecuado; esta investigación incluye el análisis del ACF de los residuos (los cuales se esperan sean ruido blanco), de residuos estandarizados, una gráfica quantil-quantil, así como un test con el método estadístico Q-estadística Ljung-Box-Pierce.

La Fig 5 muestra un resumen de todas estas evaluaciones realizadas al ajuste ARIMA(2,1,0), de donde podemos concluir que las aprueba sin mayores complicaciones.

PRONÓSTICO

La Tabla II muestra el pronóstico realizado por el modelo para 4 pasos por delante (equivalente al siguiente mes). La Fig 6 muestra el ajuste final del modelo a la serie de precios.

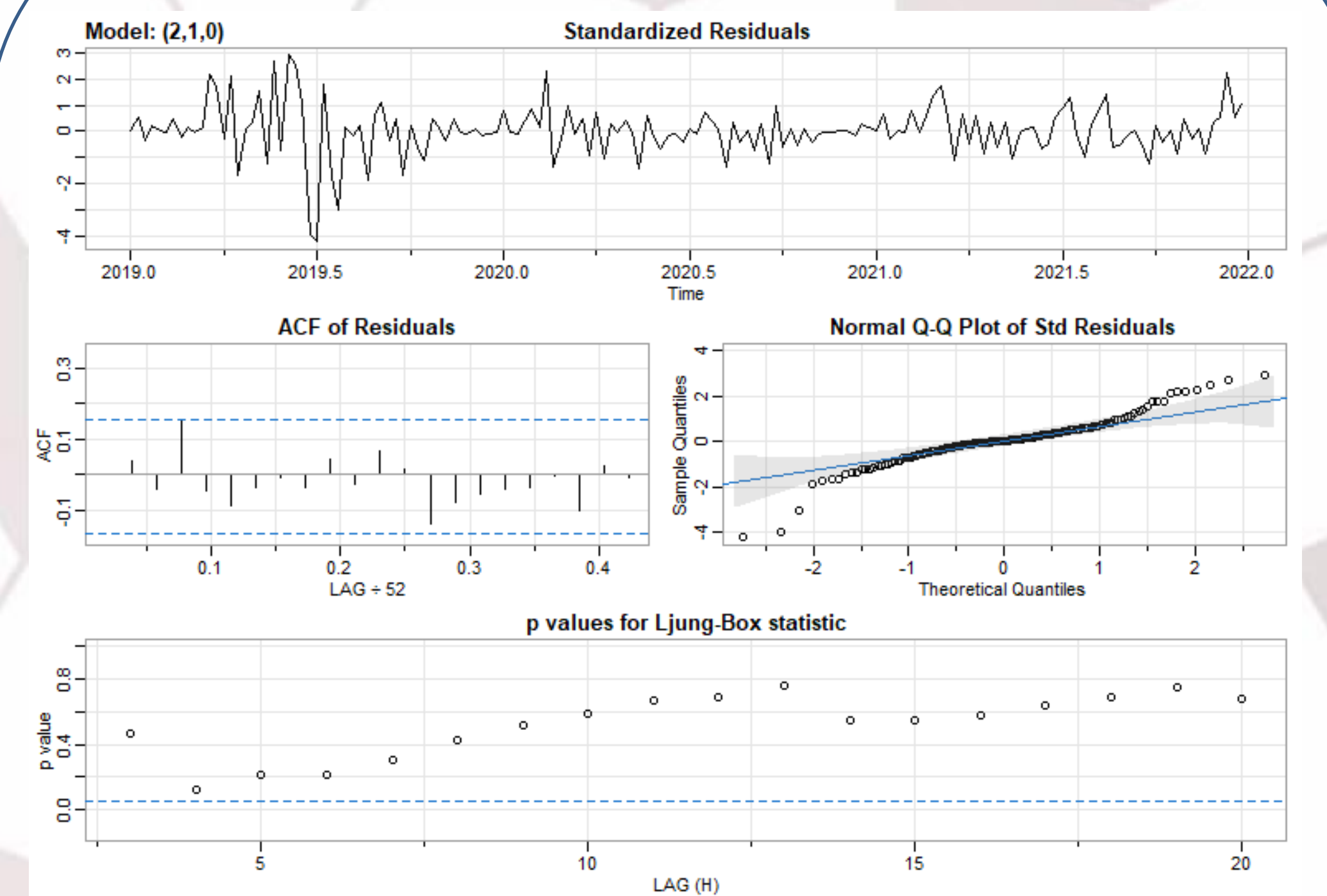


Fig 5. Análisis de residuos.

Predicciones	Valor mínimo	Valor máximo	Valor Enero	Error porcentual
$x_{157} = 61.2030$	47.0277	75.3782	60	2%
$x_{158} = 59.8840$	37.4419	82.32619	62	3.4%
$x_{159} = 59.8108$	33.3070	86.3146	47	27.2%
$x_{160} = 60.1693$	30.8943	89.4443	58	3.7%

Tabla II. Tabla de pronósticos de 4 pasos por delante.

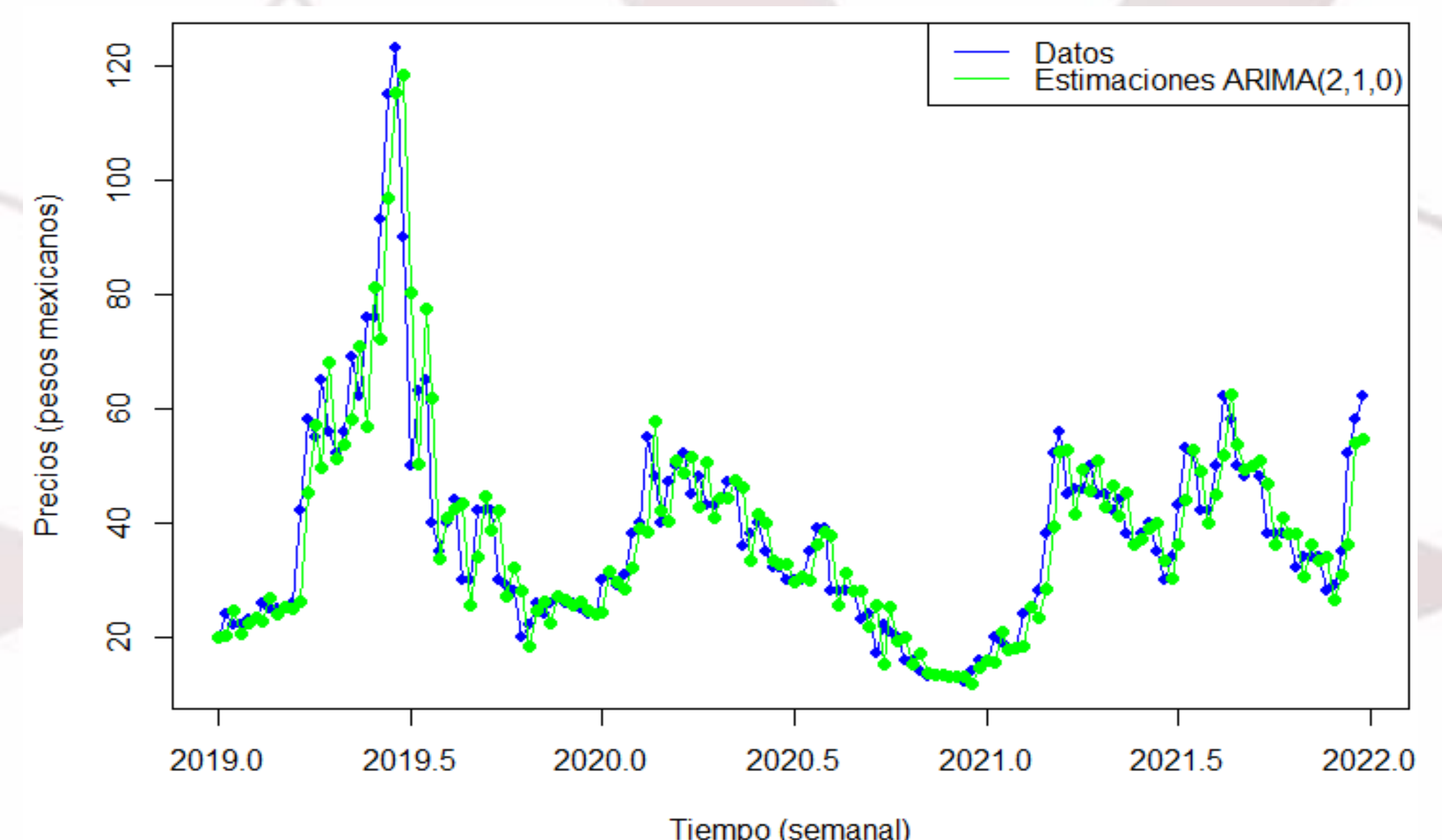


Fig 6. Ajuste del modelo ARIMA(2,1,0) a la serie de precios

RESULTADOS

Sabemos que la generación de pronósticos es satisfactoria si los valores futuros se encuentran dentro del intervalo de confianza [6]. Nuestros intervalos cumplen esta condición con un error menor al 5%, a excepción del tercer pronóstico.

DISCUSION

Pese a que aún están presentes algunas variaciones en la varianza, como se observó en la gráfica de los residuos estandarizados, se alcanzó a obtener la información necesaria de la serie para ajustar un modelo práctico, adecuado, que modela el comportamiento de la serie de manera aceptable. Lo que corresponde a el análisis numérico de residuos nos verifica esta aceptación.

Ahora, si consideramos el precio $x_{156} = 62$, y describimos el comportamiento de la serie en enero como: "baja, sube, baja, sube", el modelo falla en la segunda predicción. Para los intervalos de confianza (especialmente los valores máximos), el modelo brinda intervalos razonables. De las predicciones puntuales en la Fig 6 podemos rescatar que estos valores son bastante cercanos.

REFERENCIAS

- [1] <https://www.productoresdeaguacate.com/MODULArevista/modulos/web/www/precioseua.php>
- [2] Nielsen, A. (2019). *Practical Time Series Analysis: Prediction with Statistics and Machine Learning*. O'Reilly Media.
- [3] Copertwait, P. and Metcalfe, A. (2009). *Introductory Time Series with R*. NY, USA: Springer.
- [4] Pfaff, B. (2008). *Analysis of Integrated and Cointegrated Time Series with R* (2nd 2008 ed.). Springer.
- [5] Shumway, R. H., & Stoffer, D. S. (2017). *Time Series Analysis and Its Applications: With R Examples* (4th 2017 ed.). Springer.
- [6] Krispin, R. (2019). *Hands-On Time Series Analysis with R*. Packt Publishing.
- [7] <https://online.stat.psu.edu/stat510/lesson/5/5.1>
- [8] <https://www.rdocumentation.org/packages/stats/versions/3.6.2/topics/aggregate>
- [9] Hanke, J. (2006). *Pronósticos En Los Negocios* (8va Ed). Pearson Educación.
- [10] Pankratz, A. (1983). *Forecasting with Univariate Box - Jenkins Models*. Wiley.
- [11] Jaime, A. A. (1994). *Introducción Al Tratamiento De Series Temporales. Aplicación A Las Ciencias De Salud*. Ediciones Díaz De Santos.
- [12] Shumway, R., & Stoffer, D. (2019). *Time Series: A Data Analysis Approach Using R* (1.a ed.). CRC Press.
- [13] Bowerman, B. L., & O'Connell, R. T. (2007). *Pronósticos, series de tiempo y regresión: Un enfoque aplicado*. International Thomson Editores.
- [14] Cryer, J. D., & Chan, K. (2008). *Time Series Analysis: With Applications in R* (2nd ed.). Springer.