

**PRONÓSTICO DEL PRECIO DEL ARROZ PADDY MEDIANTE EL
MODELO ARIMA PARA COLOMBIA EN LOS AÑOS 1996 A 2018.**

PEDRO PABLO PEÑA GONZÁLEZ

OSCAR HERNÁN CASTILLO MOLINA



UNIVERSIDAD SANTIAGO DE CALI

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

MONOGRAFÍA

SANTIAGO DE CALI

2019

**PRONÓSTICO DEL PRECIO DEL ARROZ PADDY MEDIANTE EL
MODELO ARIMA PARA COLOMBIA EN LOS AÑOS 1996 A 2018.**

AUTORES:

PEDRO PABLO PEÑA GONZÁLEZ

OSCAR HERNÁN CASTILLO MOLINA

DIRECTOR TRABAJO DE GRADO

JOHYNER OBREGON MORALES

UNIVERSIDAD SANTIAGO DE CALI

FACULTAD DE CIENCIAS ECONÓMICAS Y EMPRESARIALES

MONOGRAFÍA

SANTIAGO DE CALI

2019

Resumen

El tema central del presente trabajo se basa en pronósticos del precio del cultivo de arroz paddy en Colombia para los años 1996 a 2020, la base de datos fue extraídos de Fedearroz. Mediante el modelo ARIMA prever el precio de arroz con el fin de que los cultivadores pueden reorganizar y tomar decisiones de sus cultivos.

Primero se caracterizó el cultivo dividido por departamentos y por altura, adicional por zonas que se distribuyen del siguiente modo: Bajo Cauca (Antioquia, Bolívar, Choco, Córdoba y Sucre), Centro (Caquetá, Cauca, Cundinamarca, Huila, Nariño, Tolima, Valle del Cauca), Costa Norte (Yondó – Antioquia, Atlántico, Cesar, La Guajira, Magdalena), Llanos (Arauca, Casanare, Paratebueno – Cundinamarca, Guaviare, Meta, Vichada) y Santanderes (Norte de Santander y Santander), segundo se describió el modelo y por último se estimó los precios para los últimos cuatro meses del año 2019 y todo el años 2020.

Dando como resultado del pronóstico del precio el cual descendió para de arroz paddy sugiriendo a los cultivadores disminuir los cultivos de arroz y así proceder a tomar decisiones con respecto al precio.

Palabras claves: Arroz paddy, pronostico, precio y cultivadores.

Abstract

The central theme of this paper is based on forecasts of the price of paddy rice cultivation in Colombia for the years 1996 to 2020, the database was extracted from Fedearroz. Through the ARIMA model, foresee the price of rice so that growers can reorganize and make decisions about their crops.

First, the crop divided by departments and height was characterized, additionally by areas that are distributed as follows: Bajo Cauca (Antioquia, Bolívar, Choco, Córdoba y Sucre), Centro (Caquetá, Cauca, Cundinamarca, Huila, Nariño, Tolima, Valle del Cauca), Costa Norte (Yondó – Antioquia, Atlántico, Cesar, La Guajira, Magdalena), Llanos (Arauca, Casanare, Paratebueno – Cundinamarca, Guaviare, Meta, Vichada) and Santanderes (Norte de Santander y Santander), second, the model was described and finally the prices were estimated for the last four months of the year 2019 and the whole year 2020.

As a result of the price forecast which fell for paddy rice, suggesting that growers decrease rice crops and thus proceed to make decisions regarding price.

Keywords: Paddy rice, forecast, price and growers.

Tabla de Contenidos

v

1.	Antecedentes	1
2.	Planteamiento del problema	3
3.	Objetivos	5
3.1.	General	5
3.2.	Especifico	5
4.	Justificación	6
5.	Marco de referencia	8
5.1.	Marco Contextual	8
5.1.2.	¿Qué es el arroz paddy y proceso?	8
5.1.3.	Cultivos de arroz	8
5.1.	Marco Conceptual	10
7.	Metodología	12
7.1.	Modelo	12
7.2.	Metodología de investigación	14
8.	Resultados	16
9.	Conclusiones y recomendación	21
10.	Bibliografía	23
	Anexos	25
A.	Mapa de las zonas y departamentos de siembra en Colombia	25
B.	Logaritmo del precio prueba de raíz unitaria primera diferencia	26
C.	Logaritmo del precio prueba de raíz unitaria en segunda diferencia	27

D. Correlograma del modelo	vi 28
E. Método ARMA a través de máxima verosimilitud	29
F. Raíz unitaria para los residuos	30
G. Correlograma para los residuos	31

Lista de tablas

vii

<u>Tabla 1. Departamentos arroceros divididos por zonas para el año 2016</u>	9
<u>Tabla 2. Logaritmo del precio prueba de raíz unitaria primera diferencia.</u>	26
<u>Tabla 3. Logaritmo del precio prueba de raíz unitaria en segunda diferencia.</u>	27
<u>Tabla 4. Método ARMA a través de máxima verosimilitud.</u>	29
<u>Tabla 5. Raíz unitaria para los residuos.</u>	30

Lista de gráficos

viii

<u>Gráfica 1. Se ilustra el precio promedio anual y el total nacional de producción de arroz paddy. Adaptado de datos de Fedearroz.</u>	4
<u>Grafica 2. Ilustra las hectáreas sembradas por zonas del país por años. Adaptado de datos de Fedearroz.</u>	11
<u>Grafica 3. Ilustra el resumen de los pasos del modelado ARIMA. Extraído de Gujarati y Porter (2010).</u>	15
<u>Grafica 4. Precio promedio mensual arroz paddy Verde en Colombia Peso/Tonelada. Adaptado de datos de Fedearroz.</u>	16
<u>Grafica 5. Ilustra el comportamiento del pronóstico desde agosto 2019.</u>	19
<u>Grafica 6. Ilustra los intervalos de confianza y pruebas estadísticas.</u>	20
<u>Grafica 7. Ilustra el pronóstico desde 2019 hasta 2020.</u>	20
<u>Grafica 8. Mapa de las zonas y departamentos de siembra en Colombia para el año 2016. DANE y Fedearroz (2016).</u>	25
<u>Grafica 9. Correlograma del modelo.</u>	28
<u>Grafica 10. Correlograma para los residuos.</u>	31

1. Antecedentes

En la búsqueda de este tipo trabajos nacionales se halló un solo trabajo en el cual se empleó la metodología ARIMA, esto no implica que posiblemente existan otras investigaciones con otros métodos de pronósticos. A continuación, se expondrá el documento hallado:

Castellanos y Páez (2016) a través del modelo ARIMA y metodología Box-Jenkins se pronosticó el precio promedio del arroz Paddy Verde (Enero 1996 - Septiembre de 2015), se establecieron dos series históricas mensuales de precio promedio de arroz Paddy Verde, la primera desde el año 1996 hasta el Abril de 2012, para determinar cómo habría sido la variación del IPC sin el TLC; y la segunda es la serie completa de enero de 1996 a septiembre de 2015, teniendo en cuenta el TLC. Teniendo en cuenta el TLC para los periodos comprendidos entre octubre y diciembre del 2015 tuvieron un promedio de \$7366 de diferencia a lo establecido por Fedearroz para ese mismo periodo, y con solo el IPC se puede identificar que los valores obtenidos en los tres futuros periodos a través del método Box-Jenkins tienen una desviación promedio de \$74.915 de diferencia con los valores dados por Fedearroz.

Ceballos y Pire (2015) utilizan la metodología Box-Jenkins a fin de aplicar el modelo econométrico ARIMA (1,0,1) con el fin de ajustar el comportamiento de la serie de tiempo de los precios internacionales del arroz durante el período junio 2002 a noviembre 2012. Dando como resultado que los precios de US \$648 por tonelada el primer mes de estimación y US \$665 por tonelada a los 16 meses; es decir, diciembre 2014, y arrojando valores en las bandas superior e inferior entre US \$191.7 y 2309.7 por tonelada.

Mientras Ohyver y Pudjihastuti (2018) uso la base del Servicio Provincial de Comercio extrajo los precios de la calidad media entre los periodos de enero 2015 y abril 2017, empleo modelo ARIMA para pronostica el valor del precio del arroz en Indonesia generando la ecuación final:

$$y_t = 0.8790y_{t-1} - 0.8638a_{t-1} + 0.0716a_{t-2}$$

El documento concluye que el modelo ARIMA es un método de pronostico a corto plazo y probablemente para los trabajos futuros, uno puede enfocarse en cómo obtener un método de pronóstico sólido.

Darekar y Reddy (2017) utiliza la base de datos de los precios del arroz paddy en India que es una fuente importante para este país. Dependiendo de los precios futuros durante la cosecha de kharif 2017-18 los agricultores toman decisiones para determinar el área a cultivar de paddy. Se utilizó los datos de series temporales sobre los precios promedio mensuales de arroz de enero de 2006 a diciembre de 2016 recopilados de AGMARK; empleando el modelo ARIMA (Box-Jenkins) se predijo los precios futuros del paddy y el rendimiento del modelo ajustado se examinó calculando varias medidas de bondad de ajuste *viz.*, AIC, BIC y MAPE. Concluyendo que el modelo ARIMA fue el modelo más representativo para el pronóstico del precio del arroz en la India en general y se registrarán en el rango de Rs. 1,600 - 2,200 por quintal en la temporada de cosecha de kharif, 2017-18.

Se examinó distintos documentos con fin de explorar e identificar las variables que se emplearon para la utilización del modelo ARIMA.

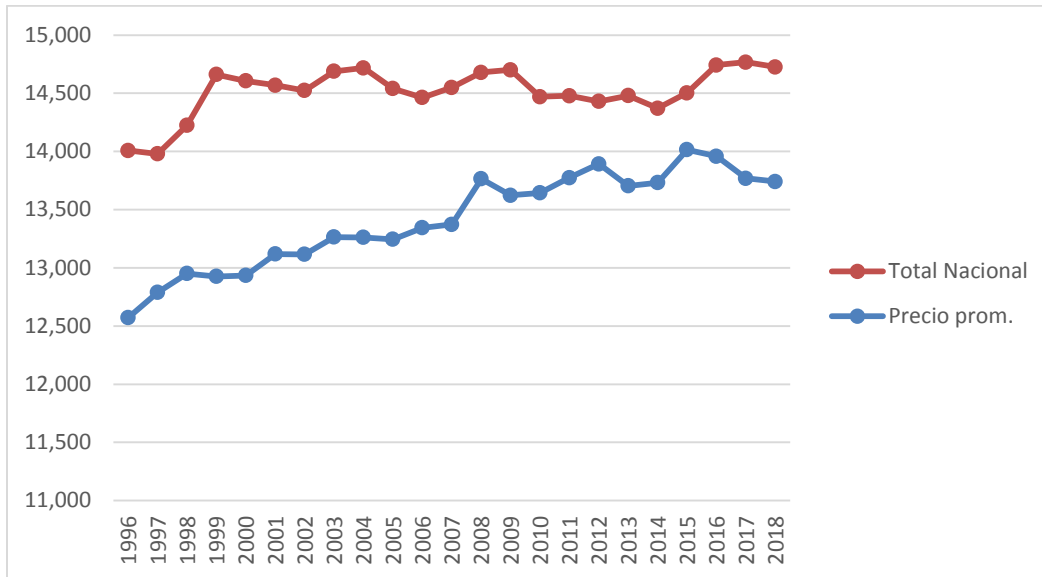
2. Planteamiento del problema

En Colombia el arroz es uno de los productos insignias, ya que es el tercer producto con mayor producción después del café y el maíz, también cuenta con fácil acceso a los precios que lo convierten en un alimento diario e indispensable en la canasta familiar y el de mayor consumo (Castellanos y Páez, 2016), que lo hacen uno de los cultivos con gran parte de área sembrada en Colombia.

Desde el punto de vista de la teoría económica sugiere que existen mecanismos los cuales pueden incidir en los niveles de precios de un bien homogéneo como el arroz (Fajardo, Avellaneda, Ramírez, Gómez y Becerra, 2013), dando como resultado que los precios sean volátiles y creando riesgos para todos los agentes económicos ya sean privados o públicos que dependan de este recurso que es el arroz.

En la gráfica 1 se observa en los años 1998, 2003, 2006, 2008 y 2015 el precio sube y la producción crece al mismo tiempo; aparentemente los cultivadores pudieron pronosticar la subida del precio para tener mayor rentabilidad. Por otro lado, en 1999, 2006, 2011, 2013 y 2016 al 2018 es cuando el precio baja y la producción aumenta generando pérdidas de beneficios para los productores; y los años que no se nombraron hubo una congruencia en la disminución el precio y la caída de la producción.

Dado lo anterior el arroz es un producto importante para los cultivadores, ya que con este cultivo generan ingresos para sus familias y no tener pérdidas de beneficios; pero los factores exógenos o endógenos que pueden producir fluctuaciones en el precio e incide en las ganancias de estas personas generando riesgos para todos los agentes económicos, en especial para el productor. Por lo tanto, el problema de investigación se centrará en el



Gráfica 1. Se ilustra el precio promedio anual y el total nacional de producción de arroz paddy. Adaptado de datos de Fedearroz.

precio del arroz paddy formulando la pregunta de investigación ¿Cómo es la evolución el precio del arroz paddy en un futuro para disminuir la incertidumbre en la toma de decisiones de la cantidad a cultivar de arroz?

3. Objetivos

General

Pronosticar el precio del arroz paddy mediante un modelo ARIMA para Colombia en el periodo 1996 a 2020.

Específico

1. Revisar el comportamiento de la serie de tiempo de la producción paddy para 2018.
2. Estimar el precio generador de la serie de tiempo de la producción de arroz paddy.
3. Sugerir la política con el fin para no perder producción ni beneficios.

4. Justificación

La obtención de arroz paddy verde es la base para la producción del arroz blanco final, pero este bien (arroz paddy) tiene muchos factores el cual hacen cambiar su precio obteniendo consecuencias positivas o negativas en especial para el productor. Estos factores que afectan el precio son exógenos o endógenos como por ejemplo los costos de producción, el clima, el TLC entre otros; que afectan las fluctuaciones del precio del cultivo de arroz paddy en Colombia. Estas alteraciones en el precio producen que los cultivadores deban ejecutar planes en los cuales la toma de decisiones se vuelva más complejas para el agente privado y/o público.

De lo anterior se ha tenido la necesidad de encontrar metodologías de prevenir situaciones a futuro para sobrevivir como agente económico. Por esta razón se han elaborado funciones más eficientes para previsión, planificación, coordinación, ejecución y control de las operaciones a fin de lograr los objetivos deseados de la empresa (Nava, 2005).

En un mercado que es cada vez más competitivo, con recursos limitados y con precios variables crean riesgos para los productores, comerciantes y consumidores involucrados en la producción, comercialización y consumo de arroz, por lo tanto es importante para la empresa (privada o pública) que disponga de metodologías o técnicas estadísticas, con las cuales permitan pronosticar acontecimientos futuros, con información proporcionada del pasado, en este caso suministrando para los cultivadores de arroz información para la toma de decisiones y efectuar los cambios pertinentes en los

planes del cultivo anticipando el precio futuro del mercado del arroz paddy, y así disminuyendo los riesgos de pérdidas de cultivos o de beneficios.

5. Marco de referencia

Marco Contextual

5.1.2. ¿Qué es el arroz paddy y proceso?

En la agricultura se emplean diferentes nombres para los diferentes procesos de los cultivos, para el caso del cultivo del arroz se explica el proceso muy brevemente a continuación:

El arroz paddy es el grano del cereal después de su recolección. Aún dispone de su cáscara natural y de todas sus propiedades nutritivas. El arroz paddy se recibe del campo y se limpia para separar posibles impurezas como pueden ser ramas, polvo o pequeñas piedras. Una vez limpio se utiliza tanto para alimentación humana como para alimentación animal. En el primer caso, el arroz paddy es molido y blanqueado para obtener el arroz blanco de consumo humano y los demás subproductos para alimentación animal. En el segundo caso, sin realizarle ningún tratamiento físico o químico, es añadido en las mezclas para pájaros, palomas y otros pequeños animales de compañía. Es una fuente de energía muy valorada, ya que debido a que aún mantiene su piel el animal puede absorber la energía del grano de una forma lenta y controlada (Indutrias Ralda).

5.1.3. Cultivos de arroz

Según cifras del DANE el sector del arroz es el tercer producto con más importancia en agricultura colombiana, además es importante para la dieta y en canasta familiar de los hogares colombianos, y cabe señalar la importancia en la canasta de consumo de los hogares más pobres (Fajardo et al., 2013).

A continuación, en Anexos A: Mapa de las zonas y departamentos de siembra en Colombia. Se ilustra las diferentes zonas divididas por departamentos y por altura donde se ubican los cultivos de arroz. En donde se resalta que la gran parte de los departamentos tienen cultivos de arroz.

Además, en la tabla 1 se observa las hectáreas de área sembrada de arroz por zonas para el 2018 dando así un total de 500.924 hectáreas sembradas en Colombia que es aproximadamente el 0,44% del territorio colombiano.

Tabla 1

Departamentos arroceros divididos por zonas para el año 2018

ZONA ARROCERA	DEPARTAMENTO	AREA SEMBRADA AÑO	
		Ha.	%
BAJO CAUCA	ANTIOQUIA	68.190	13,61%
	BOLIVAR		
	CHOCO		
	CORDOBA		
	SUCRE		
CENTRO	CAQUETA	148.214	29,59%
	CAUCA		
	CUNDINAMARCA		
	HUILA		
	NARIÑO		
	TOLIMA		
	VALLE DEL CAUCA		
COSTA NORTE	YONDÓ (ANTIOQUIA)	22.664	4,52%
	ATLANTICO		
	CESAR		
	LA GUIJIRA		
	MAGADELANA		
LLANOS	ARAUCA	222.687	44,46%

	CASANARE			
	PARATEBUENO			
	(CUND.)			
	GUAVIARE			
	META			
	VICHADA			
	NORTE	DE	39.169	7,82%
SANTANDERES	SANTAND			
	SANTANDER			

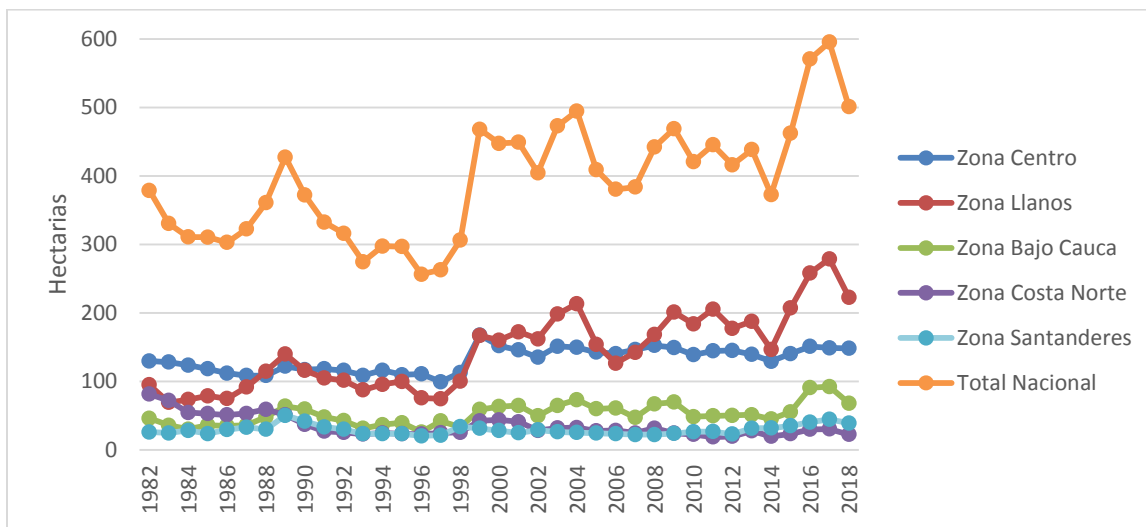
La tabla 1 muestra las hectáreas sembradas de arroz con sus respectivos porcentajes. Adaptado de datos de Fedearroz.

Para terminar en la gráfica 2 se ilustra las hectáreas sembradas por zonas del país a través de los años, en la cual se observa la zona de Llanos aporta más áreas cultivadas de arroz del país desde 1998, aunque su tendencia es creciente a través del tiempo, en el último año el 2018 hay una caída en estas áreas; igualmente la zona Centro comparten una tendencia creciente pero menos pronunciada, siendo estas dos zonas las que más aportan en cultivo de este cereal. En el caso de Las zonas Costa y Santander han tenido un comportamiento muy similar en sus tendencias y hasta áreas cultivadas y por último la zona Cauca desde 1998 su tendencia es mayor que las dos zonas anteriores.

5.1.Marco Conceptual

La palabra pronostico tiene varios significados, pero para efectos de este trabajo se tomaran conceptos de par de expertos sobre el tema; el primero que es Montemayor (2013) dice “Pronosticar es realizar un enunciado sobre el valor futuro de una variable de interés,

fundamentado ya sea por el análisis de datos históricos disponibles, por el juicio de expertos en el tema o por una combinación de ambas cosas” (p. 1).



Grafica 2. Ilustra las hectáreas sembradas por zonas del país por años. Adaptado de datos de Fedearroz.

Por ultimo Pindyck y Rubinfeld (2000) manifiesta que:

Un pronóstico es un estimado cuantitativo (o conjunto de estimados) acerca de la verosimilitud de eventos futuros, que se elabora con base en información pasada y actual. Esta información está expresada en la forma de un modelo: un modelo estructural de una sola ecuación y un modelo de ecuación múltiple o un modelo de series de tiempo. Al extrapolar nuestros modelos más allá del periodo durante el cual fueron estimados, podemos hacer pronósticos sobre eventos futuros (p. 211).

7. Metodología

Se realizó un estudio exploratorio y descriptivo. En donde se utilizó la metodología de serie de tiempo¹ denominada ARIMA para pronosticar los precios de los cultivos paddy en Colombia. Para ellos se implementó la información contenida en la base de datos de Fedearroz de años 1996 hasta 2019. Estos modelos o métodos analizan las propiedades probabilísticas o estocásticas de las series de tiempo económicas, en donde Y_t se explica por valores pasados o rezagados de sí misma y por el término del error estocástico (Gujarati y Porter, 2010).

Modelo

Ohyver y Pudjihastuti (2018) dice que el modelo ARIMA está dado por:

$$\phi_p(B)(1 - B)^d z_t = \theta_q(B)a_t \quad (1)$$

Donde ϕ_p y θ_q son parámetros desconocidos; z_t son los datos de la serie de tiempo; B es operador de cambio hacia atrás; d es la orden de diferenciación y a_t es la serie de ruido blanco.

Hay dos tipos de modelo ARIMA que se denominan estacionario ARIMA y no estacionario ARIMA. El estacionario es una serie temporal estacionaria se relaciona con características estadísticas en tiempo o distribuciones de probabilidad en un tiempo constante, y se define de la siguiente forma:

- El valor esperado de la serie temporal es independiente del tiempo.

¹ En los modelos de series de tiempo hacen el supuesto no conocer la causalidad por la cual afecta la variable que se va pronosticar, dando lugar al comportamiento pasado de una serie de tiempo con el fin de inferir su comportamiento en el futuro y son confiables en el corto plazo (Pindyck & Rubinfeld, 2000).

- La función de autocovarianza es una función de k , donde cada k , $\gamma_z = Cov(z_t, z_{t+1})$

En la ecuación (1) tiene componentes: Autoregresivos (AR) modelo p y Medias Móviles (MA) modelo q . El ARIMA puede ser un AR modelo p y MA modelo q la combinación de estos dos componentes produce el ARMA:

1. $AR_{(p)}$:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t$$

a_t es un error que se supone que no están correlacionados entre sí, donde $E\{a_t\} = 0$ y $Var\{a_t\} = \sigma_a^2$ y el coeficiente ϕ_p $i = 1, 2, \dots, p$ son los parámetros a estimar.

2. El modelo MA contiene "promedios" de los ruidos del período actual y del período anterior, el $MA_{(q)}$ se define así:

$$z_t = a_t - \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots - \phi_q a_{t-q}$$

El coeficiente ϕ_q $q = 1, 2, \dots, p$ es el parámetro a determinar.

Entonces el modelo $ARMA_{(p,q)}$ contiene AR y MA , este modelo no posee los elementos i porque es un modelo estacionario. En otras palabras, el elemento d en (1) es igual a cero. Por lo tanto el modelo $ARMA_{(p,q)}$ es:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t a_t - \phi_1 a_{t-1} - \phi_2 a_{t-2} - \dots - \phi_q a_{t-q}$$

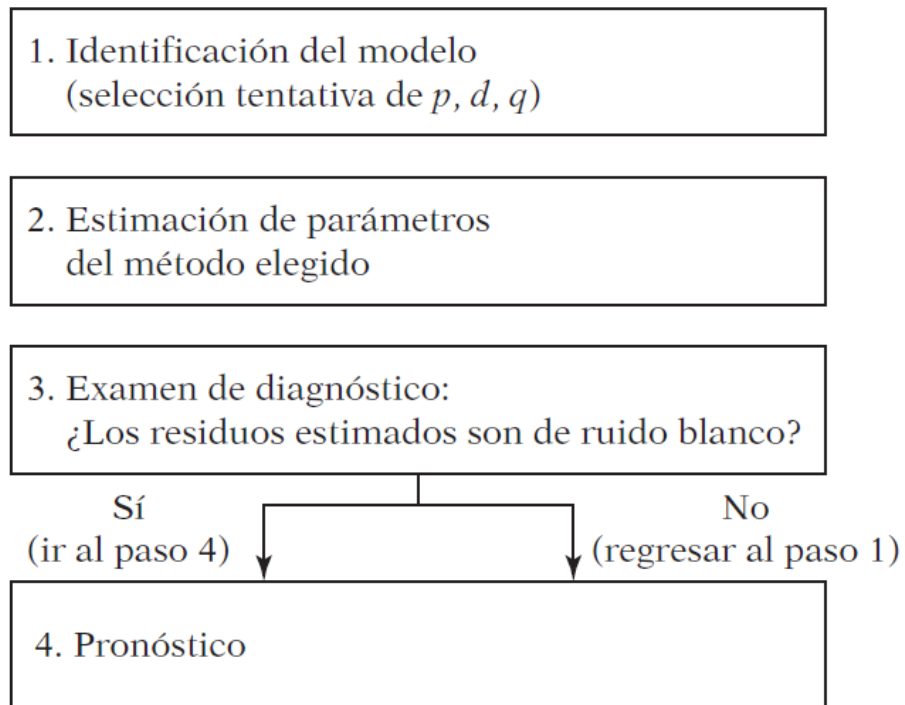
Metodología de investigación

Se evaluó el pronóstico para la base de datos del precio promedio mensual arroz paddy verde en Colombia extraída de Fedearroz para los años 1996 hasta 2019, empleando el método autorregresivo ARIMA.

Según Gujarati y Porter (2010) considera cuatro pasos que son los siguientes:

1. **Identificación:** hallar los valores apropiados de p , d y q , para esta labor el correlograma y el correlograma parcial son de gran ayuda.
2. **Estimación:** Después de identificar los valores de p , d y q , el siguiente paso es estimar los parámetros de los términos autorregresivos y de promedios móviles incluidos en el modelo.
3. **Examen de diagnóstico:** después de seleccionar en particular un modelo ARIMA y estimar sus parámetros, se ve si el modelo elegido se ajusta a los datos en buena forma o poder ser posible que exista otro modelo ARIMA que lo haga. Con una prueba del modelo sobre los residuales estimados con el fin de observar si son ruido blanco, si lo son se acepta el ajuste; en caso contrario se debe iniciar nuevamente.

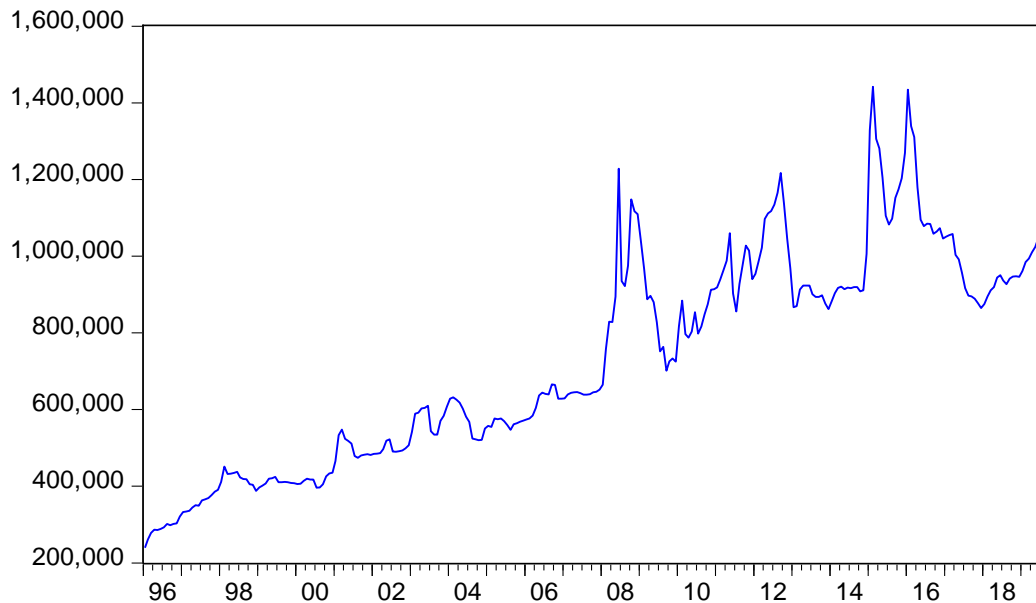
En el gráfico 3 se resume lo anterior



Gráfica 3. Ilustra el resumen de los pasos del modelado ARIMA. Extraído de Gujarati y Porter (2010)

8. Resultados

En el gráfico 4, que muestra la evolución de los precios a través de los años, se observa que la serie analizada aparentemente es no estacionaria, la muestra tiene una tendencia creciente a lo largo de los periodos mensuales de cada año; al igual que periodos con marcada estacionalidad, con lo cual, es evidente que los principales momentos de su distribución (media-varianza), van cambiando a través del tiempo.



Gráfica 4. Precio promedio mensual arroz paddy Verde en Colombia Peso/Tonelada. Adaptado de datos de Fedearroz.

Por medio de una prueba formal, se verificará la ausencia/presencia de estacionariedad de la serie, se aplicó el modelo ARIMA. Previamente se aplicará una transformación logarítmica a los precios de la serie original, con el fin de atenuar la volatilidad de la misma.

En Anexos B: Logaritmo del precio prueba de raíz unitaria primera diferencia. Se muestran los resultados del Test de Dickey-Fuller aumentado², el cual se evidencio en contra de la hipótesis (presencia de raíz unitaria en la serie), para niveles de significancia de 1% y 5% de la serie en nivel, indicando con ello que la serie en mención es estacionaria. Contrastando este resultado con la prueba KPSS³ que es más robusta, con el cual se verifico una vez más que la serie en nivel presenta una raíz unitaria (P valor 0.2617).

A continuación, se procede a aplicar una primera diferencia al logaritmo natural de la serie, con el fin de verificar que con dicha transformación se consiga hacer estacionaria la serie con el fin de verificar la posibilidad de hacerla estacionaria y proponer un modelo de estructura estacional aprovechando la información contenida en la base de datos la cual es el mensual. En Anexos C: Logaritmo del precio prueba de raíz unitaria en segunda diferencia, se observa el valor p de la prueba (0.0000) de la prueba Dickey-Fuller Aumentado, que la diferencia en logaritmo de la serie propuesta no presenta problemas de raíz unitaria.

Se aplicó una diferencia de tipo estacional a la serie y verificado que la serie es estacionaria, se optó por una diferencia estacional de orden 12 (por la periodicidad mensual de los datos), con el propósito de identificar mejor los componentes AR y MA. En Anexo D: Correlograma del modelo, se observa que la diferenciación estacional muestra un patrón de carácter estacional y regular en la misma, toda vez que los rezagos múltiples del periodo estacional (12, 24, 36), decaen muy lentamente en la función de

² En ingles se denomina Augmented Dickey-Fuller test statistic

³ Test de Kwiatkowski-Phillips-Schmidt-Shin

autocorrelación parcial; mientras que en la Función de Autocorrelación a partir del retardo 12, la función cae bruscamente a cero, sugiriendo con esto un esquema AR (12), en la parte estacional de la misma. En cuanto a la parte regular de la misma, se aprecia que resultan ser estadísticamente significativos los primeros rezagos de las dos funciones, sugiriendo esquemas AR (1) y MA (1) en la parte regular de la misma.

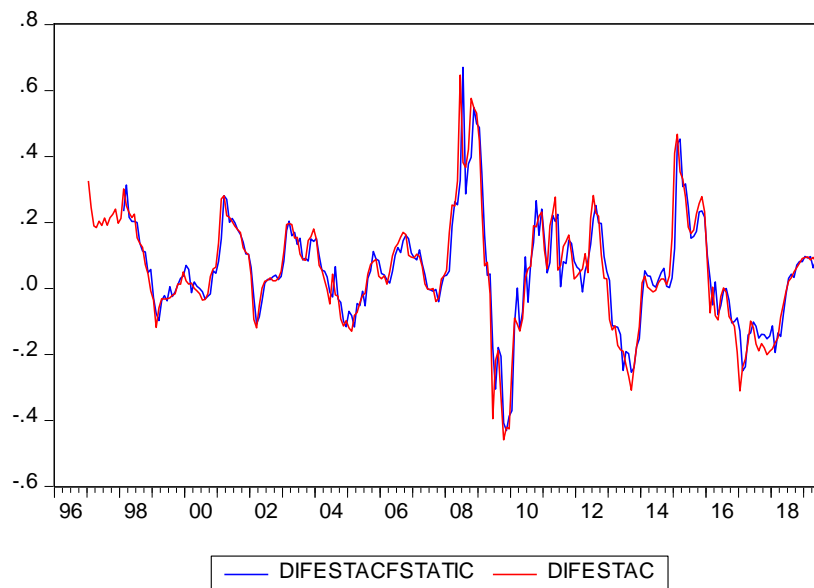
Se procedió con la estimación del modelo previamente identificado en Anexos E: Método ARMA a través de máxima verosimilitud, se observa que los componentes previamente identificados resultan ser altamente significativos (valores P menores a 1%), y el intercepto del modelo alcanza un nivel de significancia de 10%; también se destaca el elevado valor que alcanza el coeficiente de determinación⁴ (88.04%).

Una vez se identificó y se estimó el modelo apropiado para explicar el comportamiento de la serie. A continuación, en Anexos F: Raíz unitaria para los residuos se verificó que los residuos del modelo propuesto sean ruido Blanco; mediante la prueba de Dickey Fuller Aumentada no tienen raíz unitaria. Y en Anexos G: Correlograma para los residuos muestra el correlograma de dichos residuos para el modelo propuesto, se verifica que los residuos del modelo son ruido blanco y no sugieren que haya quedado procesos por modelar, ni en la parte regular ni en la parte estacional de la misma. Dando como sugerencia que el modelo ajustado y validado para la serie del precio del arroz sea un ARIMA (1,0,1) (1,1,0)₁₂.

A partir de la función estimada se pronosticará para los últimos 4 meses del año 2019 y todo el año 2020. En la gráfica 5 se observa la serie ajustada (línea color azul), es muy idéntica a la línea que representa las observaciones reales de la serie (línea color

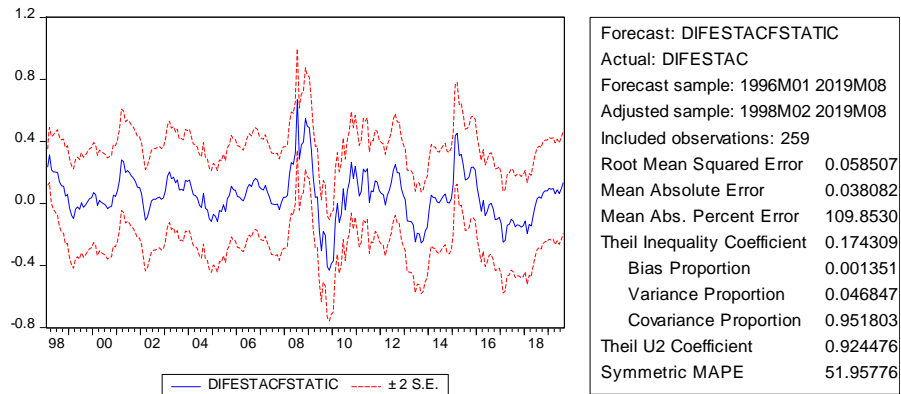
⁴ R-squared en ingles

rojo); lo anterior evidencia el buen ajuste obtenido con el modelo que, como se mencionó anteriormente, que presento un elevado valor en su coeficiente de determinación (88.04%)



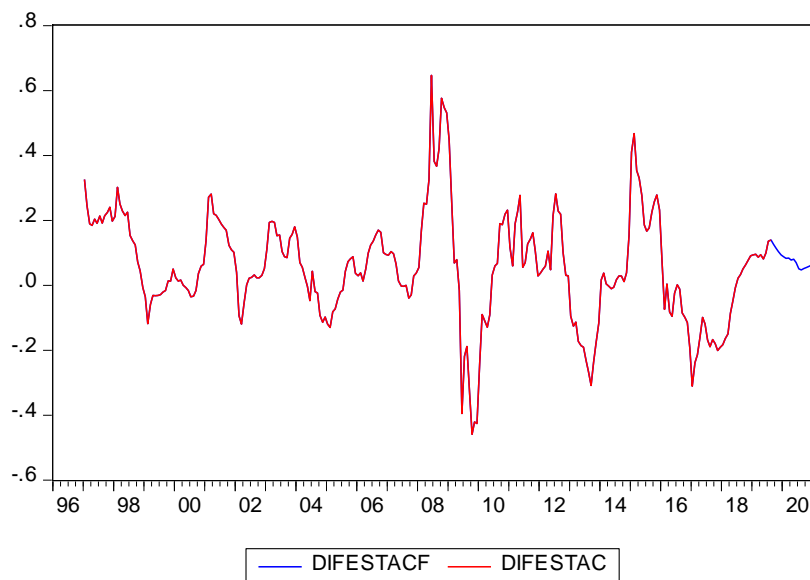
Gráfica 5. Ilustra el comportamiento del pronóstico desde agosto 2019.

Para mayor formalidad en la gráfica 6 se presenta la serie ajustada/pronosticada con un intervalo de confianza de ± 2 errores estándares. Con el *coeficiente de desigualdad de Theil*, el cual tiene un bajo valor, confirmando, una vez más, el buen ajuste alcanzado por el modelo propuesto.



Gráfica 6. Ilustra los intervalos de confianza y pruebas estadísticas.

Por último, en la gráfica 7 se observa que la serie pronosticada para los 16 periodos posteriores al mes de agosto 2019 representada por la línea color azul. Como se puede apreciar sugiere una tendencia decreciente para el precio del arroz paddy para el último trimestre del año 2019 y la totalidad del año 2020.



Gráfica 7. Ilustra el pronóstico desde 2019 hasta 2020.

9. Conclusiones y recomendación

Los datos de series temporales se encuentran en muchas áreas, como finanzas, economía, industria, seguros y agricultura, aunque existente diferentes métodos de pronósticos los cuales son aplicados con respecto al patrón de datos a analizar y hay algunos estudios que aplican algunos métodos estadísticos para pronosticar (Ohyver y Pudjihastuti, 2018).

Pero este documento pronosticó los precios promedios mensuales del arroz paddy para los años 1996 hasta 2020 mediante la serie de tiempo de los precios históricos. Hay que tener en cuenta que este tipo de técnicas no garantizan pronósticos perfectos, ya que no tienen en cuenta factores exógenos y endógenos. Sin embargo, son modelos útiles para dar una idea de prevención a futuro, en este caso para la planeación y toma de decisiones de los cultivadores de arroz.

Esta previsión se basa en datos del precio de mercado otorgados por Fedearroz y puede no ser el mismo que el previsto. En este orden de idea el modelo ARIMA es un método de extrapolación que requiere solo los datos históricos de series temporales de la variable en estudio.

Finalmente, los precios pronosticados por medio del modelo ARIMA revelaron una disminución en los precios del arroz para los meses futuros. Por lo tanto, una sugerencia latente es disminuir las hectáreas de producción de arroz con el fin de no tengan pérdidas de cultivos ni de beneficios, y también que prevean la cantidad de venta de cultivos en los mercados adecuados, puedan planificar y toma de decisiones “adecuadamente” dentro de la empresa. Por otro lado, también al conocer el valor del precio del arroz en los próximos períodos también podrá ayudar al gobierno colombiano

a tomar las decisiones en políticas necesarias anticipando las fluctuaciones del precio del arroz paddy generando un mejor bienestar en el ámbito agrario.

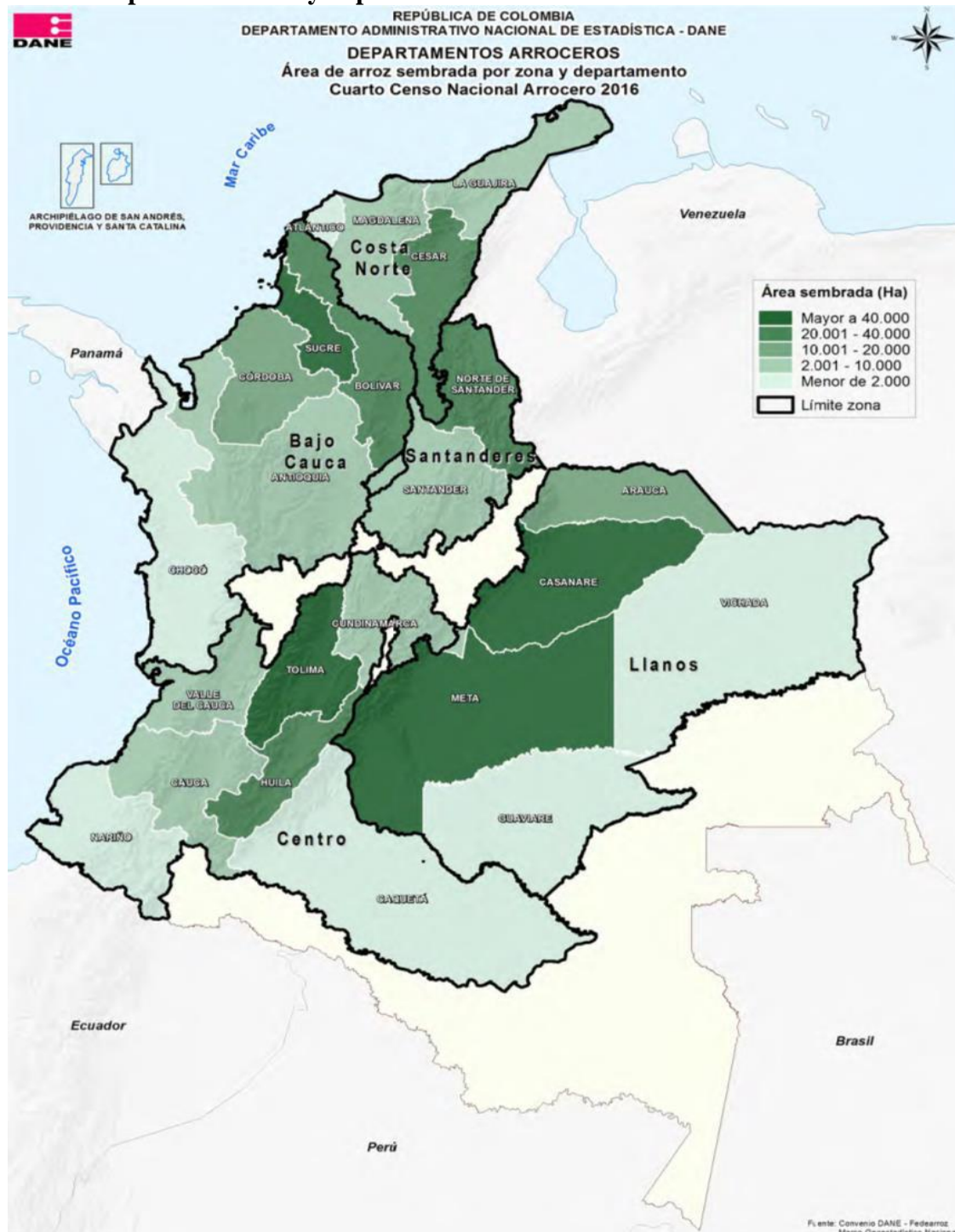
10. Bibliografía

- Castellanos, I., & Páez, I. (2016). *Predicción del comportamiento del precio promedio mensual del arroz paddy verde en Colombia teniendo en cuenta la información entre 1996 - 2015 y el TLC* (Especialización en estadística aplicada). Fundación Universitaria los Libertadores, Bogotá D.C., Colombia.
- Ceballos, S. G., & Pire, R. (2015). Estimación del precio internacional del arroz (*Oryza sativa* L.) bajo el modelo ARIMA. *Revista Mexicana de Ciencias Agrícolas*, (11), 2083–2089. <https://doi.org/10.29312/remexca.v0i11.776>
- DANE, & Fedearroz. (2016). *4° Censo Nacional Arrocerero Año 2016. Boletín técnico*. Recuperado de <https://www.dane.gov.co/files/investigaciones/agropecuaria/censo-nacional-arrocero/presentacion-4to-censo-nacional-arrocero-2016.pdf>
- Darekar, A., & Reddy, A. A. (2017). *Forecasting of common paddy prices in India*. *10*(1), 71–75.
- Fajardo, J., Avellaneda, S. M., Ramírez, J. M., Gómez, D., & Becerra, A. (2013). Política comercial para el arroz. *Fedesarrollo*, 1–72.
- Gujarati, D., & Porter, D. (2010). *Econometría* (5a ed.). Mexico D.F., Mexico: Mc. Graw Hill.
- Indutrias Ralda. (s/f). Arroz paddy o arroz con cáscara. Recuperado el 27 de agosto de 2019, de <https://www.iralda.com/productos/arroz-paddy-arroz-cascara/>
- Montemayor, E. (2013). *Métodos de pronósticos para negocios* (1a ed.). Monterrey, Mexico: Editorial digital.
- Nava, G. (2005). Modelo de pronósticos para el área comercial de una empresa distribuidora de cerveza (Tesis de maestría). Universidad de las Américas Puebla.

- Ohyver, M., & Pudjihastuti, H. (2018). Arima model for forecasting the price of medium quality rice to anticipate price fluctuations. *Procedia Computer Science*, 135, 707–711. <https://doi.org/10.1016/j.procs.2018.08.215>
- Pindyck, R., & Rubinfeld, D. (2000). *Econometría modelos y pronosticos* (4a ed.). Santiago Acahualtepec, Mexico: McGRAW-HILL.

Anexos

A. Mapa de las zonas y departamentos de siembra en Colombia



Gráfica 8. Mapa de las zonas y departamentos de siembra en Colombia para el año 2016. DANE y Fedearroz (2016).

B. Logaritmo del precio prueba de raíz unitaria primera diferencia

Tabla 2

Logaritmo del precio prueba de raíz unitaria primera diferencia.

Null Hypothesis: LNPRECIO has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 1 (Automatic - based on SIC, maxlag=15)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-3.503761	0.0409
Test critical values:		
1% level	-3.992411	
5% level	-3.426557	
10% level	-3.136516	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(LNPRECIO)
 Method: Least Squares
 Date: 09/06/19 Time: 13:06
 Sample: 1997M03 2019M08
 Included observations: 270

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNPRECIO(-1)	-0.075334	0.021501	-3.503761	0.0005
D(LNPRECIO(-1))	0.188482	0.060240	3.128832	0.0020
C	0.976400	0.276805	3.527386	0.0005
@TREND("1997M03")	0.000307	0.000102	3.007790	0.0029
R-squared	0.067399	Mean dependent var		0.004298
Adjusted R-squared	0.056881	S.D. dependent var		0.051484
S.E. of regression	0.049999	Akaike info criterion		-3.138933
Sum squared resid	0.664967	Schwarz criterion		-3.085623
Log likelihood	427.7559	Hannan-Quinn criter.		-3.117526
F-statistic	6.407913	Durbin-Watson stat		1.977935
Prob(F-statistic)	0.000331			

La tabla 2 muestra las diferentes pruebas estadísticas

C. Logaritmo del precio prueba de raíz unitaria en segunda diferencia

Tabla 3

Logaritmo del precio prueba de raíz unitaria en segunda diferencia.

Null Hypothesis: D(LNPRECIO) has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=15)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-14.33564	0.0000
Test critical values:		
1% level	-3.990935	
5% level	-3.425841	
10% level	-3.136094	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(LNPRECIO,2)
 Method: Least Squares
 Date: 09/06/19 Time: 11:05
 Sample (adjusted): 1996M03 2019M08
 Included observations: 282 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(LNPRECIO(-1))	-0.844491	0.058908	-14.33564	0.0000
C	0.008611	0.006057	1.421717	0.1562
@TREND("1996M01")	-3.12E-05	3.68E-05	-0.849945	0.3961
R-squared	0.424185	Mean dependent var		-0.000325
Adjusted R-squared	0.420057	S.D. dependent var		0.065839
S.E. of regression	0.050139	Akaike info criterion		-3.137454
Sum squared resid	0.701383	Schwarz criterion		-3.098711
Log likelihood	445.3811	Hannan-Quinn criter.		-3.121918
F-statistic	102.7653	Durbin-Watson stat		1.975167
Prob(F-statistic)	0.000000			

La tabla 3 muestra las diferentes pruebas estadísticas

D. Correlograma del modelo

Date: 09/06/19 Time: 13:43

Sample: 1997M03 2019M08

Included observations: 269

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	0.143	0.143	5.5513	0.018
		2	-0.072	-0.094	6.9565	0.031
		3	0.078	0.105	8.6060	0.035
		4	0.071	0.037	10.011	0.040
		5	-0.073	-0.078	11.481	0.043
		6	0.032	0.062	11.772	0.067
		7	0.038	-0.001	12.170	0.095
		8	-0.073	-0.066	13.657	0.091
		9	-0.071	-0.044	15.069	0.089
		10	0.069	0.064	16.409	0.089
		11	0.018	0.001	16.497	0.124
		12	-0.473	-0.475	80.030	0.000
		13	-0.105	0.054	83.153	0.000
		14	0.064	0.002	84.328	0.000
		15	-0.108	-0.078	87.697	0.000
		16	-0.095	-0.015	90.277	0.000
		17	-0.075	-0.197	91.896	0.000
		18	-0.082	0.012	93.838	0.000
		19	0.014	0.071	93.895	0.000
		20	0.100	0.030	96.842	0.000
		21	0.045	-0.007	97.442	0.000
		22	-0.018	0.007	97.541	0.000
		23	-0.082	-0.055	99.552	0.000
		24	0.054	-0.207	100.42	0.000
		25	0.009	-0.021	100.45	0.000
		26	-0.066	-0.038	101.76	0.000
		27	0.019	-0.050	101.87	0.000
		28	0.006	-0.045	101.88	0.000
		29	0.077	-0.034	103.70	0.000
		30	0.076	-0.007	105.45	0.000
		31	0.053	0.099	106.31	0.000
		32	-0.079	-0.085	108.22	0.000
		33	-0.028	-0.035	108.46	0.000
		34	-0.021	-0.033	108.59	0.000
		35	0.032	-0.034	108.91	0.000
		36	-0.123	-0.247	113.62	0.000

Gráfica 9. Correlograma del modelo.

E. Método ARMA a través de máxima verosimilitud

Tabla 4

Método ARMA a través de máxima verosimilitud.

Dependent Variable: DIFEST
 Method: ARMA Maximum Likelihood (OPG - BHHH)
 Date: 09/06/19 Time: 13:52
 Sample: 1997M03 2019M08
 Included observations: 270
 Convergence achieved after 44 iterations
 Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.057386	0.033314	1.722595	0.0861
AR(1)	0.901782	0.025788	34.96884	0.0000
SAR(12)	-0.459237	0.033676	-13.63671	0.0000
MA(1)	0.230349	0.044522	5.173867	0.0000
SIGMASQ	0.003327	0.000132	25.29230	0.0000
R-squared	0.880426	Mean dependent var		0.053029
Adjusted R-squared	0.878621	S.D. dependent var		0.167126
S.E. of regression	0.058226	Akaike info criterion		-2.813216
Sum squared resid	0.898413	Schwarz criterion		-2.746578
Log likelihood	384.7841	Hannan-Quinn criter.		-2.786457
F-statistic	487.8019	Durbin-Watson stat		1.985062
Prob(F-statistic)	0.000000			
Inverted AR Roots	.91+.24i	.91-.24i	.90	.66-.66i
	.66+.66i	.24+.91i	.24-.91i	-.24+.91i
	-.24-.91i	-.66-.66i	-.66-.66i	-.91-.24i
	-.91+.24i			
Inverted MA Roots	-.23			

La tabla 4 muestra las diferentes pruebas estadísticas

F. Raíz unitaria para los residuos

Tabla 5

Raíz unitaria para los residuos.

Null Hypothesis: RESIDUOS has a unit root
 Exogenous: Constant, Linear Trend
 Lag Length: 0 (Automatic - based on SIC, maxlag=15)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-16.51998	0.0000
Test critical values:		
1% level	-3.992411	
5% level	-3.426557	
10% level	-3.136516	

*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(RESIDUOS)
 Method: Least Squares
 Date: 09/06/19 Time: 14:06
 Sample: 1997M03 2019M08
 Included observations: 270

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESIDUOS(-1)	-1.010698	0.061180	-16.51998	0.0000
C	0.010889	0.005788	1.881232	0.0610
@TREND("1997M03")	-6.52E-05	3.72E-05	-1.752694	0.0808

R-squared	0.505473	Mean dependent var	-0.000222
Adjusted R-squared	0.501769	S.D. dependent var	0.067098
S.E. of regression	0.047361	Akaike info criterion	-3.250975
Sum squared resid	0.598905	Schwarz criterion	-3.210993
Log likelihood	441.8817	Hannan-Quinn criter.	-3.234920
F-statistic	136.4550	Durbin-Watson stat	2.001062
Prob(F-statistic)	0.000000		

La tabla 5 muestra las diferentes pruebas estadísticas

G. Correlograma para los residuos

Date: 09/06/19 Time: 14:06

Sample: 1997M03 2019M08

Included observations: 270

Autocorrelation	Partial Correlation	AC	PAC	Q-Stat	Prob	
		1	-0.001	-0.001	0.0004	0.984
		2	-0.013	-0.013	0.0484	0.976
		3	-0.006	-0.006	0.0598	0.996
		4	0.067	0.067	1.3154	0.859
		5	-0.117	-0.118	5.1049	0.403
		6	0.099	0.103	7.8126	0.252
		7	-0.006	-0.011	7.8224	0.349
		8	0.004	0.001	7.8259	0.451
		9	-0.079	-0.064	9.5932	0.384
		10	0.058	0.035	10.551	0.394
		11	0.082	0.105	12.445	0.331
		12	0.031	0.017	12.723	0.390
		13	-0.078	-0.067	14.470	0.342
		14	0.038	0.020	14.892	0.386
		15	-0.112	-0.106	18.506	0.237
		16	-0.033	-0.021	18.815	0.278
		17	-0.068	-0.079	20.167	0.266
		18	0.005	-0.023	20.173	0.323
		19	0.015	0.058	20.242	0.380
		20	0.061	0.045	21.332	0.378
		21	-0.016	0.005	21.403	0.435
		22	0.006	-0.025	21.413	0.495
		23	-0.006	0.005	21.425	0.555
		24	-0.004	-0.005	21.431	0.613
		25	-0.077	-0.076	23.212	0.565
		26	-0.069	-0.077	24.636	0.540
		27	0.005	0.026	24.643	0.594
		28	0.007	0.015	24.658	0.646
		29	0.085	0.115	26.865	0.579
		30	0.063	0.027	28.094	0.565
		31	0.029	0.016	28.358	0.603
		32	-0.047	-0.045	29.034	0.617
		33	0.002	-0.018	29.035	0.665
		34	-0.027	-0.035	29.259	0.699
		35	0.097	0.100	32.216	0.603
		36	-0.163	-0.150	40.591	0.275

Gráfica 10. Correlograma para los residuos.