



UNIVERSIDAD AUTÓNOMA CHAPINGO

---

DIVISIÓN DE CIENCIAS ECONÓMICO-ADMINISTRATIVAS

DOCTORADO EN CIENCIAS EN ECONOMÍA AGRÍCOLA

**PREDICCIÓN DEL PRECIO DE MAÍZ EN MÉXICO**

**TESIS**

Que como requisito parcial para obtener el grado de:

**DOCTOR EN CIENCIAS EN ECONOMÍA AGRÍCOLA**

**Presenta:**

**MARÍA DEL ROSARIO LÓPEZ GARCÍA**

Bajo la supervisión de:

**DR. MIGUEL ÁNGEL MARTÍNEZ DAMIÁN**

Chapingo, Estado de México, diciembre de 2020



**"PREDICCIÓN DEL PRECIO DE MAÍZ EN MÉXICO"**

Tesis realizada por **María del Rosario López García**, bajo la supervisión del Comité Asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

**DOCTOR EN CIENCIAS EN ECONOMÍA AGRÍCOLA**

DIRECTOR: \_\_\_\_\_

Dr. MIGUEL ÁNGEL MARTÍNEZ DAMIÁN

ASESOR: \_\_\_\_\_

Dr. RAMÓN VALDIVIA ALCALÁ

ASESOR: \_\_\_\_\_

Dr. JUAN HERNÁNDEZ ORTÍZ

LECTOR EXTERNO: \_\_\_\_\_

Dr. JOSÉ MIGUEL OMAÑA SILVESTRE

## CONTENIDO

1.	INTRODUCCIÓN GENERAL.....	1
1.1	Justificación .....	3
1.2	Objetivos.....	4
1.3	Hipótesis.....	4
2.	REVISIÓN DE LITERATURA .....	5
3.	SITUACIÓN ACTUAL DEL MAÍZ.....	11
3.1	Entorno Internacional.....	11
3.2	Entorno Nacional .....	12
3.3	Instrumentos financieros en el mercado de maíz .....	17
4.	MARCO TEÓRICO .....	19
4.1	Series de tiempo.....	19
4.2	Proceso estacionario .....	20
4.3	Ruido blanco.....	20
4.4	Modelos AR ( $p$ ).....	21
4.5	Modelos MA ( $q$ ) .....	21
4.6	Modelos ARMA.....	21
4.7	Procesos no estacionarios.....	22
4.8	Modelos ARIMA.....	23
4.9	Funciones de autocorrelación y autocorrelación parcial .....	24
4.10	Raíces unitarias .....	26
4.10.1	Prueba Dickey Fuller.....	26
4.10.2	Prueba de Phillips-Perron (PP) .....	28
4.11	Modelos vectoriales autorregresivos (VAR).....	28
4.12	Series cointegradas .....	31
4.13	Modelos vectoriales de corrección de error (VEC) .....	33
4.14	Regresión espuria.....	35
4.15	Discriminación de los modelos .....	36

4.16	Verificación del modelo: análisis de residuales.....	37
4.17	Pronóstico.....	37
4.18	Criterios de evaluación de la capacidad predictiva de los modelos .....	39
5.	MATERIALES Y MÉTODOS.....	41
5.1	Modelos ARIMA para el precio de maíz.....	41
5.2	Modelos VAR Y VEC .....	42
5.3	Predicción fuera de la muestra .....	45
6.	RESULTADOS .....	46
6.1	Análisis de datos de series univariadas.....	46
6.2	Identificación de modelos ARIMA.....	51
6.3	Predicción del precio de maíz mediante modelos autorregresivos integrados de promedio móvil (ARIMA) .....	55
6.4	Modelos vectoriales de corrección de error (VEC) con dos variables..	59
6.5	Predicción del precio de maíz mediante modelos VEC .....	68
6.6	Modelos vectoriales autorregresivos (VAR).....	72
6.7	Evaluación de la capacidad predictiva de los modelos .....	82
7.	CONCLUSIONES .....	84
8.	FUENTES CONSULTADAS .....	85
9.	APÉNDICE .....	90

## LISTA DE CUADROS

Cuadro 1. Prueba de raíz unitaria: Dickey Fuller Aumentada y Phillips Perron	49
Cuadro 2. Estimación de modelos ARIMA.....	52
Cuadro 3. Estimación de modelos ARIMA (continuación) .....	53
Cuadro 4. Prueba Ljung-Box a los residuales en modelos ARIMA .....	54
Cuadro 5. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie México .....	56
Cuadro 6. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie Michoacán .....	56
Cuadro 7. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie Jalisco .....	57
Cuadro 8. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie Sinaloa .....	57
Cuadro 9. Prueba de raíz unitaria a precios del mercado internacional.....	61
Cuadro 10. Contraste de cointegración de Johansen de modelos VEC .....	65
Cuadro 11. Prueba Ljung-Box a los residuales de modelos VEC.....	67
Cuadro 12. Precios observados y predichos obtenidos a partir de un modelo VEC, serie México .....	68
Cuadro 13. Precios observados y predichos obtenidos a partir de un modelo VEC, Serie Michoacán.....	68
Cuadro 14. Precios observados y predichos obtenidos a partir de un modelo VEC, Serie Jalisco .....	69
Cuadro 15. Precios observados y predichos obtenidos a partir de un modelo VAR, Serie Sinaloa .....	69
Cuadro 16. Prueba Ljung-Box a los residuales de los modelos VAR .....	75
Cuadro 17. Precios observados y predichos obtenidos a partir de los modelos VAR, serie México .....	76

Cuadro 18. Precios observados y predichos obtenidos a partir de los modelos VAR, serie Michoacán .....	76
Cuadro 19. Precios observados y predichos obtenidos a partir de los modelos VAR, serie Jalisco .....	77
Cuadro 20. Precios observados y predichos obtenidos a partir de los modelos VAR, serie Sinaloa .....	77
Cuadro 21. Criterios de discriminación de modelos .....	83

## LISTA DE FIGURAS

Figura 1. Producción mundial de maíz (Millones de toneladas).....	11
Figura 2. Producción de maíz en México (Miles de toneladas).....	12
Figura 3. Principales entidades productoras de maíz en México (2017-2018) .	13
Figura 4. Producción de maíz blanco y amarillo por ciclo comercial en México	14
Figura 5. Porcentaje de la producción mensual de maíz a nivel nacional.....	14
Figura 6. Precio de maíz promedio anual al productor en México (\$/Ton).....	15
Figura 7. Balanza comercial de maíz en México (Millones de toneladas).....	16
Figura 8. Evolución del comercio exterior de maíz en México .....	17
Figura 9. Proceso para el análisis de series de tiempo.....	35
Figura 10. Series de tiempo del precio de maíz blanco en México. ....	47
Figura 11. Logaritmo de las series de precios de maíz blanco en México .....	47
Figura 12. Series transformadas de precios de maíz blanco en México .....	48
Figura 13. Función de autocorrelación (ACF) de series de datos nacionales ...	50
Figura 14. Función de autocorrelación parcial (PACF) de series de datos nacionales.....	50
Figura 15. Función de autocorrelación (ACF) de la diferencia del logaritmo natural de series de datos nacionales.....	51
Figura 16. Función de autocorrelación parcial (PACF) de la diferencia del logaritmo natural de series de datos nacionales .....	52
Figura 17. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie MEXICO, modelo ARIMA (0,0,1) .....	54
Figura 18. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie MEXICO, modelo ARIMA (1,0,1) .....	55
Figura 19. Pronósticos del precio de Maíz blanco mediante modelos ARIMA..	58

Figura 20. Pronósticos del precio de Maíz blanco mediante modelos ARIMA (Continuación).....	59
Figura 21. Precio promedio mensual de maíz amarillo en físico (Precio FOB) .	60
Figura 22. Precio promedio de futuros de maíz .....	61
Figura 23. Precios de maíz en físico (FOB) (serie en primera diferencia) .....	62
Figura 24. Precio de maíz en el Estado de México Vs precio de maíz en físico EU.....	63
Figura 25. Precio de maíz en Michoacán Vs precio de maíz en físico EU.....	63
Figura 26. Precio de maíz en Jalisco Vs precio de maíz en físico EU .....	64
Figura 27. Precio de maíz en Sinaloa Vs precio de maíz en físico EU .....	64
Figura 28. Predicción fuera de la muestra, serie México, modelo VEC .....	70
Figura 29. Predicción fuera de la muestra, serie Michoacán, modelo VEC .....	70
Figura 30. Predicción fuera de la muestra, serie Jalisco, modelo VEC.....	71
Figura 31. Predicción fuera de la muestra, serie Sinaloa, modelo bivariado VAR(1) .....	71
Figura 32. Predicción fuera de la muestra, serie México, modelo VAR(2) .....	78
Figura 33. Predicción fuera de la muestra, serie México, modelo VAR(1) .....	78
Figura 34. Predicción fuera de la muestra, serie Michoacán, modelo VAR(3) ..	79
Figura 35. Predicción fuera de la muestra, serie Michoacán, modelo VAR(1) ..	79
Figura 36. Predicción fuera de la muestra, serie Jalisco, modelo VAR(2) .....	80
Figura 37. Predicción fuera de la muestra, serie Jalisco, modelo VAR(1) .....	80
Figura 38. Predicción fuera de la muestra, serie Sinaloa, modelo VAR(2) .....	81
Figura 39. Predicción fuera de la muestra, serie Sinaloa, modelo VAR(1) .....	81



## LISTA DE APÉNDICES

Apéndice 1. ACF Y PACF de las series originales.....	90
Apéndice 2. Verificación de ruido blanco en residuales en modelos univariados .....	90
Apéndice 3. Coeficientes estimados de modelos VEC .....	93
Apéndice 4. Diagnóstico de residuales en modelos VAR .....	97
Apéndice 5. Coeficientes estimados en modelos VAR .....	99
Apéndice 6. Prueba de raíz unitaria a modelos VAR .....	104

## SIGLAS Y ABREVIATURAS UTILIZADAS

SIGLA	SIGNIFICADO
SIAP	Servicio de Información Agroalimentaria y Pesquera
FIRA	Fideicomisos Instituidos en Relación con la Agricultura
FAO	Food and Agriculture Organization Organización de las Naciones Unidas para la Agricultura y la Alimentación
CEPAL	Comisión Económica para América Latina y el Caribe
SAGARPA	Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación
ASERCA	Agencia de Servicios a la Comercialización y Desarrollo de Mercados Agropecuarios
SINIM	Sistema Nacional de Información de Mercados
BMV	Bolsa Mexicana de Valores
MEXDER	Mercado Mexicano de Derivados S.A. de C.V.
BANXICO	Banco de México
USDA	United States Department of Agriculture
CBOT	Chicago Board of Trade
CME	Chicago Mercantile Exchange
FAOSTAT	Base de datos estadísticos corporativos de la Organización de las Naciones Unidas para la Alimentación y la Agricultura
FOB	del inglés Free On Board, Libre a bordo en puerto de carga convenido
ABREVIATURA	
<i>et al.</i>	y otros

## **DEDICATORIAS**

A mi hijo, mi más grande orgullo

y

A mi esposo, quien ha sido parte importante en el logro de mis metas  
profesionales.

Gracias por estar conmigo aun cuando el trabajo y el estudio ocuparon mi  
tiempo.

A mi familia, quienes han creído en mí siempre.

## **AGRADECIMIENTOS**

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca que me otorgo para la realización de mis estudios de doctorado.

A la Universidad Autónoma Chapingo, en particular al Posgrado en Economía, por haberme brindado la oportunidad de continuar mi formación académica.

A los integrantes de mi consejo particular:

Al Dr. Miguel Ángel Martínez Damián, por su orientación, sus observaciones, su paciencia y por el tiempo dedicado para la realización del presente trabajo de tesis.

Al Dr. Ramón Valdivia Alcalá por sus observaciones a este trabajo de tesis.

Al Dr. Juan Hernández Ortiz, por su colaboración desinteresada en el presente trabajo.

Al lector externo, el Dr. José Miguel Omaña Silvestre por su importante colaboración para concluir esta tesis.

Al Dr. José Jaime Arana Coronado, por su asesoría y tiempo dedicado.

A mis profesores, por haber contribuido a mi formación académica.

A mis compañeros de clase, que se convirtieron en amigas y amigos, por ser una parte tan importante de mi vida.

Al personal administrativo y a todas aquellas personas que de alguna manera fueron coparticipes de esta tarea.

## DATOS BIOGRÁFICOS



### Datos personales

**Nombre:** María del Rosario López García

**Lugar de nacimiento:** San Felipe del Progreso, Estado de México.

**Profesión:** Licenciado en Economía Agrícola

**Cédula profesional:** 7619908

**Maestría:** Estadística Aplicada

**Cédula de maestría:** 10155935

### Desarrollo académico

**Bachillerato:** Universidad Autónoma Chapingo. Preparatoria agrícola.

**Licenciatura:** Universidad Autónoma Chapingo. Licenciatura en Economía Agrícola. División de Ciencias Económico-Administrativas.

**Maestría:** Colegio de Postgraduados campus Montecillo. Programa de Socioeconomía estadística e Informática-Estadística (SEI-EST).

## RESUMEN GENERAL

### PREDICCIÓN DEL PRECIO DE MAÍZ EN MÉXICO<sup>1</sup>

El maíz es uno de los productos más importantes en México, porque es parte esencial en la dieta de los mexicanos y de importancia económica en la producción pecuaria y uso industrial. Este producto está sujeto a factores que causan volatilidad en sus precios, adicionalmente, los productores se enfrentan a la falta de información debido a que en México no existe una señal sobre el comportamiento de los precios futuros de maíz blanco como el Chicago Mercantile Exchange (CME). En este contexto de alta volatilidad, un pronóstico permitirá a productores, comercializadores e industriales del sector protegerse de variaciones en el precio del maíz. Por ello, en este trabajo se tuvo como objetivo el construir estimadores de predicción. Se modelaron cuatro series de tiempo correspondientes al precio de maíz en los estados de: México, Michoacán, Jalisco y Sinaloa, con base en la metodología ARIMA, modelos VAR y VEC, incorporando series de precios de futuros de maíz y precios de maíz en físico (FOB). Se evaluó la capacidad predictiva de los modelos con base en el error porcentual absoluto medio, la raíz del error cuadrático medio y la U de Theil. Los resultados muestran que, la influencia de los precios del mercado internacional, particularmente los precios de futuros y físicos (FOB) difiere en cada uno de los estados analizados. Los modelos multivariados proporcionaron un mejor predictor del precio en Sinaloa y Michoacán, sin embargo, en el Estado de México los modelos ARIMA tuvieron un mejor desempeño, por lo que se concluye que los modelos univariados también son adecuados para predecir.

**Palabras clave:** maíz, series de tiempo, ARIMA, VAR, VEC, predictor.

---

<sup>1</sup> Tesis de Doctorado en Ciencias, Universidad Autónoma Chapingo

Autor: María del Rosario López García.

Director: Dr. Miguel Ángel Martínez Damián

## GENERAL ABSTRACT

### PREDICTION OF THE PRICE OF CORN IN MEXICO<sup>2</sup>

Corn is one of the most important products in Mexico, because it is an essential part of the Mexican diet and of economic importance in livestock production and industrial use. This product is subject to factors that cause volatility in its prices, additionally, producers face a lack of information because in Mexico there is no signal on the behavior of future prices of white corn as the Chicago Mercantile Exchange (CME). In this context of high volatility, a forecast will allow producers, marketers and industrialists in the sector to protect themselves from variations in the price of corn. Therefore, the objective of this work was to construct prediction estimators. Four time series corresponding to the price of corn were modeled in the states of: Mexico, Michoacan, Jalisco and Sinaloa, based on the ARIMA methodology, VAR and VEC models incorporating series of corn futures prices and physical corn prices (FOB). The predictive capacity of the models was evaluated based on the absolute mean percentage error, the root mean square error and Theil's U. The results show that the influence of international market prices, particularly future and physical prices (FOB) differs in each of the analyzed states. The multivariate models provided a better price predictor in Sinaloa and Michoacán, however, in the State of Mexico, the ARIMA models had a better performance, so it is concluded that the univariate models are also adequate to predict.

**Key words:** corn, time series, ARIMA, VAR, VEC, predictor.

---

<sup>2</sup> Tesis de Doctorado en Ciencias, Universidad Autónoma Chapingo

Autor: María del Rosario López García.

Director: Dr. Miguel Ángel Martínez Damián

## **1. INTRODUCCIÓN GENERAL**

El maíz grano es el cultivo con el mayor valor económico generado por su venta en México; tiene un 89% de participación nacional en la producción de granos, aunque su importancia nacional va más allá de la esfera productiva (SIAP, 2019).

El maíz, como producto agrícola, está sujeto a factores que causan volatilidad en sus precios, algunos de estos son: cambios en la oferta y demanda, efectos negativos del clima en los cultivos, ajustes en la política económica interna y externa y en general el contexto internacional del mercado de este producto. Adicionalmente a todos los factores perjudiciales inherentes al proceso productivo, los productores se enfrentan a la falta de información sobre el comportamiento de los mercados; esto afecta la toma de decisiones, por ello, es importante proporcionar predicciones y/o estimaciones sobre los precios del maíz que les permitan a los productores, comercializadores e industriales tomar acciones para protegerse de variaciones en el precio,

Los productores pueden prever el precio de venta de su producto, y tomar decisiones anticipadas sobre la siembra, del mismo modo, pueden plantear escenarios sobre sus posibles ingresos tomando en cuenta el precio pronosticado durante la época de cosecha y/o fecha de venta de su producto. Los industriales y empresas que utilizan el maíz en su actividad productiva u otro producto agropecuario cuyo precio este altamente correlacionado con el precio del maíz, pueden tener información sobre el precio de abasto de su insumo.

Desde el año 2012, los precios de maíz en el mercado spot y futuros del maíz amarillo que cotiza en la bolsa de valores de Chicago (Chicago Mercantile Exchange) han sido tomados como predictores del precio de maíz blanco y amarillo en México (tomando en cuenta los diferenciales de precios debido a costos de comercialización). En México ya existen instrumentos derivados (contratos de futuros) que ofrece el MEXDER únicamente para maíz amarillo.



Cabe resaltar que el maíz blanco no cuenta con dicho instrumento financiero. En 2015, el 85.9% de la producción nacional de maíz correspondió a maíz blanco, 13.6% a maíz amarillo y el restante 0.5% a otros tipos de maíz (FIRA, 2016), esto pone en evidencia la importancia de contar con un predictor específico para este producto.

En este trabajo de investigación se realizó un análisis de los precios del maíz blanco en México. Se modelaron cuatro series de tiempo correspondientes a los estados de: México, Michoacán, Jalisco y Sinaloa, usando la metodología de modelos autorregresivos de media móvil (ARIMA), modelos vectoriales autorregresivos (VAR) y modelos vectoriales de corrección de error (VEC), incorporando series de precios de futuros de maíz<sup>3</sup> y precios de maíz en físico<sup>4</sup>, con el objetivo de proporcionar un predictor de los precios del maíz blanco en México. Se realizaron estimaciones puntuales y se evaluó la capacidad predictiva de los modelos con base en los criterios: error porcentual absoluto medio (MAPE, del inglés Mean Absolute Percentage Error), la raíz del error cuadrático medio (RECM) y la U de Theil.

En el primer capítulo se exponen los objetivos, justificación y las hipótesis probadas en este trabajo. En el segundo capítulo se muestra una revisión de las investigaciones recientes en las que se utilizaron los modelos ARIMA, VAR, VEC, entre otras. Posteriormente, en el tercer capítulo se describe brevemente el entorno internacional del maíz con respecto a la producción y el entorno nacional con respecto a la producción, disponibilidad, consumo, la balanza comercial e instrumentos financieros utilizados por los productores de maíz en México. En el cuarto capítulo se exponen los conceptos teóricos de los modelos autorregresivos integrados de media móvil (ARIMA), vectoriales autorregresivos (VAR) y vectoriales de corrección de error (VEC). En capítulo número cinco, se describen los datos utilizados y la metodología para la construcción de los modelos ARIMA, VAR y VEC. En el sexto capítulo se muestran los resultados de

---

<sup>3</sup> Precio de liquidación mensual promedio del contrato de futuros de maíz más cercano

<sup>4</sup> Precios promedio mensuales de maíz amarillo No. 2, FOB en el Golfo de México

las pruebas estadísticas, así como, las predicciones realizadas usando los modelos ajustados comparadas con los valores observados para el periodo de julio a diciembre de 2018 y los criterios de discriminación de los modelos. Finalmente, en el capítulo siete se exponen las conclusiones derivadas de este trabajo de investigación

## **1.1 Justificación**

“Las fluctuaciones de precios son un atributo normal y un requisito necesario para el funcionamiento competitivo del mercado”, sin embargo, “la volatilidad extrema de los precios tiene un costo ya que los actores del mercado tendrán dificultades para planificar con anticipación y adaptarse a las señales fluctuantes del mercado” (FAO, 2010), por otra parte, los agricultores basan sus decisiones de inversión y planificación en los precios futuros esperados. Dado lo anterior, los predictores contruidos mediante las metodologías propuestas son una herramienta para conocer el comportamiento del mercado del maíz, mediante la estimación de pronósticos a corto plazo y además proporcionan información sobre el rumbo del mercado, lo que ayudará en la toma de decisiones.

La volatilidad de los precios de las materias primas agrícolas tiene costos para los países, tales como pérdidas en la eficiencia económica, reducción de la seguridad alimentaria, efectos negativos sobre la balanza comercial, problemas sociales y riesgos elevados especialmente para los pequeños agricultores, pues se incrementa la incertidumbre sobre sus ingresos esperados (CEPAL *et al.*, 2011).

Particularmente para los países de América Latina y el Caribe (LAC) la volatilidad de precios tiene diferentes connotaciones, dependiendo de la relación entre la estructura de sus exportaciones e importaciones y el comportamiento de los precios de los productos que exporta o importa (CEPAL *et al.*, 2011).

## **1.2 Objetivos**

### Objetivo General

- Construir estimadores de predicción para el precio de maíz blanco en México en los estados: México, Michoacán, Jalisco y Sinaloa.

### Objetivos particulares

- Comparar los predictores univariados ARIMA y los multivariados VAR y VEC con base en los criterios de discriminación establecidos.
- Presentar evidencia empírica de la relación entre el precio de futuros de maíz, el precio de maíz en físico (FOB) y los precios nacionales.

## **1.3 Hipótesis**

- Los modelos multivariados VAR y VEC proporcionan predictores con menores: U de Theil, raíz del error cuadrático medio y error porcentual absoluto medio, comparados con los modelos univariados ARIMA, debido a que incorporan información del mercado internacional.
- Existe influencia de los precios del maíz en el mercado Internacional sobre el mercado interno en México.

## 2. REVISIÓN DE LITERATURA

Los modelos autorregresivos integrados de promedios móviles (ARIMA) han sido aplicados para pronosticar variables por ejemplo en: Delgadillo *et al.* (2016), Contreras *et al.* (2016) , Barreras *et al.* (2013), Sánchez *et al.* (2013) y Marroquín y Tovar (2010); otros estudios se enfocan en analizar y modelar variables, por ejemplo: Contreras *et al.* (2003), Botero y Cano (2008), Gil y Maya (2008) y Morana (2001), entre otros.

Disponer de pronósticos de variables económicas particularmente en el sector agropecuario resulta útil a los productores, comercializadores e inversionistas financieros con el fin de tomar decisiones de forma anticipada al mercado; la aplicación de técnicas de modelación de series de tiempo para obtener predicciones ha sido de interés en Delgadillo *et al.* (2016) quienes compararon modelos de series de tiempo ARIMA para pronosticar el rendimiento de granos básicos: maíz, frijol, trigo y arroz en México, con el objetivo de predecir dicha variable en el corto plazo.

Contreras *et al.* (2016) hicieron uso de modelos de promedio móvil simple, promedio móvil ponderado, suavización exponencial y suavización exponencial ajustada para pronosticar la demanda de almacenamiento en productos perecederos, el volumen de ingreso y egreso de los productos en una cámara frigorífica, para prever los requerimientos de instalaciones, personal y materiales necesarios. Determinaron que la técnica de promedios móviles ponderados tuvo mejores resultados.

En el mismo sentido de construcción de modelos univariados Barreras *et al.* (2013) mostraron que mediante los modelos Autorregresivos de Media Móvil (ARMA) es posible construir predictores de la producción porcícola en Baja California, México. En su estudio aplicaron la metodología de Box-Jenkins para ajustar modelos ARMA (12,12) y AR(12) a datos mensuales de producción de carne de cerdo.

Otro ejemplo de predicción de variables en el sector agropecuario se hizo en Sánchez *et al.* (2013) quienes plantearon un modelo ARIMA para pronósticar el comportamiento de la producción de leche de bovino en el estado de Baja California y concluyeron que un modelo ARMA (1,1) proporciona pronósticos adecuados a corto plazo.

Del mismo modo, Marroquín y Tovar (2010) ajustaron un modelo ARIMA (23,0,1) a una serie de precios nominales de venta al mayoreo de jitomate “bola” en México, usando la metodología Box-Jenkins y posteriormente construyeron predicciones del comportamiento del precio del jitomate en los siguientes 12 meses.

Los modelos ARIMA han sido aplicados en el análisis y modelación de variables que pertenecen al sector energético; (Gil y Maya, 2008) emplearon un procedimiento para modelar el precio de la energía eléctrica en el mercado de contado y de derivados sobre este subyacente en Colombia, usando un modelo de volatilidad condicional autorregresivo generalizado GARCH ( $p, q$ ). Particularmente el estudio de estos autores se enfocó en la modelación de los precios de energía eléctrica y la modelación de la media y la varianza de los precios de ésta. Explican que, en el caso de los precios de la energía eléctrica, las series de variaciones logarítmicas de precios se caracterizan por presentar exceso de curtosis, conglomerados de volatilidad y convergencia o reversión a la media y concluyeron que el mejor ajuste para la volatilidad del precio de energía eléctrica se logra con un modelo exponencial generalizado autorregresivo, condicionalmente heterocedástico (EGARCH) (1,1) asumiendo una distribución generalizada del error (GED).

Por otra parte, Botero y Cano (2008) hicieron un análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia usando modelos autorregresivos y modelos estacionales. Así mismo, Contreras *et al.* (2003) analizaron los mercados de energía eléctrica de España peninsular y California y propusieron dos modelos ARIMA para la predicción de precios. Sus

resultados mostraron que las predicciones de precios obtenidas para los mercados de energía eléctrica, español y californiano son lo suficientemente precisas para ser utilizadas por productores y consumidores para preparar sus estrategias de licitación correspondientes.

Una aplicación de la metodología semiparamétrica basada en aproximaciones Bootstrap se mostró en Morana (2001) donde se aplicaron modelos GARCH para modelar el precio del petróleo a corto plazo. En su análisis se explica que el nivel del precio del petróleo no muestra una tendencia global, por el contrario, muestra una tendencia local estocástica, y la tasa absoluta de variación del precio del petróleo presenta períodos alternos de alta volatilidad, seguida de períodos de relativa tranquilidad, las cuales, son características comunes en la especulación de precios de los activos. Con base en sus resultados sugirió modelar el precio del petróleo en términos de un proceso de martingala<sup>5</sup> por considerarlo más apropiado que usar un modelo de paseo aleatorio.

Los modelos VAR (Vectoriales Autorregresivos) y VEC (Vectoriales de Corrección de Errores) son metodologías para predecir y modelar el comportamiento de series temporales, éstos pueden ofrecer ventajas sobre los modelos Box-Jenkins (modelos univariados) ya que los modelos VAR permiten que la variable dependa de sus propios rezagos y los rezagos de otras variables que influyen sobre el comportamiento de las variables de interés<sup>6</sup>. Estudios relacionados con la modelación de los precios del maíz en México aplicando estas metodologías se muestran en Zavaleta (2019), Araujo (2011), Godínez (2006) y Lederman (2004).

Zavaleta (2019) analizó la dinámica estocástica de las variaciones porcentuales del precio del maíz, mediante una estructura Autorregresiva y de Promedios Móviles (ARMA) y mediante un Vector Autorregresivo (VAR), para modelar

---

<sup>5</sup> Una aplicación de la hipótesis de martingala puede encontrarse en De la Uz (2002)

<sup>6</sup> Las variables que influyen sobre la variable de interés especificadas en un modelo VAR son determinadas con base en evidencia empírica o teoría relacionada con el fenómeno objeto de estudio.

posteriormente la volatilidad a través de estructuras GARCH (Generalized Autoregressive Conditionally Heteroscedasticity), esto con la finalidad de medir el impacto en la volatilidad de los precios, que ha tenido la introducción de “futuros” de maíz amarillo en el mercado de derivados y afirma que la volatilidad del precio físico del maíz amarillo disminuyó posterior a la introducción del contrato futuro.

En su investigación Araujo (2011) utilizó modelos de corrección de errores para analizar la transmisión de precios entre los mercados de maíz mexicanos y el mercado estadounidense; por otra parte, Godínez (2006) uso modelos VAR para probar la inexistente causalidad y liderazgo del precio futuro sobre los precios físicos de maíz blanco en México.

Se han propuesto modelos univariados y multivariados con fines comparativos en Zavaleta (2019) y Arias y Torres (2004) quienes propusieron modelos VAR y VEC para el pronóstico de las importaciones de Costa Rica, como una alternativa a los modelos ARIMA usados previamente por el Banco Central de Costa Rica (BCCR), y determinaron que, en su estudio, los VEC tuvieron un mejor desempeño para pronosticar a corto plazo en comparación con los modelos ARIMA y los VAR.

Otras variables macroeconómicas son estudiadas en Morán (2014) donde se construyeron modelos VAR usando las variables: nivel de precios, inflación, salario nominal, oferta monetaria, gasto público y tipo de cambio real; posteriormente realizaron una estimación y proyección de la evolución futura de dichas variables.

Investigaciones similares se centran en evidenciar relaciones entre variables, tal es el caso de Galindo y Cardero (1997) mediante la construcción de modelos VAR a partir de las variables: precios, un agregado monetario, el ingreso y la tasa de interés nominal en México. Análogamente, Rodríguez (2002) modeló el comportamiento de la economía de Puerto Rico haciendo uso de modelos VAR y vectores de cointegración. El estudio se centró en la relación de equilibrio a

largo plazo de la tasa de interés, el índice de precios, la producción real y la oferta monetaria de Estados Unidos.

Rodríguez *et al.* (2019) analizó el efecto de la tasa representativa de mercado (TRM) e Índice de tasa de cambio real (TCR) en el Índice de Precios al Consumidor para Colombia (IPC) mediante un modelo VEC. Estos modelos han sido usados para realizar análisis de series de tiempo multivariados, en los que destaca el hecho de que las variables presentan cointegración, es decir, una relación de equilibrio en el largo plazo.

Otras aplicaciones de la metodología de modelos VEC se desarrollaron en Gutiérrez y Vásquez (2008) en un análisis de cointegración con base en las metodologías de Johansen (1988) y Lütkepohl (2005) para medir la sensibilidad del riesgo de crédito ante cambios en la actividad económica, la tasa de desempleo y la tasa de interés. En Castillo (2017) se estableció una relación de largo plazo entre las importaciones, exportaciones y los títulos de deuda pública sobre el riesgo país de Colombia, mediante la estimación de un modelo VEC.

Una herramienta importante proporcionada por los modelos VAR y VEC es el análisis de las funciones impulso respuesta (FIR) que indica la respuesta dinámica de la variable dependiente en el modelo VAR ante choques en los términos de error o innovaciones de todas las variables endógenas, prescindiendo de los efectos de las variables que explícitamente se establecen como exógenas. Nájera (2019) presentó un análisis sobre el comportamiento de seis criptomonedas usando modelos VAR y VEC, funciones de impulso respuesta, prueba de cointegración de Johansen y causalidad de Granger, y estableció la existencia de una interrelación a corto plazo entre comportamiento de las criptomonedas.

Se han hecho comparaciones del desempeño de modelos multivariados en Lanteri (2010) mediante la realización de pronósticos de las exportaciones de Argentina con base en modelos VAR sin restricciones, VAR Bayesianos (BVAR y BVAR-Kalman), modelos FAVAR (Factor aumentado VAR), resultando estos



últimos con mejor desempeño en la predicción comparados con los modelos VAR sin restricciones. Por su parte, Guerrero (2007) propuso modelos VAR, VEC y BVAR para verificar la viabilidad de las metas establecidas por el gobierno mexicano para algunas variables de importancia macroeconómica y concluyó que los modelos Bayesianos proporcionaron mejores pronósticos, sin embargo, advirtió que no se debe generalizar el uso de un modelo BVAR para la obtención de pronósticos de cualquier variable macroeconómica, por el contrario, cada modelo es útil para pronosticar las variables que inicialmente se estudian.

Otras metodologías utilizadas para predicción de series de tiempo son las redes neuronales artificiales Kaastra y Boyd (1996), Mercado *et al.* (2015), y Velásquez y Franco (2010) quienes hicieron uso de redes neuronales con arquitectura dinámica. De acuerdo con su investigación los modelos de redes neurales pronosticaron el precio de contratos de electricidad en Colombia para el siguiente mes con una mayor precisión comparados con la metodología ARIMA clásica.

### 3. SITUACIÓN ACTUAL DEL MAÍZ

El maíz es uno de los productos agrícolas más importantes a nivel mundial debido a su uso para consumo humano, producción pecuaria, y en industrias como la farmacéutica, bioenergética, cosmética entre otras. Además, el cultivo de este producto genera fuentes de empleo en México y en el mundo. En este capítulo se describe brevemente el panorama general de la producción a nivel internacional, así como la producción, disponibilidad, precios y balanza comercial a nivel nacional.

#### 3.1 Entorno Internacional

De acuerdo con FIRA (2016) entre los ciclos comerciales 2006 a 2016, la producción de maíz en el mundo mostró un crecimiento promedio anual de 3.3%, teniendo en el ciclo 2016-2017 una producción de 959.1 millones de toneladas, siendo los principales productores: Estados Unidos, con el 37%, China, 21% y Brasil, con una participación cercana a 8% del total (Figura 1).

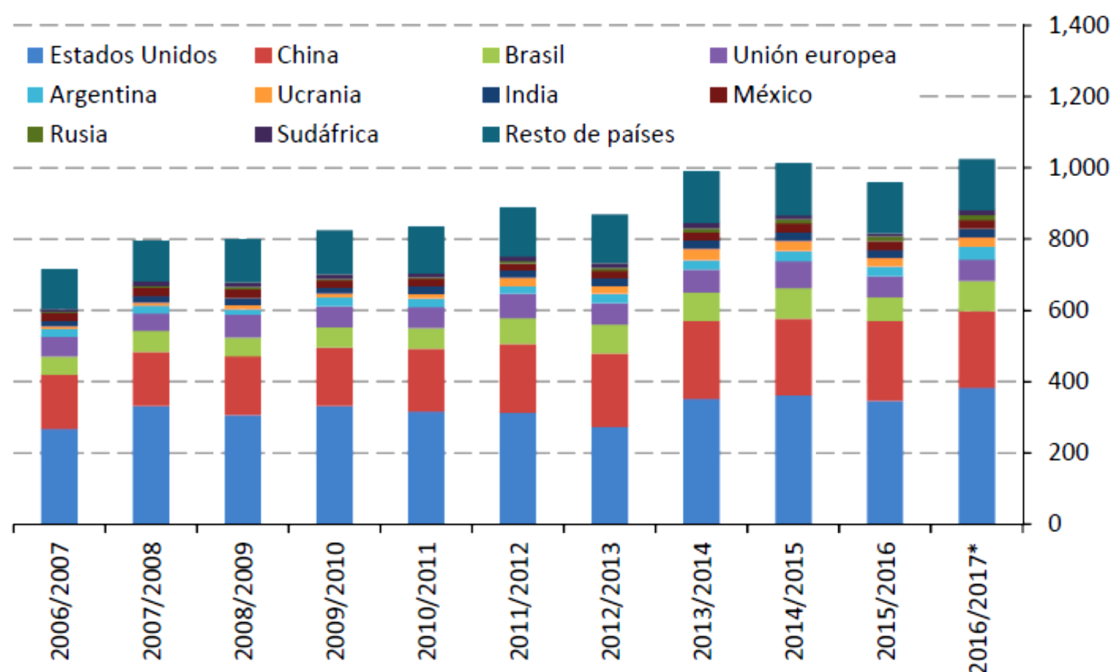


Figura 1. Producción mundial de maíz (Millones de toneladas)

Fuente: FIRA (2016). \*Dato estimado

México es el 7° productor mundial de maíz con 27,169,400 toneladas y aporta 2.4% al volumen internacional de maíz. Se estima que el consumo mundial de maíz crecerá a una tasa promedio anual de 1.2% entre 2016 y 2025 (FIRA, 2016).

### 3.2 Entorno Nacional

El maíz es el cultivo más representativo de México debido a su importancia económica, social y cultural. Las condiciones climáticas permiten que gran parte del territorio nacional sea propicio para la producción de maíz grano (SAGARPA, 2017). En las últimas décadas la producción de maíz en México ha tenido una tendencia creciente superando los 27 millones de toneladas en 2018 (Figura 2).

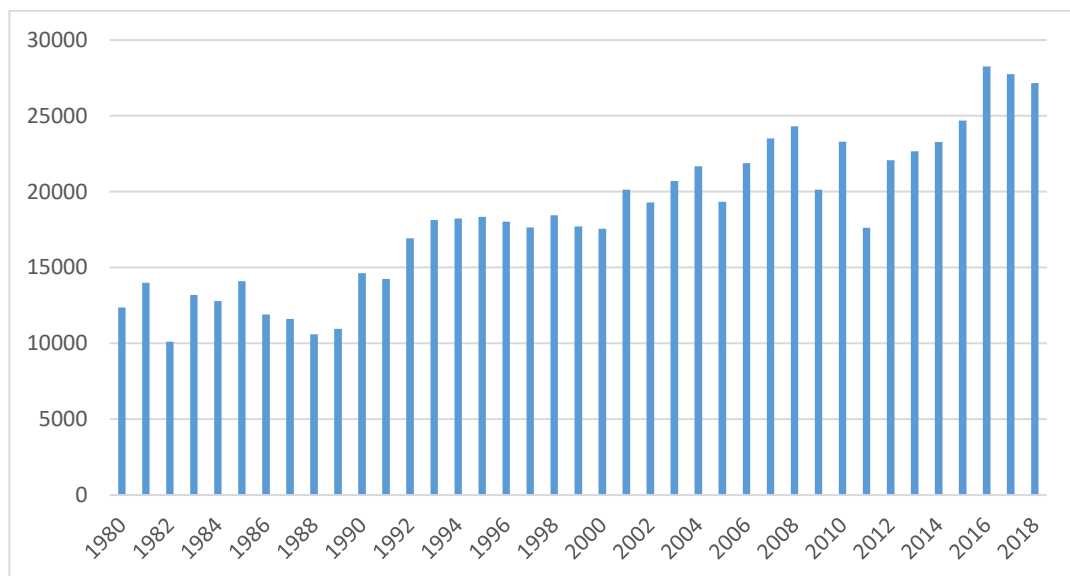


Figura 2. Producción de maíz en México (Miles de toneladas)

FUENTE: Elaboración propia con datos de FAOSTAT

En 2017 y 2018 los estados con mayor participación en el total de la producción nacional de maíz en México fueron los estados de Sinaloa, Jalisco, Michoacán y Estado de México con una participación equivalente a 49% de la producción total (Figura 3).

Los estados con el mayor porcentaje del valor (\$) de la producción son: Sinaloa 21.4%, líder nacional en el cultivo, lo cual se sustenta en los niveles de productividad que tiene en la cosecha del maíz blanco, en segundo lugar, Jalisco 14.4% el cual deriva de la extensa superficie que destina para el cultivo y Michoacán 7.4%, (SIAP, 2019).

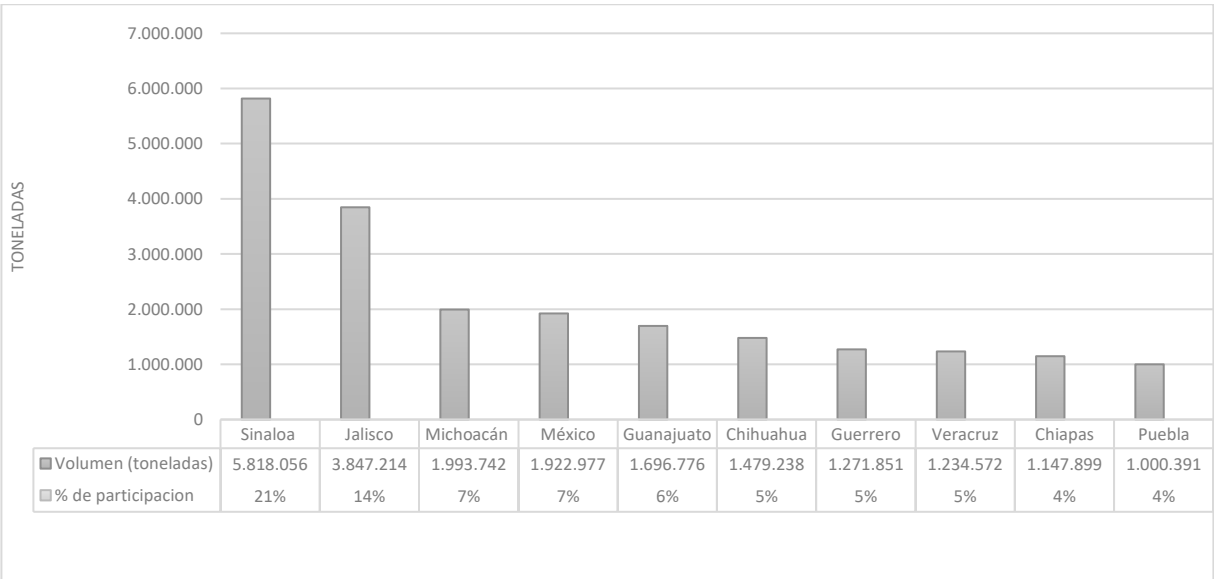


Figura 3. Principales entidades productoras de maíz en México (2017-2018)

Fuente: Elaboración propia con datos de SIAP

En 2015, el 85.9% de la producción nacional correspondió a maíz blanco, 13.6% a maíz amarillo y el restante 0.5% a otros tipos de maíz. Cabe destacar que la proporción de maíz amarillo incrementó pasando de 6.9% en 2005 a 13.6% en 2015 (FIRA, 2016).

En México la producción de maíz blanco durante 2013 a 2017 superó los 20 millones de toneladas, por otra parte, la producción de maíz amarillo fue alrededor de 3 millones de toneladas.

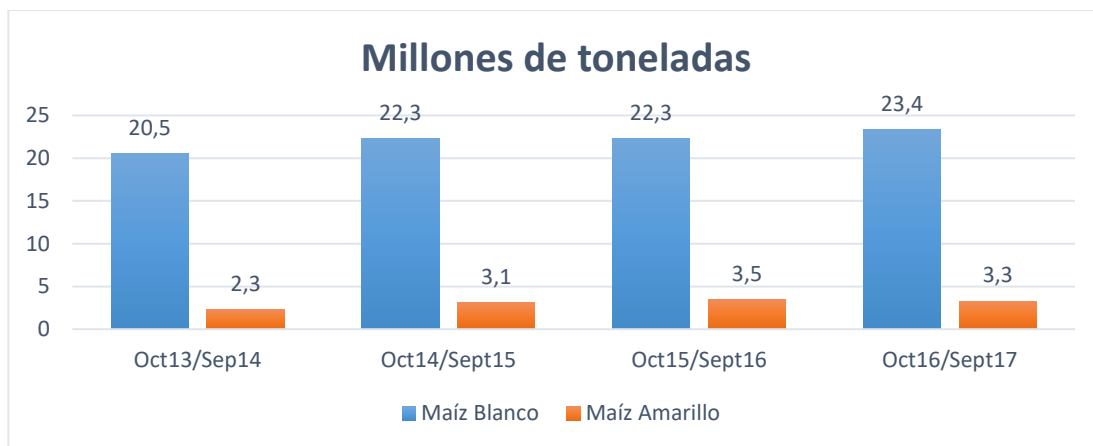


Figura 4. Producción de maíz blanco y amarillo por ciclo comercial en México

Fuente: Elaboración propia con datos de FIRA (2016)

La mayor disponibilidad del grano se da en dos periodos, noviembre-enero y mayo-junio.

La estacionalidad de la producción del maíz está relacionada con los periodos de cosechas. El 71.5% de la producción nacional se obtiene entre octubre y febrero (P-V) y el 22.7% entre mayo y julio (O-I). Las tres principales entidades cosechan más del 90% de su producción en las siguientes fechas: Sinaloa entre mayo y julio; Jalisco, entre noviembre y febrero; y Estado de México, entre noviembre y enero (FIRA, 2016).

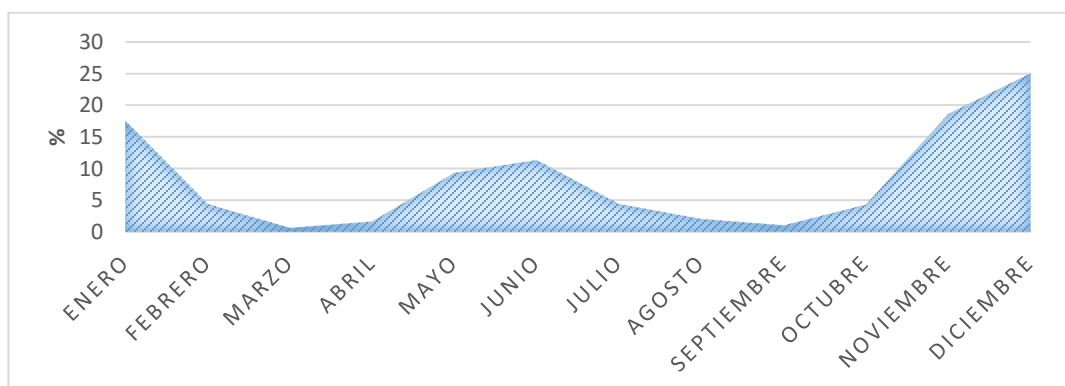


Figura 5. Porcentaje de la producción mensual de maíz a nivel nacional

Fuente: Elaboración propia con datos de FIRA (2016).

El precio al productor promedio anual de maíz en México ha mantenido una tendencia creciente en los últimos años (Figura 6), mostrando un comportamiento estable en periodos cortos (5 años aproximadamente), sin embargo, mostró un comportamiento variable a partir del año 2010, manteniendo un precio promedio al productor entre 3000 y 4000 pesos por tonelada. Cabe recordar que el precio a nivel nacional está sujeto a las condiciones de oferta y demanda, así como las condiciones del mercado y las perspectivas a nivel internacional.

De acuerdo con FIRA (2015), el precio al productor de maíz se incrementó en 2011 y 2012 debido a la poca disponibilidad del grano en el mercado nacional ante la sequía y condiciones de estiaje que prevalecieron en casi todo el país. A partir del 2013 el precio al productor presenta una tendencia a la baja ante periodos de producción más favorables.

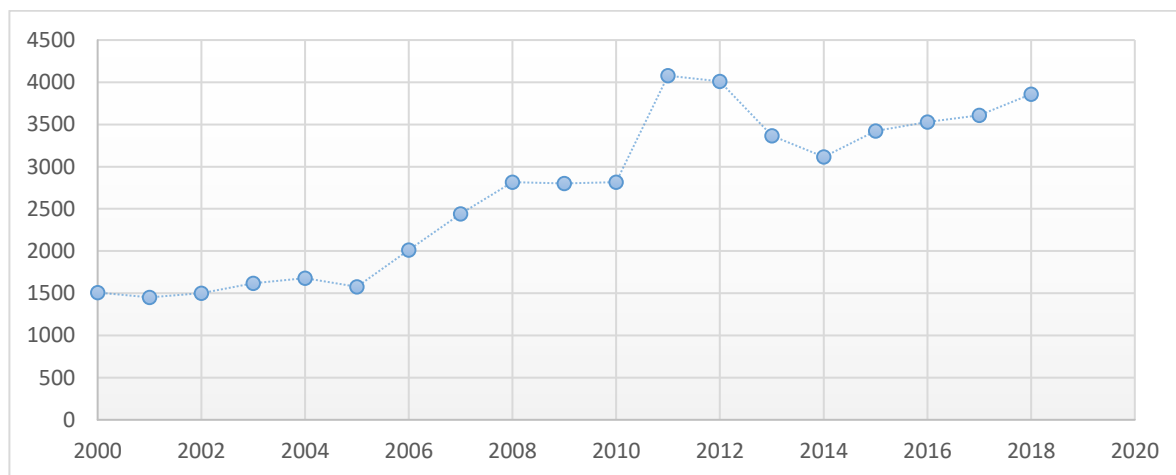


Figura 6. Precio de maíz promedio anual al productor en México (\$/Ton)

Fuente: Elaboración propia con datos de FAOSTAT

De acuerdo con estimaciones de SAGARPA se tuvo un incremento de 5.8% con respecto a 2014 en el nivel de consumo nacional situándose en 35.6 millones de toneladas durante 2015. El 64% del consumo total de maíz correspondió a maíz blanco, y el 36% restante a maíz amarillo (FIRA, 2015).

El consumo promedio per cápita anual en México es de 196.4 kg de maíz blanco, (SAGARPA, 2017). El 90% del total de maíz procedente del exterior es amarillo y fue utilizado en actividades ganaderas y agroindustriales. Los principales socios comerciales de México en la adquisición de este grano son Estados Unidos, Brasil y Argentina (SIAP, 2019).

Las importaciones de maíz mostraron una tendencia creciente en volumen y valor (\$) de la producción. Durante 2015 se importó el volumen de maíz más alto registrado hasta 2015, más de 12 millones de toneladas. Estados Unidos ha sido el principal proveedor de maíz grano en los últimos años, por lo tanto, la balanza comercial de maíz de México fue deficitaria en el periodo de 2005 a 2015, presentando un saldo negativo superior a 10 y 12 millones de toneladas, en 2014 y 2015 respectivamente.

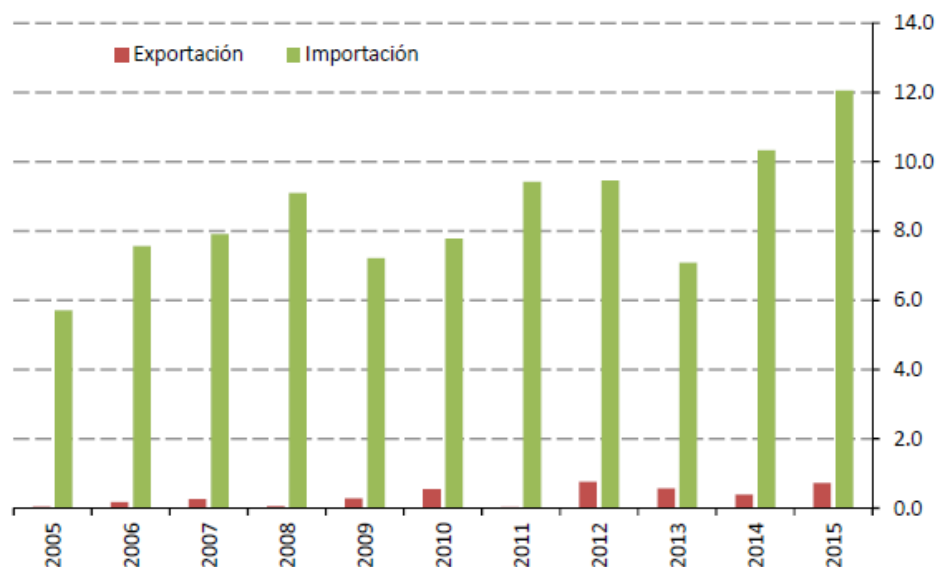


Figura 7. Balanza comercial de maíz en México (Millones de toneladas)

Fuente: FIRA (2016)

Del mismo modo, el volumen de importaciones ha tenido una tendencia creciente en los últimos años superando los dos mil novecientos millones de dólares en 2018.



Figura 8. Evolución del comercio exterior de maíz en México

Fuente: Elaboración propia con datos de SIAP (2019).

### 3.3 Instrumentos financieros en el mercado de maíz

Un contrato de futuros es un acuerdo para comprar o vender un activo en una fecha específica en el futuro a un precio determinado (Hull, 2009). En otras palabras, un futuro permite a los inversionistas conocer con certeza el precio al cual podrán adquirir o vender un bien o activo subyacente en una fecha posterior (BMV, 2018) .

En 2012 se introdujo el futuro de maíz en el mercado de derivados en México, en la Bolsa Mexicana de Valores (BMV). Este instrumento tiene las siguientes ventajas: cada contrato es de 25 toneladas, tiene fecha de vencimiento y calidad (amarillo número 2) específicos que permiten la gestión de riesgos de precios, teniendo como contraparte el mercado de derivados (MexDer). El precio de este



instrumento está denominado en pesos (BMV, 2012). Las posiciones de compra o venta se hacen ante el MexDer<sup>7</sup> y es esta entidad quien realiza las compensaciones a través de su cámara de compensación (ASIGNA)<sup>8</sup>.

El contrato de maíz amarillo no. 2 FOB (free on board) puesto en el Golfo que cotiza en la bolsa de valores de Chicago (CME) ha sido tomado como referente del precio internacional del maíz (FIRA, 2016), sin embargo, si se desea comparar los precios en el mercado mexicano, para la toma de decisiones, adicionalmente deben considerarse los costos de transporte.

Desde 1993 el gobierno federal mexicano, a través del programa Apoyos y Servicios a la Comercialización Agropecuaria (ASERCA), ha promovido la adquisición de coberturas anticipadas que permiten a los productores garantizar los precios de sus cosechas. Mediante este programa gubernamental se permite a los usuarios el acceso a líneas de financiamiento garantizado por la Banca de Desarrollo, con la cual los productores pueden adquirir “futuros” y “opciones” de la Bolsa de Chicago de los Estados Unidos con la finalidad de cubrir el riesgo de movimientos adversos de los precios físicos y al mayoreo de maíz blanco a través de la adquisición de futuros de maíz amarillo.

---

<sup>7</sup> Mercado Mexicano de Derivados, S.A. de C.V. se encarga de proveer las instalaciones y demás servicios para que se coticen y negocien contratos de derivados.

<sup>8</sup> ASIGNA, es un Fideicomiso de administración y pago constituido en 1998 con el objeto de compensar y liquidar las operaciones de productos derivados realizadas en MexDer (grupo BMV).

## 4. MARCO TEÓRICO

Existen diferentes metodologías para el análisis de series temporales, sin embargo, esta investigación se enfocó en el uso de modelos Autorregresivos Integrados de Promedio Móvil (ARIMA), Vectoriales Autorregresivos (VAR) y Vectoriales de Corrección de Error (VEC) para la generación de pronósticos. En este capítulo se expone brevemente los conceptos teóricos de estos modelos.

### 4.1 Series de tiempo

Un proceso estocástico es una secuencia de variables aleatorias  $\{Y_t: t = 0, \pm 1, \pm 2, \pm 3, \dots\}$  ordenadas y equidistantes cronológicamente. Donde  $Y_t$  es una variable aleatoria escalar referida a la unidad observable considerada en el momento  $t$ . Se sabe que la estructura probabilística completa de dicho proceso está determinada por el conjunto de distribuciones de todas las colecciones finitas de las  $Y$ . Gran parte de la información en estas distribuciones conjuntas se puede describir en términos de su primer y segundo momento (Cryer y Sik, 2008).

Una serie de tiempo se refiere a las observaciones de una variable que ocurren en una secuencia de tiempo. Se usa el símbolo  $Y_t$  para el valor numérico de una observación; el subíndice  $t$  indica el periodo de tiempo en que ocurre la observación, por lo tanto, una secuencia de  $n$  observaciones puede representarse de esta manera:  $Y_1, Y_2, \dots, Y_n$  (Pankratz, 1983).

Sea el proceso estocástico:

$$\dots, Y_{-1}, Y_0, Y_1, Y_2, \dots, Y_n, Y_{n+1}, \dots$$

Y la serie temporal:

$$Y_1, Y_2, \dots, Y_n$$

## 4.2 Proceso estacionario

Un proceso  $\{Y_t\}$  es estrictamente estacionario si la distribución conjunta de  $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$  es igual a la distribución conjunta de  $Y_{t_1-k}, Y_{t_2-k}, \dots, Y_{t_n-k}$  para cualquier punto en el tiempo  $t_1, t_2, \dots, t_n$  y para cualquier rezago  $k$ . Por lo tanto, si  $n = 1$  la distribución univariada de  $\{Y_t\}$  es igual a la distribución de  $Y_{t-k}$  para todo  $t$  y  $k$ , es decir, las  $Y$ 's (distribuciones marginales) estarán idénticamente distribuidas, además:

$E(Y_t) = E(Y_{t \pm k})$  y  $Var(Y_t) = Var(Y_{t \pm k})$ , para todo  $t$  y  $k$ , es decir, la media y la varianza son constantes a través del tiempo.

Un proceso estocástico es estacionario si el valor de la covarianza entre dos periodos depende sólo de la distancia o rezago entre estos dos periodos, y no del tiempo en el cual se calculó la covarianza (Gujarati y Porter, 2010).

El supuesto de estacionariedad es el más importante para hacer inferencia estadística sobre la estructura de un proceso estocástico. La idea básica de estacionariedad es que las leyes de probabilidad que gobiernan el comportamiento del proceso no cambian con el tiempo (Cryer y Sik, 2008).

## 4.3 Ruido blanco

Un ejemplo de un proceso estacionario es el denominado “ruido blanco” definido como una secuencia de variables aleatorias idéntica e independientemente distribuidas  $\{\varepsilon_t\}$  para  $t = 1, 2, \dots, n$ , con media 0 y varianza  $\sigma^2$  lo cual suele representarse como  $\{\varepsilon_t\} \sim \text{niid}(0, \sigma^2)$ . Esto implica que todas las variables tienen la misma varianza  $\sigma^2$  y  $\text{corr}(\varepsilon_i, \varepsilon_j) = 0$  para  $i \neq j$ . Cuando cada  $\varepsilon_t$  sigue una distribución Normal,  $\{\varepsilon_t\}$  se denomina un proceso de ruido blanco normal o gaussiano.

#### 4.4 Modelos AR ( $p$ )

Un proceso estocástico univariante estacionario ( $Y_t$ ) sigue un modelo autorregresivo de orden  $p$  de la forma:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t$$

Un modelo autorregresivo de orden (1), AR(1) sigue el siguiente modelo:

$$y_t = \phi_1 y_{t-1} + \varepsilon_t$$

donde  $\phi_1$  es un parámetro,  $\phi_1 < 1$  (condición de estacionariedad),  $(\varepsilon_t) \sim \text{IID}(0, \sigma^2)$ .

#### 4.5 Modelos MA ( $q$ )

Un modelo de media móvil de orden  $q$  se puede representar como:

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

Del mismo modo, un proceso estocástico univariante estacionario ( $Y_t$ ) sigue un modelo MA(1) (media móvil de orden 1) cuando:

$$Y_t = \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

donde  $\theta_1$  es un parámetro,  $\theta_1 < 1$  (condición de invertibilidad), y  $(\varepsilon_t) \sim \text{iid}(0, \sigma^2)$ .

#### 4.6 Modelos ARMA

Los modelos formados por una parte autorregresiva y una de promedio móvil forman el siguiente modelo general ARMA ( $p, q$ ):

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

Un modelo ARMA (1,1) tiene la siguiente forma:

$$Y_t = \phi_1 y_{t-1} + \varepsilon_t - \theta \varepsilon_{t-1}$$

#### 4.7 Procesos no estacionarios

Un proceso estocástico  $(Y_t)$  es no estacionario cuando las propiedades estadísticas<sup>9</sup> de al menos una secuencia finita  $Y_{t_1}, Y_{t_2}, \dots, Y_{t_n}$  ( $n \geq 1$ ) de componentes de  $(Y_t)$ , son diferentes de las de la secuencia  $Y_{t_1+h}, Y_{t_2+h}, \dots, Y_{t_n+h}$  para al menos un número entero  $h > 0$  (Alberto, 2007).

$$Y_t = Y_{t-1} + \varepsilon_t, \text{ donde } (\varepsilon_t) \sim \text{iid}(0, \sigma^2).$$

Un modelo no estacionario  $Y_t = \phi Y_{t-1} + \varepsilon_t$ , donde  $\phi = 1$  puede convertirse a estacionario si  $\nabla Y_t = \varepsilon_t$ , donde  $\nabla Y_t = Y_t - Y_{t-1}$ , es la primera diferencia de  $Y_t$

La diferencia de cualquier serie temporal estacionaria también es estacionaria. Sin embargo, la sobrediferenciación introduce correlaciones innecesarias en una serie, dificulta el proceso de modelado y requiere la estimación parámetros innecesariamente introducidos (Cryer y Sik, 2008).

Pueden aplicarse transformaciones para obtener series estacionarias como son: diferenciación simple o estacional para eliminar la tendencia o inducir estacionariedad en media. Asimismo, se puede realizar una transformación de box-Cox o logaritmo natural para inducir estacionariedad en varianza.

Un ejemplo de un proceso no estacionario es el modelo de caminata aleatoria, donde el valor de  $y_t$  es igual a su valor en el periodo anterior más un error aleatorio. La caminata aleatoria recuerda los choques por siempre; es decir, tiene memoria infinita. La suma  $\sum_{t=1}^n \varepsilon_t$  se conoce también como tendencia estocástica (Gujarati y Porter, 2010).

Por lo tanto, una caminata aleatoria puede representarse como:  $y_t = y_{t-1} + \varepsilon_t$   
Donde  $\varepsilon_t$  se comporta como ruido blanco.

---

<sup>9</sup> Un proceso aleatorio estacionario en sentido amplio o débil únicamente requiere que el primer y segundo momento no varíe con respecto al tiempo.

#### 4.8 Modelos ARIMA

Los modelos ARIMA (Autorregresive Integrate Moving Average) fueron popularizados por George Box y Gwilym Jenkins (1976 y Reinsel (1994) como procedimientos para entender y utilizar series de tiempo univariadas

Una serie de tiempo  $\{Y_t\}$  sigue un proceso integrado autorregresivo de promedio móvil si la  $d$ -ésima diferencia  $w_t = \nabla^d Y_t$  es un proceso estacionario ARMA  $(p, q)$ , es decir, si  $\{w_t\}$  sigue un proceso ARMA  $(p, q)$ , entonces  $\{Y_t\}$  es un proceso ARIMA  $(p, d, q)$  y puede representarse como sigue:

$$Y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

Si se obtiene la primera diferencia,  $d = 1$ , se obtiene:

$$Y_t - Y_{t-1} = \phi_1 (y_{t-1} - y_{t-2}) + \phi_2 (y_{t-2} - y_{t-3}) + \cdots + \phi_p (y_{t-p} - y_{t-p-1}) + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

$$W_t = \phi_1 W_{t-1} + \phi_2 W_{t-2} + \cdots + \phi_p W_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

Sí es necesario diferenciar una serie de tiempo  $d$  veces para hacerla estacionaria se dice que la serie de tiempo original es ARIMA  $(p, d, q)$ , donde  $p$  denota el número de términos autorregresivos y  $q$  el número de términos de promedios móviles (Cryer y Sik, 2008).

La metodología Box-Jenkins considera cuatro pasos para ajustar un modelo de series de tiempo.

En primer lugar, se debe identificar si la serie de tiempo analizada tiene raíz unitaria en cuyo caso debe realizarse un proceso de transformación para volverla estacionaria. Y posteriormente se deben identificar los valores  $p$ ,  $d$  y  $q$ . La función de autocorrelación (ACF) y la función de autocorrelación parcial (PACF) son una herramienta útil para la identificación de estos valores (Box *et al.*, 2016).

Posterior a la especificación de los valores de  $p$ ,  $d$  y  $q$ , se deben estimar los parámetros de los términos autorregresivos y de promedios móviles incluidos en el modelo (pueden estimarse mediante el método de mínimos cuadrados ordinarios o máxima verosimilitud). Para discriminar entre diferentes modelos propuestos se puede hacer uso de los criterios AIC, BIC, etc., (tomando el modelo con el mínimo valor en cada uno de estos estadísticos).

Como tercer paso se realiza un examen de diagnóstico para corroborar que los residuales estimados a partir del modelo propuesto son de ruido blanco, en cuyo caso se acepta el modelo. Finalmente, se pone a prueba el modelo para obtener predicciones.

#### 4.9 Funciones de autocorrelación y autocorrelación parcial

Las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) permiten identificar si la serie constituye un proceso estacionario y ayudan a identificar el modelo ARIMA que se ajusta al conjunto de datos observado.

La auto covarianza de orden  $k$ , ( $k > 0$ ) de un proceso  $Y_t$  estacionario se define como:

$$\gamma_k = cov [Y_t, Y_{t+k}] \text{ o bien, } \gamma_k = cov [Y_{t-k}, Y_t] \text{ para } (k \geq 0)$$

Es la covarianza entre cualquier par de componentes de  $(Y_t)$  separados entre sí por un intervalo temporal o retardo  $k > 0$ . Para  $k = 0$ ,  $\gamma_0 = cov [Y_t, Y_t] = Var(Y_t)$ .

Para un proceso estacionario la autocorrelación simple de orden  $k$ , ( $k > 0$ ) se define como:

$$\rho_k = \frac{cov [Y_t, Y_{t+k}]}{Var(Y_t)^{1/2}} = \frac{\gamma_k}{\gamma_0}, (k = 1, 2, \dots); \rho_0 = 1$$

Bajo ciertas condiciones  $\hat{\rho}_k \sim n^{-1/2}$ . A partir de la secuencia de valores de  $k$  ( $\gamma_k: k = 1, 2, \dots$ ) se genera la función de autocorrelación muestral (ACF) de la

serie temporal  $Y_1, Y_2, Y_3, \dots, Y_n$ . La ACF de una serie se muestra mediante un gráfico llamado correlograma.

En lugar de probar la significancia estadística de cualquier coeficiente de autocorrelación individual, podemos utilizar el estadístico Q desarrollado por Box y Pierce para probar la hipótesis conjunta de que todos los  $\rho_k$  hasta ciertos rezagos son simultáneamente iguales a cero.

$$Q_{BP} = n \sum_{k=1}^m \hat{\rho}_k^2$$

donde  $n$  es el tamaño de la muestra y  $m$  = longitud del rezago. Una variante del estadístico Q de Box-Pierce es el estadístico Ljung-Box:

$$Q_{LB} = n(n+2) \sum_{k=1}^m \frac{\hat{\rho}_k^2}{n-k} \sim \chi^2(m)$$

Bajo  $H_0 = \rho_1, \rho_2, \rho_3, \dots, \rho_k = 0$

Las autocorrelaciones parciales (PACF) se utilizan para medir el grado de asociación entre  $Y_t$  e  $Y_{t-k}$  cuando los efectos de otros rezagos en el tiempo 1, 2, 3, ...,  $k-1$ , son eliminados.

El coeficiente de autocorrelación parcial de orden  $k$  se denota por  $\alpha_k$  y se puede calcular haciendo la siguiente regresión:

$$y_t = \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \dots + \beta_k y_{t-k}$$

La autocorrelación parcial,  $\alpha_k$ , es el coeficiente estimado  $\beta_k$  de esta regresión múltiple. Usualmente la función de autocorrelación parcial o PACF se muestra mediante un gráfico.

Al igual que con el ACF, las autocorrelaciones parciales deberían ser todas cercanas a cero para una serie de ruido blanco. Quenouille (1949) demostró que si la serie temporal es ruido blanco, las autocorrelaciones parciales estimadas



son aproximadamente independientes, con una distribución normal y un error estándar  $\frac{1}{\sqrt{n}}$ . Por lo tanto, los mismos valores críticos de  $\pm 1.96/\sqrt{n}$  pueden usarse con la PACF para evaluar si los datos son ruido blanco (Makridakis y Wheelwright, 1997).

#### 4.10 Raíces unitarias

Una forma de comprobar la presencia de raíz unitaria en una serie es mediante un análisis gráfico (si el gráfico de los datos presenta tendencia puede ser un indicio de que la media y/o varianza no son constantes), la prueba del correlograma (para un proceso puramente de ruido blanco, las autocorrelaciones en distintos rezagos se ubican alrededor del cero) o alguna de las siguientes pruebas: prueba Dickey Fuller Aumentada (DFA) y Phillips Perrón (PP).

##### 4.10.1 Prueba Dickey Fuller

Sea  $Y_t$  el proceso estocástico autorregresivo de orden 1:

$$Y_t = \rho Y_{t-1} + \varepsilon_t$$

$$-1 \leq \rho \leq 1$$

Donde  $\varepsilon_t$  es un término de error de ruido blanco

sí  $\rho = 1$ , el proceso de raíz unitaria se convierte en un modelo de caminata aleatoria, el cual es un proceso estocástico no estacionario.

No se puede probar la hipótesis de que  $\rho = 1$  por medio de la prueba  $t$  acostumbrada, ya que esa prueba tiene un sesgo muy marcado en el caso de una raíz unitaria (Gujarati y Porter, 2010). Por lo tanto, se reescribe la ecuación anterior como:

$$\begin{aligned} Y_t - Y_{t-1} &= \rho Y_t - Y_{t-1} + \varepsilon_t \\ &= (\rho - 1)Y_t + \varepsilon_t \end{aligned}$$

Usando el operador de primeras diferencias:

$$\Delta Y_t = \delta Y_{t-1} + \varepsilon_t$$

Donde:  $\delta = (\rho - 1)$  y  $\Delta$  es el operador de primeras diferencias

Se prueba la hipótesis (nula) de que  $\delta = 0$ , y la hipótesis alternativa es que  $\delta < 0$ . Si  $\delta = 0$ , entonces  $\theta = 1$ , es decir, se tiene una raíz unitaria, lo cual significa que la serie de tiempo en consideración es no estacionaria.

Dickey y Fuller (1976) mostraron que dada la hipótesis nula:  $\delta = 0$  el valor estimado  $t$  del coeficiente de  $Y_{t-1}$  sigue el estadístico  $\tau$  (*tau*), por lo tanto, puede probarse, comparando el estadístico *tau* con los valores críticos en tablas de este estadístico o prueba *tau*.

Los valores críticos de estas tablas corresponden a tres posibles formas de estimación conforme a tres hipótesis nulas:

Modelo sin constante:  $\Delta Y_t = \delta Y_{t-1} + \varepsilon_t$

Modelo con constante:  $\Delta Y_t = \mu + \delta Y_{t-1} + \varepsilon_t$

Modelo con constante y tendencia:  $\Delta Y_t = \mu + \delta Y_{t-1} + \beta t + \varepsilon_t$

La prueba DFA consiste en tomar las primeras diferencias de  $Y_t$  y hacer la regresión sobre  $Y_{t-1}$ , con la finalidad de saber si el coeficiente estimado de la pendiente en esta regresión  $\hat{\delta}$  es cero, en cuyo caso se concluye que  $Y_t$  es no estacionaria; si es negativa, se infiere que  $Y_t$  es estacionaria.

La prueba (DFA) puede aplicarse siguiendo el siguiente modelo

$$\Delta y_t = \beta_1 + \beta_2 t + \delta y_{t-1} + \sum_{i=1}^n \alpha_i \Delta y_{t-i} + \varepsilon_t$$

Donde:  $\varepsilon_t$  es un término de error puro de ruido blanco

#### 4.10.2 Prueba de Phillips-Perron (PP)

La prueba de Dickey Fuller (DF) contempla el supuesto de que los términos de error  $\varepsilon_t$  están idéntica e independientemente distribuidos. La prueba de Dickey Fuller Aumentada (DFA) ajusta la prueba DF y considera una posible correlación serial en los términos de error. Phillips y Perron utilizan estadística no paramétrica<sup>10</sup> para evitar la correlación serial en los términos de error (Gujarati y Porter, 2010).

#### 4.11 Modelos vectoriales autorregresivos (VAR)

El modelo VAR desarrollado por Sims (1980 y 1986) se fundamentan en una estructura vectorial que supone una interdependencia entre las variables del modelo y sus rezagos ((Sims, 1980), citado por (Morán, 2014), de esta forma, un modelo VAR( $p$ ) es un sistema en el que todas las variables son consideradas endógenas y cada una es expresada como una combinación lineal de  $p$  de sus propios valores rezagados y de  $p$  valores rezagados de las otras variables en el modelo.

El valor de  $p$ , es definido por la longitud del rezago incluido en el modelo y éste puede determinarse mediante criterios como el de Akaike o el de Schwarz, seleccionando el modelo que proporcione los valores más bajos de dichos criterios. No obstante, la inclusión de muchos términos rezagados consumirá muchos grados de libertad, y existe la posibilidad de presencia de multicolinealidad. Por otra parte, la inclusión de pocos rezagos provoca errores de especificación (Gujarati y Porter, 2010).

---

<sup>10</sup> Los métodos no paramétricos suelen requerir suposiciones menos restrictivas acerca del nivel de medición de los datos y menos suposiciones acerca de la forma de las distribuciones de probabilidad generadas por los datos muestrales en comparación con los métodos paramétricos (Anderson *et al.*, 2008) .

Dos series de tiempo,  $\{y_t\}$  y  $\{z_t\}$ , siguen un proceso de vector autorregresivo de orden 1, denotado VAR (1) si:

$$y_t = \phi_{11}y_{t-1} + \phi_{12}z_{t-1} + \varepsilon_{yt}$$

$$z_t = \phi_{21}y_{t-1} + \phi_{22}z_{t-1} + \varepsilon_{zt}$$

Donde  $y_t$  y  $z_t$  son estacionarias<sup>11</sup>, además  $\varepsilon_{yt}$  y  $\varepsilon_{zt}$  son perturbaciones de ruido blanco no correlacionadas, con desviaciones estándar de  $\sigma_y$  y  $\sigma_z$ .

De forma matricial:

$$\begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} \\ z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

$$X_t = \Gamma_1 X_{t-1} + \varepsilon_t$$

Donde:

$$x_t = \begin{bmatrix} y_t \\ z_t \end{bmatrix}, \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}, \varepsilon_t = \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

El proceso se puede escribir como<sup>12</sup>:

$$\left( \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} [B] \right) \begin{bmatrix} y_t \\ z_t \end{bmatrix} = \begin{bmatrix} \varepsilon_{yt} \\ \varepsilon_{zt} \end{bmatrix}$$

Donde B es un operador de retardo que aplicado a vectores se define como:

$$Bx_t = x_{t-1}$$

De forma general un modelo con  $k$  valores rezagados de las variables  $Y_t$  y  $z_t$

$$Y_t = \alpha + \sum_{j=1}^k \beta_j y_{t-j} + \sum_{j=1}^k \gamma_j z_{t-j} + \varepsilon_{1t}$$

$$z_t = \alpha' + \sum_{j=1}^k \theta_j y_{t-j} + \sum_{j=1}^k \gamma_j z_{t-j} + \varepsilon_{2t}$$

<sup>11</sup> Si las variables son estacionarias es correcto estimar un modelo VAR, por otra parte, si las variables son no estacionarias cointegradas es posible estimar un modelo VEC.

<sup>12</sup> (Peña, 2010)

Si  $y_t$  y  $z_t$  son no estacionarias  $I(1)$  y no cointegradas, entonces es posible estimar un modelo VAR en primeras diferencias:

$$y_t - y_{t-1} = \phi_{11}(y_{t-1} - y_{t-2}) + \phi_{12}(z_{t-1} - z_{t-2}) + \varepsilon_{\Delta y t}$$

$$z_t - z_{t-1} = \phi_{21}(y_{t-1} - y_{t-2}) + \phi_{22}(z_{t-1} - z_{t-2}) + \varepsilon_{\Delta z t}$$

Es decir:

$$\Delta y_t = \phi_{11}\Delta y_{t-1} + \phi_{12}\Delta z_{t-1} + \varepsilon_{\Delta y t}$$

$$\Delta z_t = \phi_{21}\Delta y_{t-1} + \phi_{22}\Delta z_{t-1} + \varepsilon_{\Delta z t}$$

En forma matricial

$$\begin{bmatrix} y_t - y_{t-1} \\ z_t - z_{t-1} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} y_{t-1} - y_{t-2} \\ z_{t-1} - z_{t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\Delta y t} \\ \varepsilon_{\Delta z t} \end{bmatrix}$$

$$\begin{bmatrix} \Delta y_t \\ \Delta z_t \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ \Delta z_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{\Delta y t} \\ \varepsilon_{\Delta z t} \end{bmatrix}$$

En una generalización multivariada un proceso vectorial autorregresivo puede representarse como:

$$y_t = A_0 + A_1 y_{t-1} + A_2 y_{t-2} + \cdots + A_p y_{t-p} + \varepsilon_t$$

Donde:

$y_t$  es un vector de dimensión  $(n \times 1)$  que contiene cada una de las variables incluidas en el VAR

$A_0$  es vector de términos de intercepción de dimensión  $(n \times 1)$

$A_i = (n \times n)$  matrices de coeficientes

$\varepsilon_t$  es un vector de perturbaciones de ruido blanco de dimensión  $(n \times 1)$

Las ecuaciones en el sistema se pueden estimar usando MCO y dichas estimaciones son consistentes y asintóticamente eficientes.

Un modelo VAR( $p$ ) con  $n$  ecuaciones puede ser representado como:

$$\begin{bmatrix} A_{1t} \\ \vdots \\ A_{nt} \end{bmatrix} = \begin{bmatrix} A_{10} \\ \vdots \\ A_{n0} \end{bmatrix} + \begin{pmatrix} A_{11}(L) & A_{12}(L) \cdots & A_{1n}(L) \\ \vdots & \ddots & \vdots \\ A_{n1}(L) & A_{n2}(L) \cdots & A_{nn}(L) \end{pmatrix} \begin{bmatrix} y_{1t-1} \\ \vdots \\ y_{nt-p} \end{bmatrix} + \begin{pmatrix} A_{11}(L) & A_{12}(L) \cdots & A_{1n}(L) \\ \vdots & \ddots & \vdots \\ A_{n1}(L) & A_{n2}(L) \cdots & A_{nn}(L) \end{pmatrix} \begin{bmatrix} y_{1t-p} \\ \vdots \\ y_{nt-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{nt} \end{bmatrix}$$

$A_{it}$ : es el vector de variables incluidas en el modelo

$A_{i0}$  = son los interceptos

$A_{ij}(L)$  = son los polinomios en el operador de rezago  $L$

Los coeficientes individuales de  $A_{ij}(L)$  se denotan por  $a_{ij}(1), a_{ij}(2), \dots$

Como todas las ecuaciones tienen la misma longitud de retraso, los polinomios  $A_{ij}(L)$  son todos del mismo grado.

Los términos  $\varepsilon_{it}$  son perturbaciones de ruido blanco que pueden estar correlacionadas entre sí.

#### 4.12 Series cointegradas

Las series  $I(d)$  integradas  $y_t$  y  $z_t$  están cointegradas si existe una combinación lineal entre ellas que es de orden de integración menor que  $d$  (Peña, 2010). Es decir, a pesar de no ser estacionarias en un nivel individual, una combinación lineal de dos o más series de tiempo puede ser estacionaria, por lo tanto, existe una relación de largo plazo, o de equilibrio, entre ambas (Gujarati y Porter, 2010). Si la variable  $z_t$  es capaz de explicar totalmente el comportamiento no estacionario de  $y_t$ , se dice entonces que las variables están cointegradas (Peña, 2010).

Engle y Granger (1987) propusieron un procedimiento para determinar si dos variables  $I(1)$  están cointegradas de orden CI (1,1). En primer lugar, se debe probar el orden de integración de las variables<sup>13</sup> y posteriormente se estima la relación de equilibrio a largo plazo en la forma:

$$Y_t = \beta_0 + \beta_1 z_t + \varepsilon_t$$

La regresión anterior se denomina regresión cointegrante, y el parámetro de pendiente  $\beta_1$  parámetro cointegrante. Teóricamente, es posible que existan relaciones no lineales a largo plazo entre un conjunto de variables integradas. También se debe tener en cuenta que el vector de cointegración no es único (Enders, 2015), por lo que en una aplicación se escoge una normalización.

Se debe verificar que los residuos de la regresión anterior son  $I(0)$ , es decir, estacionarios y se aproximan al ruido blanco<sup>14</sup>. Si esto sucede, se puede afirmar que las series estudiadas están cointegradas (Gujarati y Porter, 2010).

La cointegración por el método de Johansen se prueba al contrastar que los estadísticos de la traza y los eigenvalores máximos sean mayores a los valores críticos. La hipótesis nula es que hay como máximo  $r$  vectores de cointegración contra la alternativa de que hay  $m$ , donde  $r \leq m$ .

El contraste de hipótesis consiste en el siguiente procedimiento:

- 1) se prueba la hipótesis nula  $H_0: r = 0$  (no existe cointegración), Vs  $H_a: r = 1$  (existe una relación de cointegración).
- 2) Si se rechaza la hipótesis anterior a partir de los estadísticos mencionados (traza y los eigenvalores máximos), se contrastan las siguientes hipótesis:  $H_0: r = 1$  Vs  $H_a: r = 2$ , y así sucesivamente hasta no rechazar  $H_0$ , o bien hasta que se tenga que aceptar la hipótesis alternativa de  $r = m - 1$ .

---

<sup>13</sup> Por definición, la cointegración requiere que se integren dos variables del mismo orden.

<sup>14</sup> Se puede verificar mediante la prueba DFA aplicada a los residuos de la regresión

#### 4.13 Modelos vectoriales de corrección de error (VEC)

Un modelo de corrección de errores para las variables  $I(1)$  necesariamente implica cointegración. Esto está determinado por el teorema de representación de Granger (1980). Si existe una relación de cointegración entre las variables, es posible ajustar un modelo a largo plazo y la dinámica a corto plazo simultáneamente.

Sean dos variables  $y_t$   $I(1)$  y  $z_t$   $I(1)$  no estacionarias integradas de orden 1, cointegradas mediante la ecuación:

$$Y_t = \beta_1 z_t + \varepsilon_t$$

Donde los errores estimados son:

$$\hat{\varepsilon}_t = Y_t - \hat{\beta}_1 z_t$$

Tal que:

$$\hat{\varepsilon}_t \sim I(0)$$

Y los cambios en  $Y_t$  y  $z_t$  dependen de las desviaciones de este equilibrio en el periodo  $t - 1$ .

$$\Delta y_t = \alpha_1(y_{t-1} - \beta_1 z_{t-1}) + u_{1t}$$

$$\Delta z_t = \alpha_2(y_{t-1} - \beta_1 z_{t-1}) + u_{2t}$$

De forma general, en un modelo de corrección de error,  $\Delta y_t$  y  $\Delta z_t$ , dependen de cambios previos en ambas variables de la siguiente forma:

$$\Delta y_t = \alpha_1(y_{t-1} - \beta_1 z_{t-1}) + \gamma_{11}\Delta y_{t-1} + \gamma_{12}\Delta z_{t-1} + u_{1t}$$

$$\Delta z_t = \alpha_2(y_{t-1} - \beta_1 z_{t-1}) + \gamma_{21}\Delta y_{t-1} + \gamma_{22}\Delta z_{t-1} + u_{2t}$$

Así, el modelo VEC se puede escribir en forma matricial como:

$$X_t - X_{t-1} = \alpha\beta'X_{t-1} + \Gamma(X_{t-1} - X_{t-2}) + u_t$$

Donde:

$$X_t = [Y_t, Z_t]' \quad \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix} \quad \beta' = [1, -\beta_1] \quad \Gamma = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \quad u_t = [u_{1t}, u_{2t}]' \quad \Pi = \alpha\beta'$$



De forma equivalente:

$$\Delta X_t = \Pi X_{t-1} + \sum_{i=1}^{p-1} \Gamma_i \Delta X_{t-i}$$

Donde:

$$\Pi = \alpha\beta'$$

Un sistema cointegrado puede verse como una forma restringida de un modelo VAR general.

sea  $X_t = [Y_t, Z_t]'$  y  $\varepsilon_t = (\varepsilon_{y_t}, \varepsilon_{z_t})'$ , entonces:

$$\Delta X_t = \Pi X_{t-1} + \varepsilon_t$$

Si se estima un VAR de variables cointegradas utilizando solo las primeras diferencias, sin la expresión  $\Pi X_{t-1}$ , esto eliminaría la porción de corrección de errores del modelo. El análisis de cointegración indica que la estimación de un modelo un VAR en primeras diferencias, estará mal especificado, cuando hay algunos vectores cointegrantes presentes en la serie  $I(1)$ , por consiguiente, las estimaciones de los parámetros son sesgadas debido a la omisión de variables (Cuthbertson, 2002).

El proceso de construcción de modelos multivariados se puede resumir como sigue:

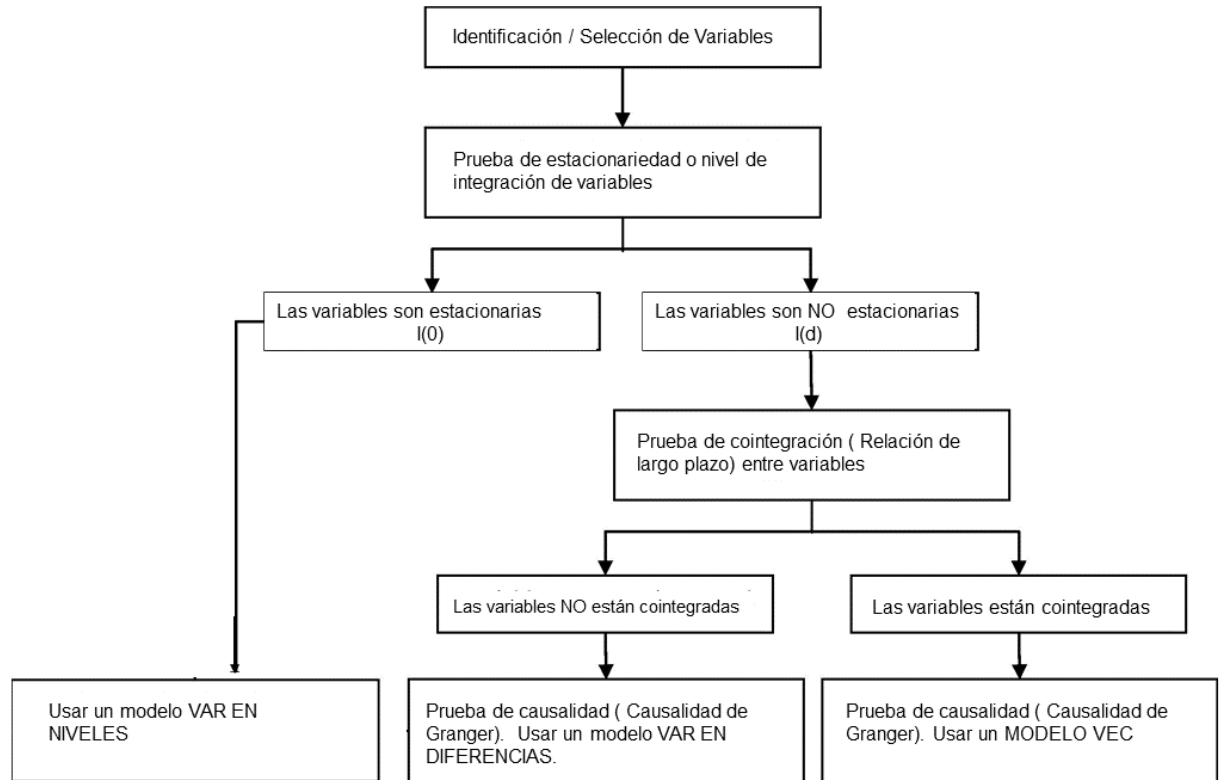


Figura 9. Proceso para el análisis de series de tiempo

Fuente: Modificado de Maparu y Mazumder (2017).

#### 4.14 Regresión espuria

La regresión de una serie no estacionaria sobre otra no estacionaria puede conducir a una regresión espuria, sin embargo, si las dos series comparten una tendencia común la regresión podría no ser espuria y los errores resultantes de dicha regresión son estacionarios. Una prueba para la cointegración puede considerarse como una preprueba para evitar las situaciones de regresiones espurias (Granger, 1986).

#### 4.15 Discriminación de los modelos

Algunos criterios como el Criterio de Información de Akaike (1973) AIC y Criterio de Información Bayesiano de Schwarz (BIC) pueden ser empleados para discriminar entre varios modelos ARIMA propuestos para la misma serie, seleccionando los modelos con menor AIC y BIC. El criterio de información de Akaike se define como:

$$AIC = -2L + 2k$$

Donde:

$k = p + q + 1$  si el modelo contiene un intercepto o término constante y  $k = p + q$  en caso contrario<sup>15</sup>.

Y el criterio de información bayesiano de Schwarz se define como:

$$BIC = -2L + k \log(n)$$

En ambos criterios,  $L$  es la *log-verosimilitud* asociada con el modelo, evaluada en las estimaciones obtenidas de los parámetros,  $n$  es el tamaño de muestra y  $k$  es el número de parámetros.

Si el proceso verdadero sigue un modelo ARMA  $(p, q)$ , entonces se sabe que los valores de  $(p, q)$  especificados al minimizar el BIC son consistentes (Cryer y Sik, 2008).

---

<sup>15</sup> Los términos  $2(p + q + 1)$  o  $2(p + q)$  sirven como un factor de penalización y ayudan en la construcción de modelos parsimoniosos

#### 4.16 Verificación del modelo: análisis de residuales

Si el modelo se especifica correctamente y las estimaciones de los parámetros son razonablemente cercanas a los valores verdaderos, entonces los residuos deberían tener casi las propiedades del ruido blanco, normales e independientes distribuidas idénticamente, con medias cero y desviaciones estándar comunes (Cryer y Sik, 2008).

La verificación del modelo adecuado puede realizarse mediante las siguientes formas: inspección del gráfico de residuales, evaluación de normalidad mediante gráficos cuantil-cuantil o alguna prueba de normalidad como Shapiro-Wilk. El estadístico Box y Pierce (1970) permite probar la significancia conjunta de las autocorrelaciones entre residuos. En la inspección del gráfico de los residuales a través del tiempo, se espera que la gráfica muestre una dispersión dentro de una banda establecida, alrededor de cero, sin tendencias evidentes. En las gráficas cuantil-cuantil de los residuos, los puntos deben parecer seguir una línea recta.

#### 4.17 Pronóstico

Dada la serie de tiempo  $Y_1, Y_2, \dots, Y_t$  y bajo el supuesto de que el modelo ha sido correctamente especificado y los parámetros verdaderos para cada serie son conocidos, es posible pronosticar el valor de  $Y_{t+k}$  que ocurrirá en  $k$  unidades de tiempo en el futuro. Se puede denotar el pronóstico como:

$$\hat{y}_t(k) = E(y_{t+k} | y_1, y_2, \dots, y_t)$$

La observación futura  $y_{t+k}$  estará contenida dentro de los límites de predicción con  $(1 - \alpha) \%$  de confiabilidad:

$$\hat{y}_t(k) \pm z_{1-\alpha/2} \sqrt{\text{var}(\varepsilon_t(k))}$$

El supuesto habitual para construir intervalos de predicción es que los errores de pronóstico tienen una distribución normal con media cero (Makridakis y Wheelwright, 1997). Bajo este supuesto, un intervalo de predicción aproximado para la siguiente observación es:

$$\hat{y}_{n+1} \pm z_{1-\alpha/2} \sqrt{MSE}$$

Donde:  $\hat{y}_{n+1}$ : es el valor predicho en el periodo  $n + 1$  y  $MSE$ : es el error cuadrático medio.

Después de identificar los parámetros del modelo y estimarlos es posible determinar predicciones puntuales mediante el siguiente procedimiento en un modelo MA ( $q$ ):

Dados los valores observados  $y_1, y_2, \dots, y_{t-1}$  es posible utilizar estos valores para obtener una predicción puntual  $\hat{y}_t$  de valor  $y_t$  mediante el siguiente cálculo:

$$y_t = y_{t-1} + \epsilon_t - \theta_1 \epsilon_{t-1}$$

$$\hat{y}_t = y_{t-1} + \hat{\epsilon}_t - \hat{\theta}_1 \hat{\epsilon}_{t-1}$$

Para determinar la predicción puntual del valor futuro  $y_t$  se calcula la perturbación aleatoria:

$$\hat{\epsilon}_t = y_t - \hat{y}_t \quad \text{y} \quad \hat{\epsilon}_{t-1} = y_{t-1} - \hat{y}_{t-1}$$

Donde:  $\hat{\epsilon}_t$ : es el residuo  $t$ -ésimo

La predicción puntual  $\hat{\epsilon}_t$  de la perturbación o choque futuros  $\epsilon_t$  es cero y la predicción puntual  $\hat{\epsilon}_{t-1}$  de la perturbación aleatoria pasada  $\epsilon_{t-1}$  es el residuo  $(t - 1)$  -ésimo. En caso de no poder estimarse  $\hat{y}_{t-1}$  se considera igual a cero (Bowerman *et al.*, 2007). Siguiendo un procedimiento similar es posible construir pronósticos para un modelo AR ( $p$ ).

#### 4.18 Criterios de evaluación de la capacidad predictiva de los modelos

La capacidad predictiva de los modelos puede ser medida utilizando estadísticas como: el error promedio de pronóstico (ME), el error promedio absoluto de pronóstico (MAE), el error cuadrático medio del pronóstico (MSE), la raíz del error cuadrático medio (RMSE, del inglés Root Mean Squared Error), el porcentaje medio de error (MPE), el error porcentual absoluto medio (MAPE, del inglés Mean Absolute Percentage Error) y la U de Theil (Makridakis, 1997).

Sea  $\hat{y}_1, \hat{y}_2, \hat{y}_3, \dots, \hat{y}_n$  una secuencia de  $n$  pronósticos:

El error de pronóstico se define como:

$$e = y_i - \hat{y}_i$$

Por lo tanto, los errores del pronóstico serán:  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$

Y el error cuadrático medio (ECM) es:

$$\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}$$

La raíz del error cuadrático medio (RECM) asociado con dicha secuencia es:

$$RECM = \sqrt{1/n \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

Es decir, es la desviación típica muestral de los errores del pronóstico  $\hat{e}_1, \hat{e}_2, \dots, \hat{e}_n$ .

Es una medida que sólo toma valores positivos; cuando esta medida tiende a cero indica que los pronósticos son mejores.

El error medio absoluto (MAE, del inglés Mean Absolute Error) asociado con dicha secuencia es:

$$MAE = \frac{1}{n} \sum_{i=1}^n |(y_i - \hat{y}_i)| = \frac{1}{n} \sum_{i=1}^n |\widehat{e}_n|$$

es decir, la media de los valores absolutos de los errores del pronóstico.

El error porcentual absoluto medio (MAPE, del inglés Mean Absolute Percentage Error) se define como:

$$MAPE = \frac{100}{n} \sum_{i=1}^n \frac{|(y_i - \hat{y}_i)|}{y_i}$$

La U de Theil permite una comparación relativa de los métodos de pronóstico formales con métodos “ingenuos” y también pondera los errores involucrados para que los errores grandes tengan mucho más peso que los errores pequeños (Makridakis *et al.*, 1997).

La U de Theil se define como:

$$CDT = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i - y_i)^2}}{\sqrt{\frac{1}{n} \sum_{i=1}^n (\hat{y}_i)^2} + \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i)^2}}$$

Sí el coeficiente de Theil es igual 0, significa que el pronóstico es perfecto (exacto). Sí el coeficiente de Theil es igual a 1, indica que la precisión del método de pronóstico bajo estudio es igual a la del método simplista<sup>16</sup>, que asigna el valor del pronóstico igual al valor de  $y_t$  por lo que no se justifica implementar el método bajo estudio. Si el coeficiente de Theil es mayor que 1, significa que el modelo no es útil para fines predictivos. Sí  $0 < U \text{ de Theil} < 1$  el método usado para pronosticar es superior al método simple. Cabe destacar que la U de Theil es una medida de precisión relativa que toma valores no negativos.

---

<sup>16</sup> En (Makridakis *et al.*, 1997) se expone este método denominado “ingenuo” o simplista, que utiliza la observación más reciente disponible como pronóstico.

## 5. MATERIALES Y MÉTODOS

Con base en los conceptos teóricos de los modelos ARIMA, VAR y VEC expuestos anteriormente, se construyeron modelos para las series que a continuación se describen.

### 5.1 Modelos ARIMA para el precio de maíz

Las series de tiempo analizadas corresponden a los precios promedio mensuales en los estados con mayor participación porcentual en la producción nacional y con mayor valor (\$) de la producción, estos estados son: México, Michoacán Jalisco y Sinaloa. Los datos fueron obtenidos del portal del Sistema Nacional de Información de Mercados (SIIM). El periodo de análisis comprende desde enero del año 2000 a diciembre del año 2018, teniendo 228 observaciones en cada serie.

Dado que los valores futuros de las series son desconocidos, se puede retener una parte de las observaciones del proceso de estimación y estimar modelos alternativos sobre un conjunto de datos reducido y usar estas estimaciones para pronosticar las observaciones del período en retención. Luego se pueden comparar las propiedades de los errores de pronóstico de los modelos estimados (Enders, 2015).

En la práctica se han usado particiones del conjunto de datos dejando uno fuera de la estimación y posteriormente realizando el pronóstico en Sánchez *et al.* (2013). En este caso, se realizó una partición del conjunto de datos de cada serie, en dos subconjuntos. El primer subconjunto corresponde al conjunto de “entrenamiento” y consta de 222 observaciones; éste fue usado para ajustar un modelo ARIMA a cada una de las series. El segundo subconjunto corresponde al conjunto de “validación”, consta de las 6 las últimas observaciones de cada serie y fue utilizado para evaluar la capacidad predictiva de los modelos propuestos.



Nótese que la metodología Box-Jenkins se aplica a series estacionarias, antes de ajustar un modelo ARIMA es necesario verificar las tres condiciones de estacionariedad en cada una de las series: media y varianza finitas y constantes respecto al tiempo y, covarianza finita que dependa del tiempo en la definición de proceso autorregresivo.

Se aplicaron dos pruebas para verificar la presencia de raíces unitarias: Dickey Fuller Aumentada (DFA) y Phillips-Perron (PP). Las series que resultaron no estacionarias fueron transformadas con la finalidad de inducir estacionariedad. Se obtuvo el logaritmo natural para estabilizar la varianza y se diferenciaron las series para estabilizar la media alrededor de cero. Posteriormente se aplicó la metodología de series de tiempo ARIMA para ajustar modelos a cada serie analizada.

La estimación de los parámetros de los modelos ARIMA para las cuatro series analizadas se realizó mediante MCO y posteriormente se hizo un análisis de residuales para verificar que éstos presenten un comportamiento similar a ruido blanco.

## **5.2 Modelos VAR Y VEC**

La hipótesis probada en la construcción de los modelos VAR fue que los precios de maíz blanco en el mercado nacional están influenciados por los precios en el mercado internacional, en cuyo caso, los modelos VAR y VEC proporcionan un mejor predictor en comparación con el predictor propuesto por los modelos ARIMA.

Una hipótesis similar fue propuesta por Lederman (2004) quien afirmó que “los precios de maíz mexicanos y estadounidenses están cointegrados”; en ese mismo sentido Araujo (2011) argumentó que “existe integración entre los mercados de maíz en México y los Estados Unidos, y dicha integración varía entre los distintos mercados regionales en México”, además agregó que “es

posible observar que las regiones con una mayor grado de autosuficiencia presentarán un menor grado de integración con el mercado estadounidense”.

Por otra parte, Godínez (2006) y (Godínez y Fuentes , 2008) probaron la hipótesis de que el precio futuro del maíz amarillo n°2 de la Bolsa de Chicago no mantiene una causalidad y liderazgo sobre los precios físicos de maíz blanco en México.

Resultados similares fueron obtenidos por Ortiz y Montiel (2017) quienes mostraron mediante un análisis de volatilidad estocástica multivariante que el precio de mercado de futuros de maíz no se encuentra fuertemente relacionado con los precios registrados en algunos estados del país, e infiere que la cobertura del programa ASERCA no cumple adecuadamente con su propósito de proteger a los agricultores nacionales que siembran maíz blanco.

En la construcción de modelos VAR y VEC se emplearon los precios de futuros<sup>17</sup> del maíz amarillo y los precios de maíz en físico<sup>18</sup> (FOB). Los precios fueron transformados de dólares por bushel<sup>19</sup> a pesos por Kg usando el tipo de cambio FIX<sup>20</sup>. Los datos del tipo de cambio se obtuvieron del portal del Banco de México (BANXICO) y los precios de maíz en físico y precios de futuros se obtuvieron del sitio de internet del Departamento de Agricultura de Estados Unidos (USDA). El periodo de análisis fue de enero del año 2000 a diciembre de 2018.

Variables similares a las utilizadas en este estudio fueron propuestas en Godínez (2006) quien utilizó series semanales de precios al mayoreo del maíz blanco en 15 centrales de abasto en México y el precio futuro semanal del maíz amarillo US #2 de la Bolsa de Chicago.

Zavaleta (2019) utilizó las cotizaciones del maíz amarillo en el mercado spot, del Chicago Mercantile Exchange, las cotizaciones semanales del maíz blanco en el

---

<sup>17</sup> Precio de liquidación mensual promedio del contrato de futuros de maíz más cercano

<sup>18</sup> Precios promedio mensuales de maíz amarillo No. 2, FOB en el Golfo de México

<sup>19</sup> Equivalencia: 1bushel de maíz = 25.4 kg

<sup>20</sup> Tipo de cambio pesos por dólar E.U.A., para solventar obligaciones denominadas en moneda extranjera

mercado spot que provee la Bolsa Mexicana de Valores a través de SIBOLSA en la construcción de modelos VAR. Araujo (2011) utilizó los precios de EU de exportación “franco a bordo” para el Golfo de Luisiana en dólares americanos y seis series de precios de maíz en mercados locales en México. De forma similar Ortiz y Montiel (2017) hicieron uso de los precios de los contratos de futuros del maíz amarillo #2 del Chicago Board of Trade (CBOT) con vencimiento en marzo y precios spot de los estados con altos niveles de producción de maíz en México en los que ASERCA brinda apoyos directos a coberturas de precios para maíz.

Se realizó una partición análoga a la aplicada a los precios nacionales. Se obtuvieron dos subconjuntos: un conjunto de “entrenamiento” compuesto de 222 observaciones y un conjunto de “validación”, compuesto de las 6 últimas observaciones de cada serie.

Como primer paso se verificó el orden de integración de las series: precio de maíz en físico y precio de futuros. Si las series son integradas se evaluó la pertinencia de establecer un modelo VEC verificando previamente la existencia de una relación de cointegración mediante el contraste de Johansen. Si las series son estacionarias se establecieron modelos VAR.

Una vez contruidos los modelos, fue posible generar predicciones y finalmente se puso a prueba el desempeño de los modelos propuestos.

Algunas medidas que han sido usadas en la práctica para verificar la capacidad predictiva de los modelos son: el error cuadrático medio del pronóstico (MSE), el error promedio de pronóstico (ME), el error promedio absoluto de pronóstico (MAE), el porcentaje de error al estimar, el porcentaje medio de error (MPE) y el error porcentual absoluto medio (MAPE) en Botero y Cano (2008) y Contreras *et al.* (2016), la U de Theil en Botero y Cano (2008), Guerrero (2007), Lanteri (2010), Sánchez *et al.* (2013) y Barreras *et al.* (2013), la raíz del error cuadrático medio (RMSE), en Arias y Torres (2004), Guerrero (2007), Lanteri (2010), Sánchez *et al.* (2013), y Barreras *et al.* (2013). Otras medidas para probar la bondad del ajuste en sus modelos son la desviación media absoluta (MAD) en

Velásquez y Franco (2010) y Contreras *et al.* (2016), la proporción de sesgo (PS), la proporción de la varianza (PV) y la proporción de covarianza (PCV) en Barreras S. *et al.* (2013) y Sánchez *et al.* (2013).

En esta investigación se utilizaron los siguientes criterios de comparación: Raíz de Error cuadrático Medio (RECM), el error porcentual absoluto medio (MAPE) y la U de Theil, como medidas para discriminar entre los modelos propuestos.

La estimación de los modelos, así como las pruebas de hipótesis y gráficos fueron realizados usando el programa R versión 3.6.1 y el programa SAS Statistical Analysis System.

### **5.3 Predicción fuera de la muestra**

El rango de datos utilizados para la estimación fue de enero del año 2000 a julio de 2018, denominado anteriormente como “conjunto de entrenamiento” o muestra. La predicción fuera de la muestra consiste en estimar los datos que han sido reservados como “conjunto de validación” correspondientes a periodo de julio de 2018 a diciembre de 2018. De esta forma es posible medir la capacidad predictiva de los modelos propuestos con base en los criterios antes mencionados.

## **6. RESULTADOS**

En las primeras tres secciones de este capítulo se muestran los resultados del análisis de las series de datos nacionales: México, Michoacán, Jalisco y Sinaloa, la construcción de modelos Autorregresivos Integrados de Promedio Móvil (ARIMA) y los pronósticos puntuales realizados con dichos modelos. Posteriormente con base en la metodología de modelos multivariados se construyeron modelos Vectoriales de Corrección de Error (VEC) utilizando las series de datos de México y los datos del precio de maíz en físico (FOB) de Estados Unidos; los resultados del análisis, construcción y pronósticos de los modelos VEC se muestran en las secciones cuatro y cinco de este capítulo. Finalmente, en la sección seis se muestran los resultados correspondientes a los modelos Vectoriales Autorregresivos (VAR).

### **6.1 Análisis de datos de series univariadas**

El comportamiento de las series de México, Michoacán y Sinaloa sugiere que éstas no cumplen con las condiciones de estacionariedad, es decir, media y varianzas constantes, debido a la presencia de tendencia a través del tiempo y periodos de volatilidad en los años posteriores a 2010 (Figura 10) y las series de precios transformadas en logaritmos tampoco muestran medias y varianzas constantes (Figura 11). Se obtuvo la primera diferencia del logaritmo de las series para inducir estacionariedad en media y varianza (Figura 12).

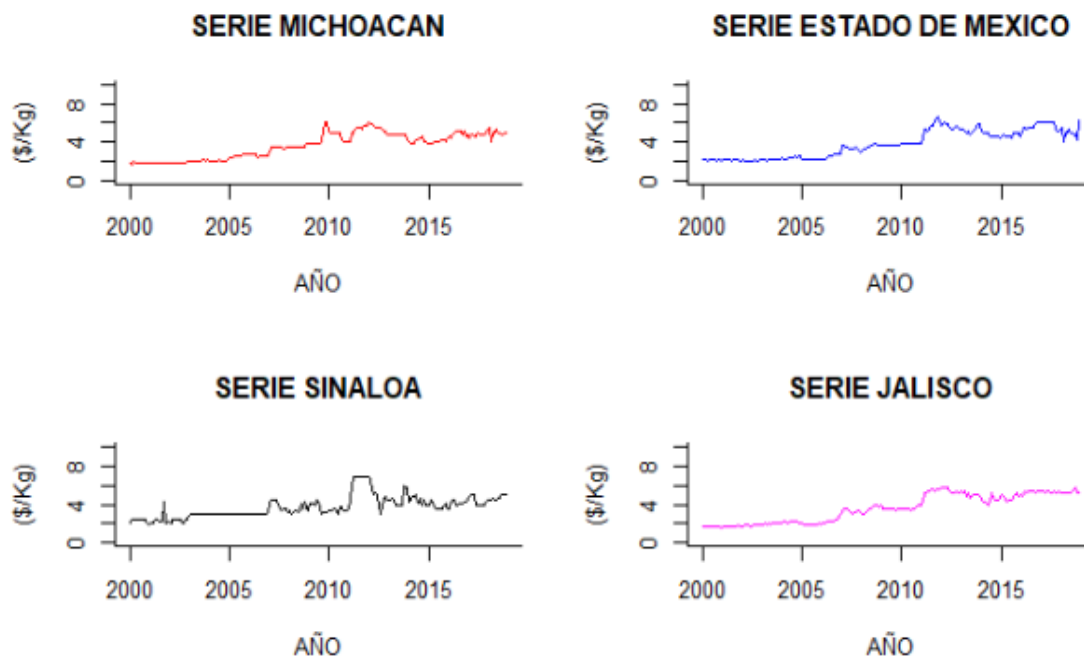


Figura 10. Series de tiempo del precio de maíz blanco en México.

Fuente: Elaboración propia con datos de SNIIM \*Series en niveles

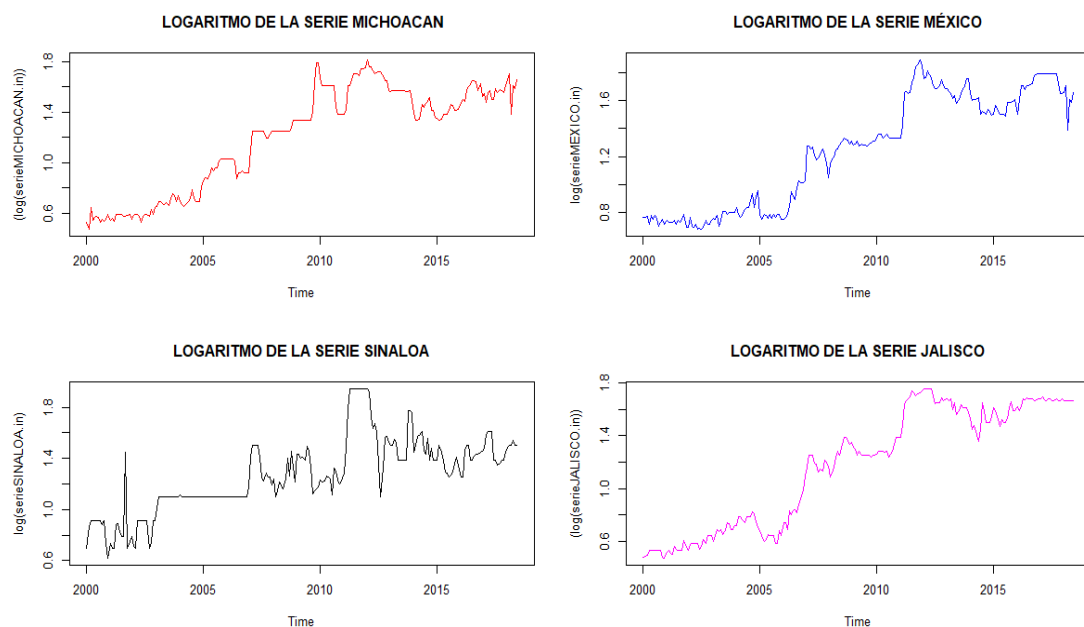


Figura 11. Logaritmo de las series de precios de maíz blanco en México

Fuente: Elaboración propia con datos de SNIIM

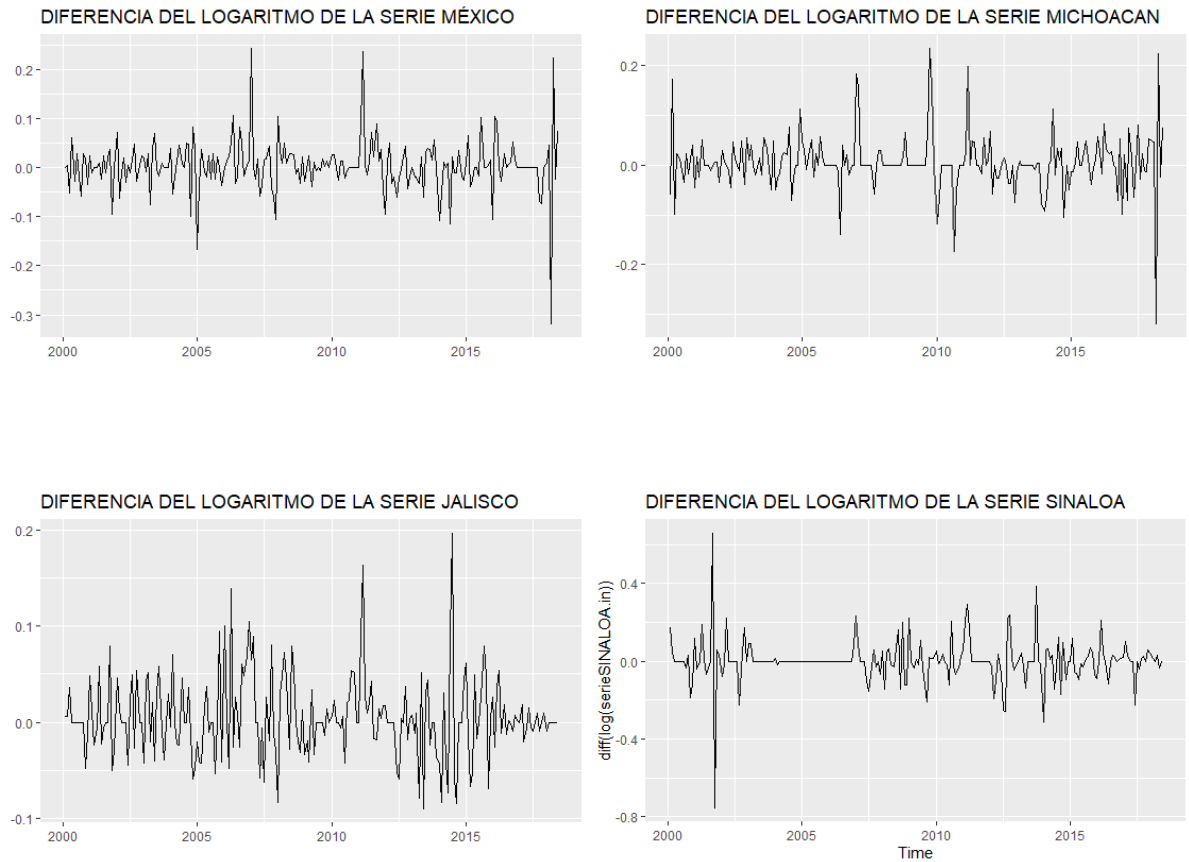


Figura 12. Series transformadas de precios de maíz blanco en México

Fuente: Elaboración propia

Los resultados de las prueba de Dickey-Fuller Aumentada (DFA) y Phillips Perron (Cuadro 1), indican que la serie correspondiente al precio de maíz blanco en el estado de Sinaloa es estacionaria en niveles y las series: Estado de México, Michoacán y Jalisco, son integradas de orden uno  $I(1)$ , es decir, son no estacionarias en niveles, sin embargo, las series son estacionarias mediante transformaciones de logaritmo natural y primera diferencia.

Cuadro 1. Prueba de raíz unitaria: Dickey Fuller Aumentada y Phillips Perron

VARIABLE	Estadística de Prueba		Orden de integración
	H <sub>0</sub> : Raíz Unitaria H1: No Raíz Unitaria		
	DFA	PP	
	Pr < Tau	Pr < Tau	
MEXICO	0.1997	0.2220	I (1)
Ln(PMEXICO)	0.3771	0.3496	I (1)
ΔLn(PMEXICO)	0.0005	<0.0001	I(0)
MICHOACAN	0.7946	0.3146	I (1)
Ln(PMICHOACÁN)	0.7988	0.5695	I(1)
ΔLn(PMICHOACÁN)	<0.0001	<0.0001	I(0)
JALISCO	0.4597	0.4268	I (1)
Ln(PJALISCO)	0.6396	0.6436	I (1)
ΔLn(PJALISCO)	0.0002	<0.0001	I(0)
SINALOA	0.0535	0.0190	I (0)
Ln(SINALOA)	0.0878	0.0031	I (0)
ΔLn(PSINALOA)	<0.0001	0.0010	I(0)

Fuente: Elaboración propia. \*Indica rechazo de la hipótesis nula a un nivel de significancia de 5%.  $\Delta$  denota al operador de primeras diferencias y Ln denota logaritmo natural.

Los del Cuadro 1 pueden verificarse gráficamente mediante el análisis del correlograma de cada serie de tiempo (Figura 13 y Figura 14) donde se muestra que los coeficientes de correlación (ACF) disminuyen lentamente a través del tiempo en las series originales.



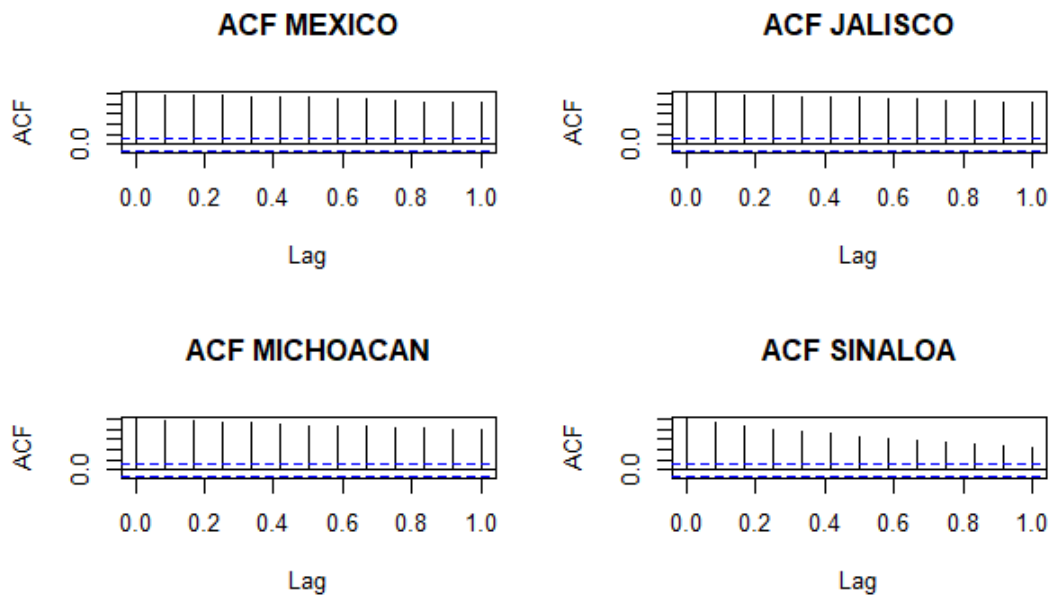


Figura 13. Función de autocorrelación (ACF) de series de datos nacionales

Fuente: Elaboración propia. \*series en niveles

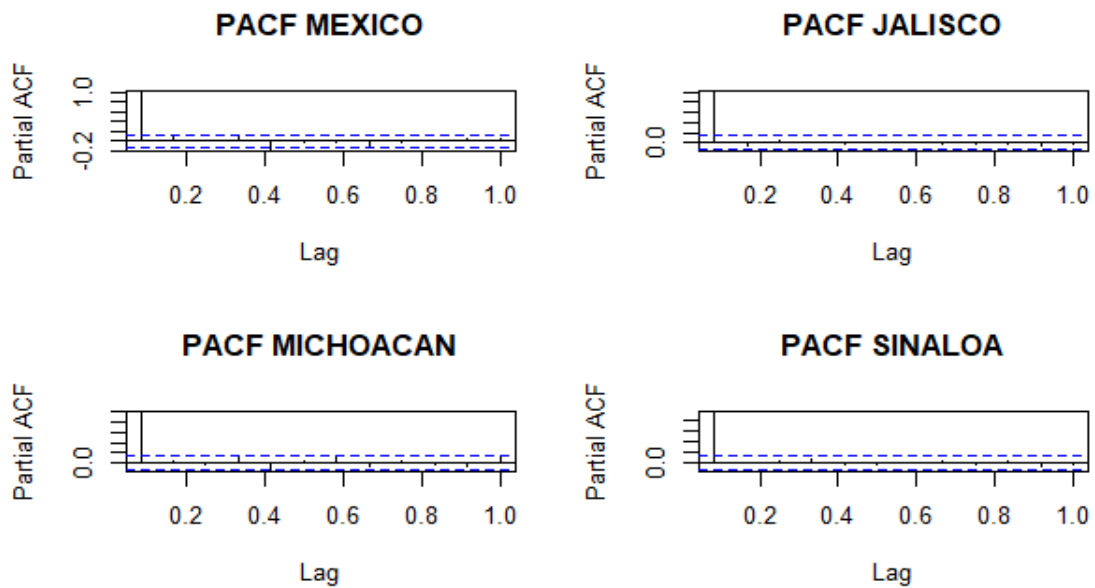


Figura 14. Función de autocorrelación parcial (PACF) de series de datos nacionales

Fuente: Elaboración propia. \*series en niveles

## 6.2 Identificación de modelos ARIMA

Siguiendo la metodología ARIMA, se realizó la fase de identificación de modelos. Mediante la inspección de las funciones de autocorrelación (ACF) y funciones de autocorrelación parcial (PACF) (Figura 15 y Figura 16), se definieron los valores  $(p, d, q)$  y se eligieron dos modelos para cada serie discriminando con base en los criterios AIC y Log Likelihood. Los coeficientes estimados y sus correspondientes estadísticas AIC y Log likelihood se muestran en el Cuadro 2 y Cuadro 3.

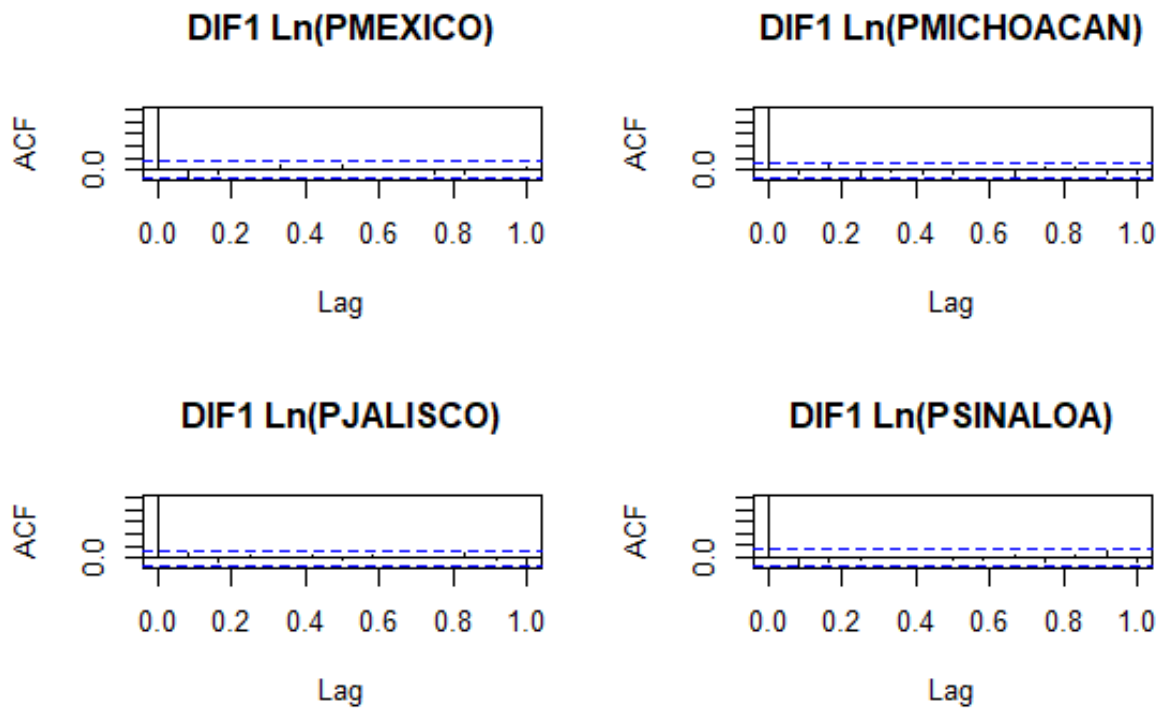


Figura 15. Función de autocorrelación (ACF) de la diferencia del logaritmo natural de series de datos nacionales

Fuente: Elaboración propia.

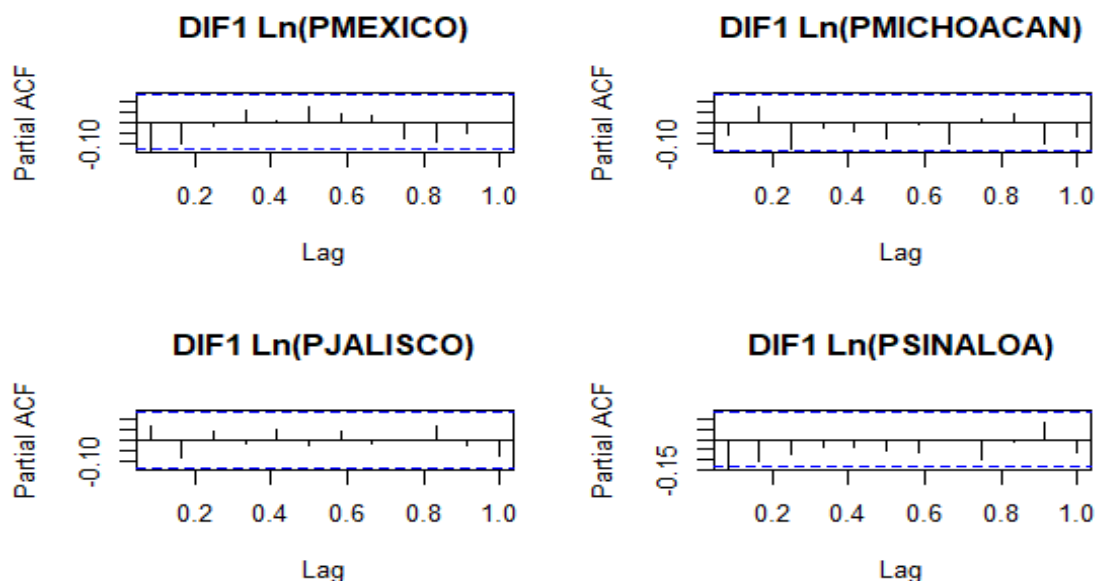


Figura 16. Función de autocorrelación parcial (PACF) de la diferencia del logaritmo natural de series de datos nacionales

Fuente: Elaboración propia.

Cuadro 2. Estimación de modelos ARIMA

SERIE	MODELO AJUSTADO	COEFICIENTE ESTIMADO			LOG LIKELIHOOD	AIC
		INTERCEPTO	AR (1)	MA (1) MA (2)		
$\Delta \text{Ln}(\text{PMÉXICO})$	MA (1)	0.004		-0.1669	337.95	-669.9
ee		-0.0715		-0.0029		
	ARIMA (1,0,1)	0.004	0.2288	-0.3884	338.3	-668.59
ee		-0.0028	-0.2386	-0.2568		
$\Delta \text{Ln}(\text{PMICHOCÁN})$	AR (1)	0.0051	-0.0563		326.49	-646.9
ee		-0.0035	-0.0675			
	MA (1)	0.0051		-0.0493	326.44	-646.89
ee		-0.0035		-0.0636		

FUENTE: Elaboración propia. ee denota Error Estándar

Cuadro 3. Estimación de modelos ARIMA (continuación)

SERIE	MODELO AJUSTADO	COEFICIENTE ESTIMADO			LOG LIKELIHOOD	AIC
		INTERCEPTO	AR (1)	MA (1)	MA (2)	
$\Delta \ln(\text{PJALISCO})$	MA (2)	0.0054		0.0822	-0.088	389.23
ee		-0.0028		-0.067	-0.0679	
	MA (1)	0.0054		0.0816		388.4
ee		-0.003		-0.0737		
$\Delta \ln(\text{PSINALOA})$	ARIMA (1,0,1)	0.0033	0.804	-0.9684		178.9
ee		-0.0013	-0.0638	-0.0342		
	MA (1)	0.0035		-0.1834		173.9
ee		0.0061		0.076		

FUENTE: Elaboración propia. ee denota Error Estándar

Además de observar las correlaciones residuales en los rezagos individuales, es útil tener una prueba que tenga en cuenta sus magnitudes como grupo (Cryer y Sik, 2008). La verificación de autocorrelación en los residuales puede realizarse mediante la estadística de Box y Pierce (1970) o Ljung-Box (1978).

Los resultados de la estadística de Ljung-Box indican que no se rechaza la hipótesis nula, por lo que se concluye que en los errores no existe autocorrelación (Cuadro 4).

Cuadro 4. Prueba Ljung-Box a los residuales en modelos ARIMA

SERIE	MODELO AJUSTADO	Ljung-Box test
		$H_0$ : No existe autocorrelación
$\Delta \text{Ln}(\text{PMÉXICO})$	MA (1)	$Q^* = 17.782$ , $df = 22$ , $p\text{-value} = 0.7188$
	ARIMA (1,0,1)	$Q^* = 16.378$ , $df = 21$ , $p\text{-value} = 0.748$
$\Delta \text{Ln}(\text{PMICHOACÁN})$	AR (1)	$Q^* = 20.74$ , $df = 22$ , $p\text{-value} = 0.5369$
	MA (1)	$Q^* = 20.96$ , $df = 22$ , $p\text{-value} = 0.5232$
$\Delta \text{Ln}(\text{PJALISCO})$	MA (2)	$Q^* = 18.458$ , $df = 21$ , $p\text{-value} = 0.6199$
	MA (1)	$Q^* = 21.139$ , $df = 22$ , $p\text{-value} = 0.5122$
$\Delta \text{Ln}(\text{PSINALOA})$	ARIMA (1,0,1)	$Q^* = 16.109$ , $df = 21$ , $p\text{-value} = 0.7635$
	MA (1)	$Q^* = 20.335$ , $df = 22$ , $p\text{-value} = 0.5621$

Fuente: Elaboración propia

Se realizaron gráficas de residuales, función de auto correlación (ACF) e histograma de frecuencias con línea de tendencia normal para los modelos ARIMA como se ejemplifica para la serie del precio de maíz en el estado de México (Figura 17).

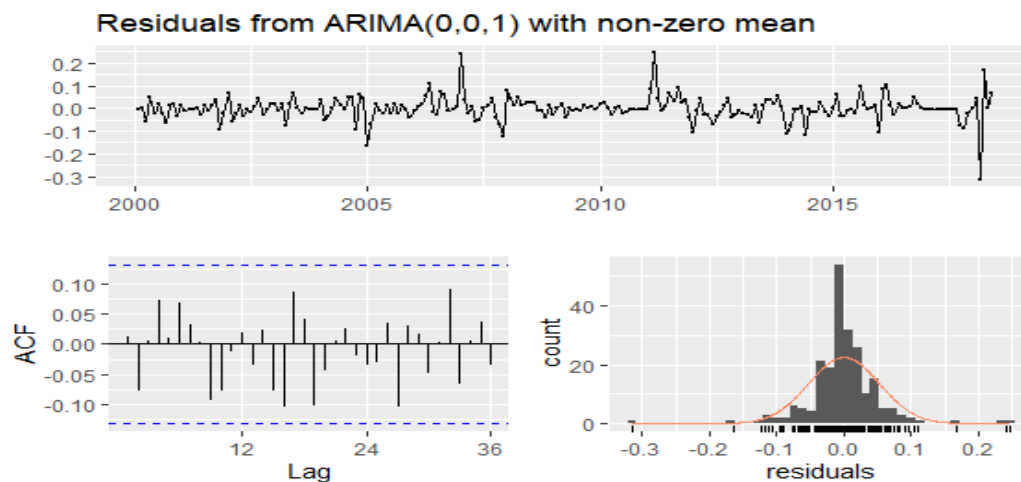


Figura 17. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie MEXICO, modelo ARIMA (0,0,1)

Fuente: Elaboración propia

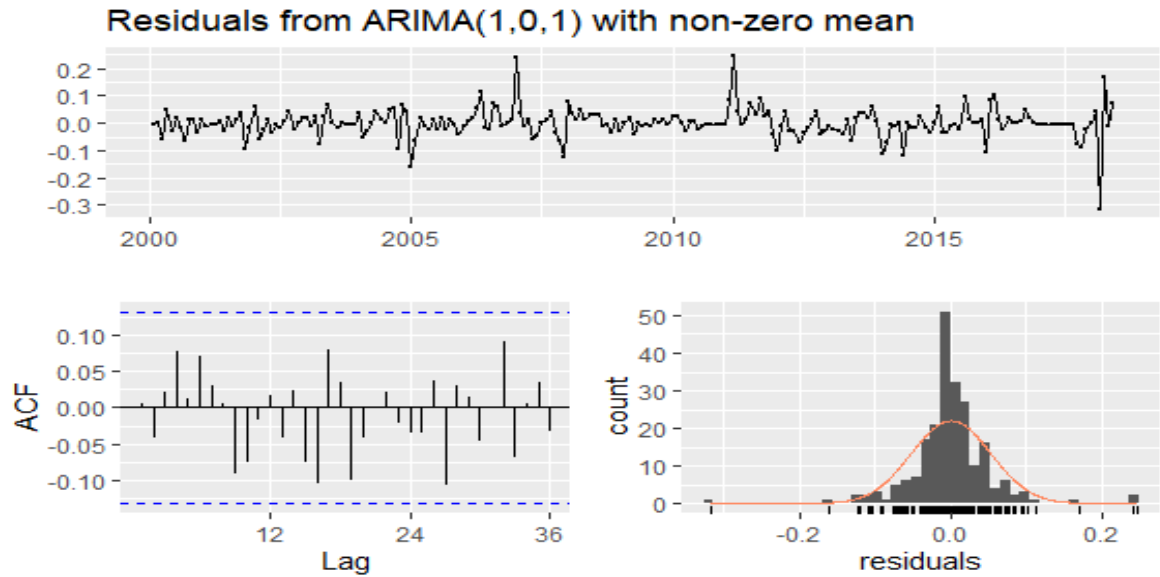


Figura 18. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie MEXICO, modelo ARIMA (1,0,1)

Fuente: Elaboración propia

Se realizaron gráficos similares para las series de precios en el estado de Michoacán, Jalisco y Sinaloa, éstos son mostrados en el Apéndice 2.

### 6.3 Predicción del precio de maíz mediante modelos autorregresivos integrados de promedio móvil (ARIMA)

Se realizó una comparación entre los valores observados y las estimaciones puntuales fuera de la muestra para los 6 meses inmediatos siguientes (julio a diciembre de 2018), los resultados se muestran en los cuadros 5-8.

Cuadro 5. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie México

MODELO	MA (1)	ARIMA (1,0,1)	
	PREDICHOS	PREDICHOS	OBSERVADOS
Jul 2018	5.210622	5.209732	5
Ago 2018	5.231461	5.216541	5
Sept 2018	5.252384	5.234133	4.75
Oct 2018	5.27339	5.254261	4.75
Nov 2018	5.294481	5.275036	4.3
Dic 2018	5.315655	5.296023	6.3

Fuente: Elaboración propia con datos de SNIIM

Cuadro 6. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie Michoacán

MODELO	AR (1)	MA (1)	
	PREDICHOS	PREDICHOS	OBSERVADOS
Jul 2018	5.256702	5.259425	5
Ago 2018	5.284727	5.286318	5
Sept 2018	5.311693	5.313348	4.75
Oct 2018	5.338864	5.340516	4.75
Nov 2018	5.366171	5.367823	5
Dic 2018	5.393618	5.39527	5

Fuente: Elaboración propia con datos de SNIIM

Cuadro 7. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie Jalisco

MODELO	MA (2)	MA (1)	
	PREDICHOS	PREDICHOS	OBSERVADOS
Jul 2018	5.32874	5.326339	5.3
Ago 2018	5.359954	5.354975	5.34
Sept 2018	5.388794	5.383764	5.52
Oct 2018	5.41779	5.412709	5.8
Nov 2018	5.446941	5.441809	5.2
Dic 2018	5.47625	5.471065	5.3

Fuente: Elaboración propia con datos de SNIIM

Cuadro 8. Precios observados y predichos obtenidos a partir de modelos ARIMA, serie Sinaloa

MODELO	ARIMA (1,0,1)	MA (1)	
	PREDICHOS	PREDICHOS	OBSERVADOS
Jul 2018	4.524409	4.522363	4.75
Ago 2018	4.550904	4.546677	4.60
Sept 2018	4.575391	4.571121	5.00
Oct 2018	4.600010	4.595696	5.00
Nov 2018	4.624761	4.620404	5.00
Dic 2018	4.649646	4.645244	5.00

Fuente: Elaboración propia con datos de SNIIM



Gráficamente los valores observados y predichos se muestran a continuación.

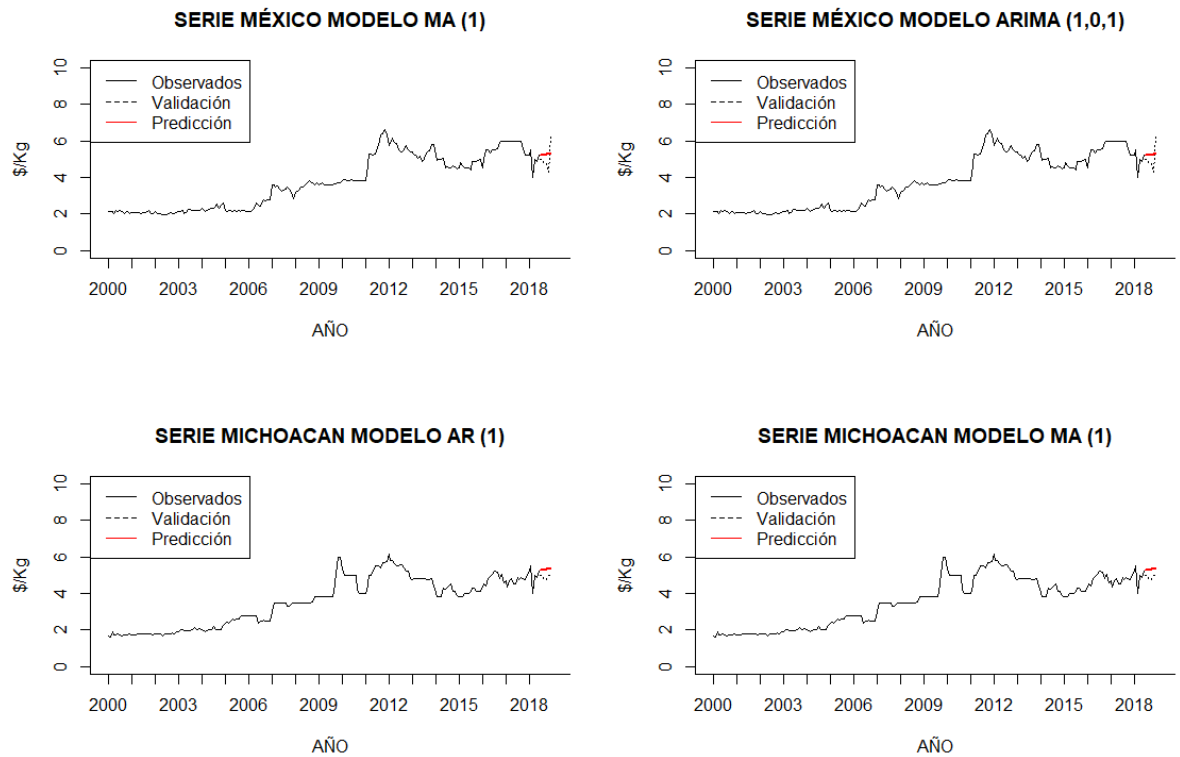


Figura 19. Pronósticos del precio de Maíz blanco mediante modelos ARIMA

Fuente: Elaboración propia con datos de SNIIM

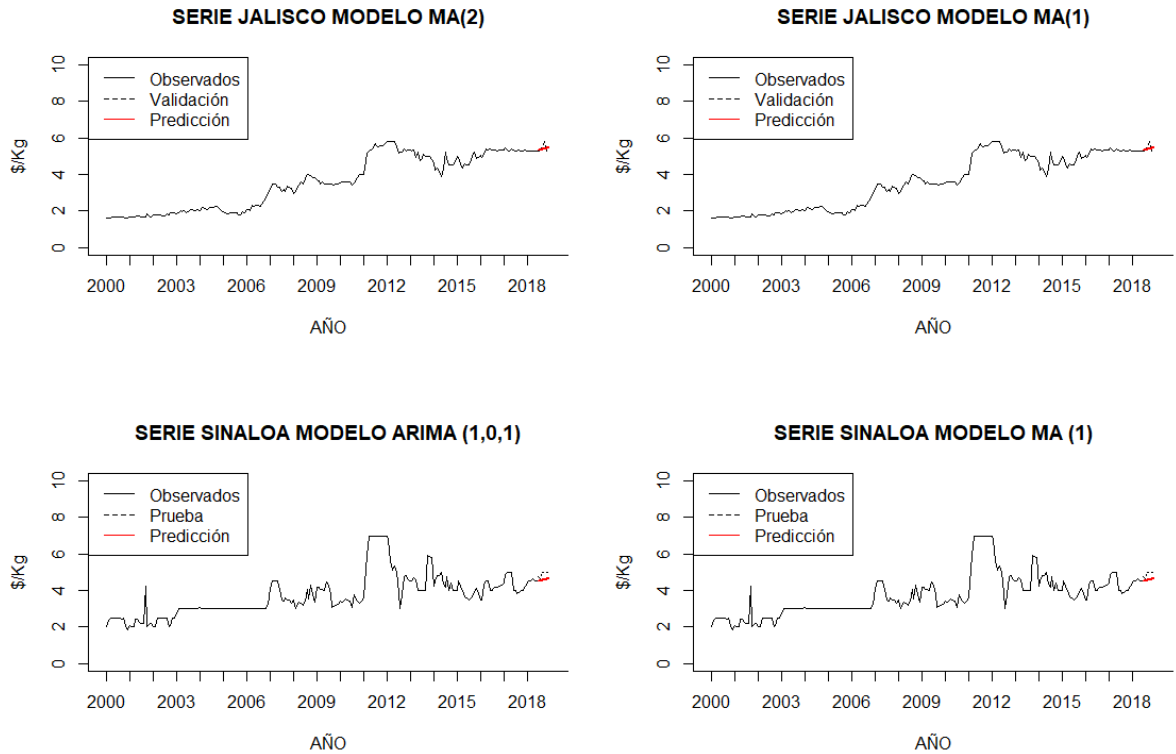


Figura 20. Pronósticos del precio de Maíz blanco mediante modelos ARIMA (Continuación)

Fuente: Elaboración propia con datos de SNIIM

#### 6.4 Modelos vectoriales de corrección de error (VEC) con dos variables

Los modelos VAR y VEC fueron construidos usando las series: precio promedio mensual de maíz amarillo en físico (precio FOB), precio de futuros y las series de precios nacionales: precio en estado de México, Michoacán, Jalisco y Sinaloa.

En las series presentadas en la Figura 21 correspondiente al precio promedio mensual de maíz amarillo (precio FOB) y Figura 22 correspondiente al precio promedio de futuros se muestra una tendencia creciente en todo el periodo de análisis. En el periodo comprendido entre año 2000 a 2005 se presentó una relativa estabilidad en el que el precio de futuros y el precio de maíz en físico no superaron el precio de 3.16 dólares/bushel = 1.27\$/kg y 2.89 dólares/bushel =

1.17\$/kg respectivamente; posterior a 2005 y hasta 2015 se tuvo un periodo caracterizado por una mayor volatilidad comparado con los años anteriores, donde se tuvieron aumentos y disminuciones en ambos precios oscilando entre 1 y 3.9 \$/kg; a partir del año 2015 ambos precios<sup>21</sup> tuvieron comportamientos más estables que oscilaron entre los dos y tres pesos / kg.

En las figuras 21 y 22 resalta el hecho de que los precios físicos y precios futuros presentan una tendencia creciente, además, la media y la varianza parecen cambiar con el tiempo, lo que sugiere que las series presentan raíz unitaria.

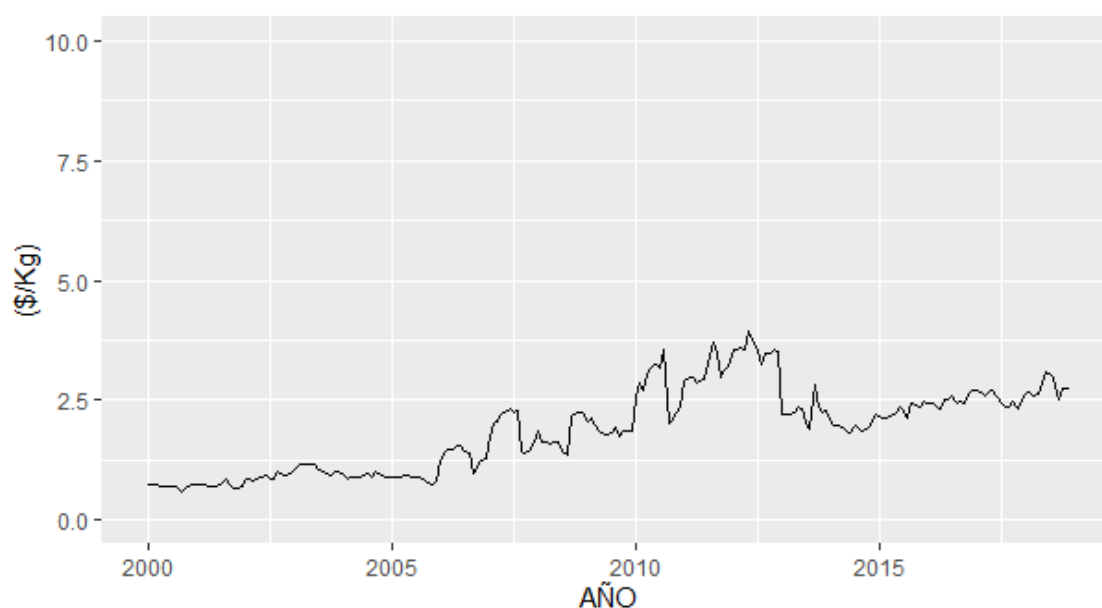


Figura 21. Precio promedio mensual de maíz amarillo en físico (Precio FOB)

Fuente: Elaboración propia con datos de USDA. \* maíz amarillo N° 2, FOB en el Golfo de México.

<sup>21</sup> Los precios originales se encuentran en dólares por bushel y fueron transformados con fines comparativos con los datos nacionales en pesos/Kg

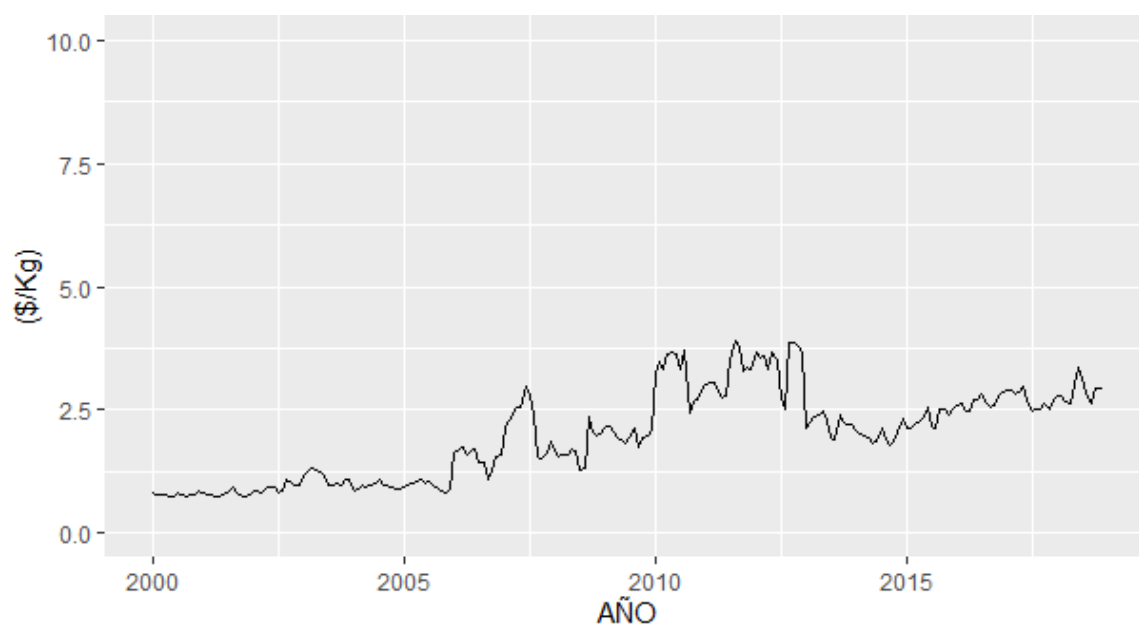


Figura 22. Precio promedio de futuros de maíz

Fuente: Elaboración propia con datos de USDA. \*Precios de liquidación mensual promedio del contrato de futuros de maíz más cercano

Cuadro 9. Prueba de raíz unitaria a precios del mercado internacional

VARIABLE	Estadística de Prueba		Orden de integración
	H <sub>0</sub> : Raíz Unitaria Ha: No Raíz Unitaria		
	DFA	PP	
	Pr < Tau	Pr < Tau	
PRECIO FISICO	0.2589	0.1179	I (1)
Δ (PRECIO FISICO)	<0.0001	<0.0001	I (0)
PRECIO FUTURO	0.0338	0.0320	I (0)

Fuente: Elaboración propia. \*Indica rechazo de la hipótesis nula a un nivel de significancia de 5%. Δ denota al operador de primeras diferencias.

Mediante las pruebas Dickey-Fuller Aumentada (DFA) y Phillips Perron (PP) se comprobó que la serie de precios de futuros es estacionaria. Por otra parte, la serie de precios de maíz en físico presenta raíz unitaria, sin embargo, es estacionaria en primera diferencia (Figura 23).

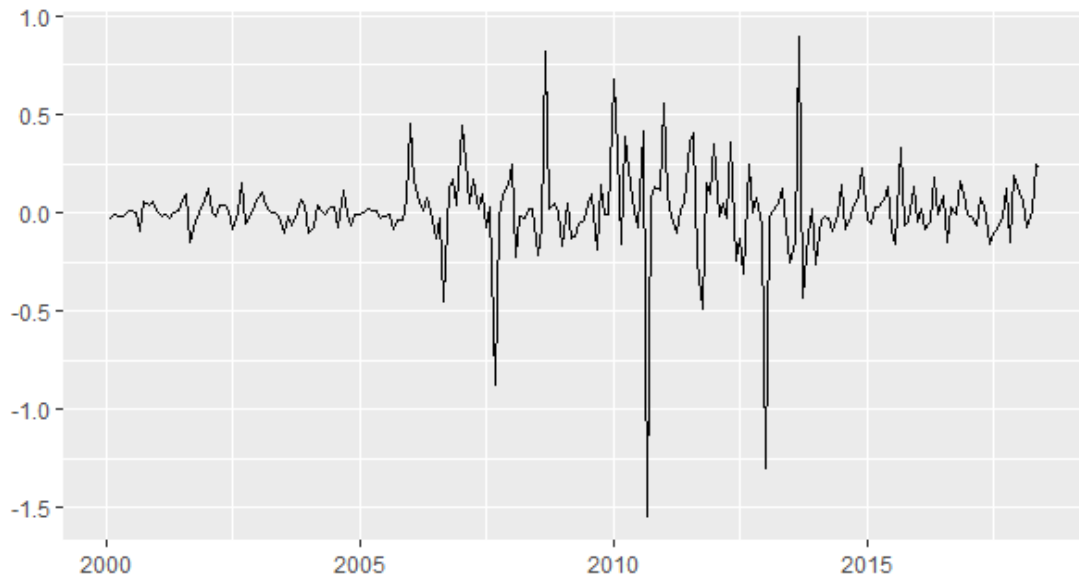


Figura 23. Precios de maíz en físico (FOB) (serie en primera diferencia)

Fuente: Elaboración propia

La construcción de un modelo vectorial de corrección de error (VEC) requiere el establecimiento del orden de integración  $I(d)$  y la existencia de cointegración bajo el enfoque del teorema de Granger (Maparu y Mazumder, 2017). Mediante las pruebas Dickey- Fuller y Phillis Perron se comprobó que las series: México, Michoacán, Jalisco y precio de maíz en físico son procesos integrados de orden (1).

Se construyeron los siguientes modelos VEC con dos variables: precio de maíz en el estado de México y precio de maíz en físico (precio FOB), precio de maíz en el estado de Michoacán y precio de maíz en físico, precio de maíz en el estado de Jalisco y precio físico, precio de maíz en el estado de Sinaloa y precio físico.

El teorema de representación de Granger (Engel y Granger, 1987; Granger, 1981) establece que si un conjunto de series cointegran, se identifican entonces como un bloque que tiende a moverse simultáneamente en el tiempo, situación que puede identificarse como una relación de “equilibrio” (Galindo y Catalán, 2003).

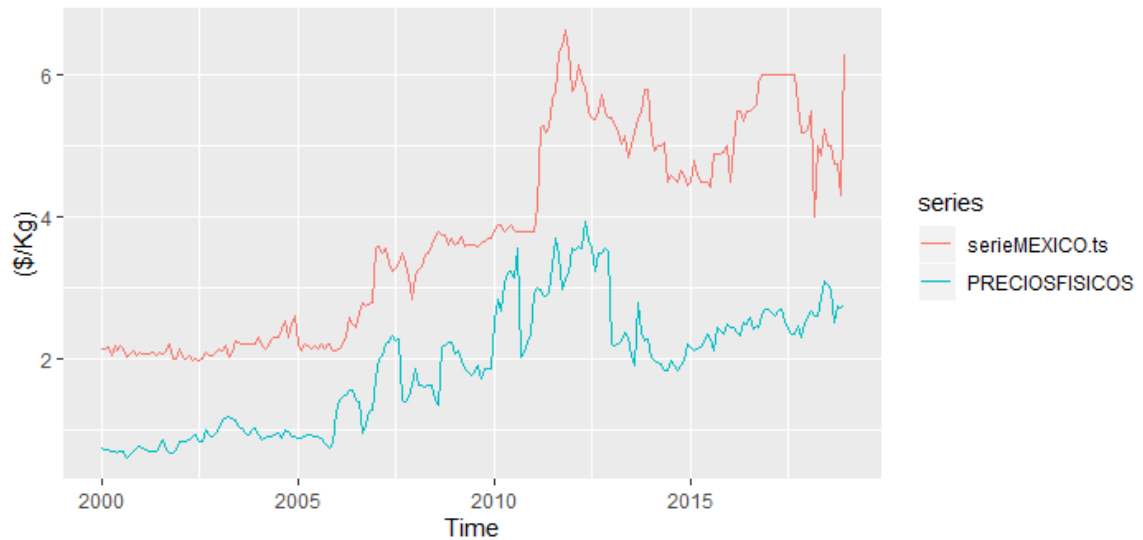


Figura 24. Precio de maíz en el Estado de México Vs precio de maíz en físico EU

Fuente: Elaboración propia con datos de USDA y SNIIM

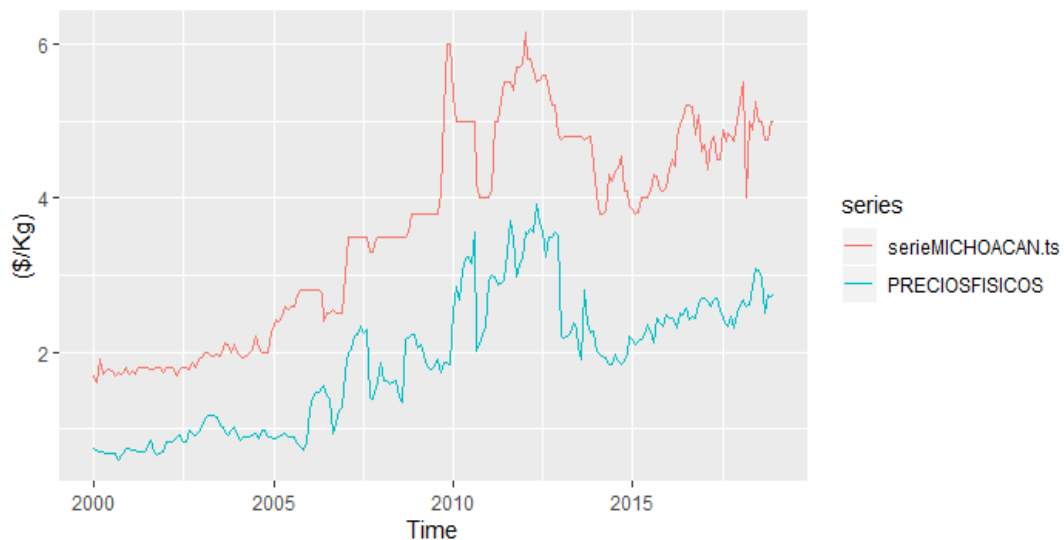


Figura 25. Precio de maíz en Michoacán Vs precio de maíz en físico EU

Fuente: Elaboración propia con datos de USDA y SNIIM

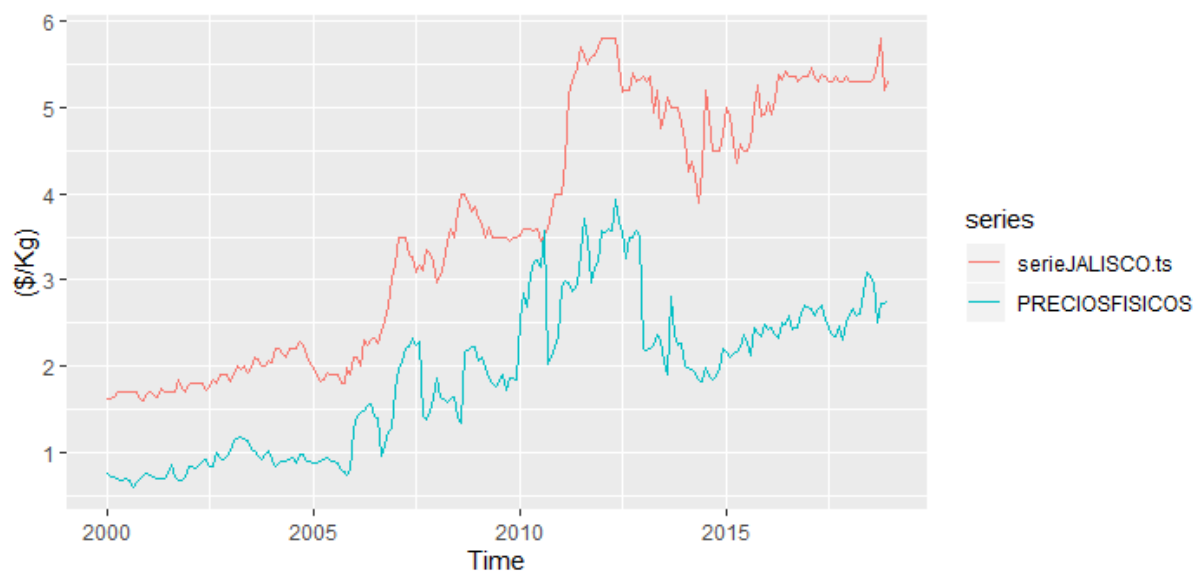


Figura 26. Precio de maíz en Jalisco Vs precio de maíz en físico EU

Fuente: Elaboración propia con datos de USDA y SNIIM

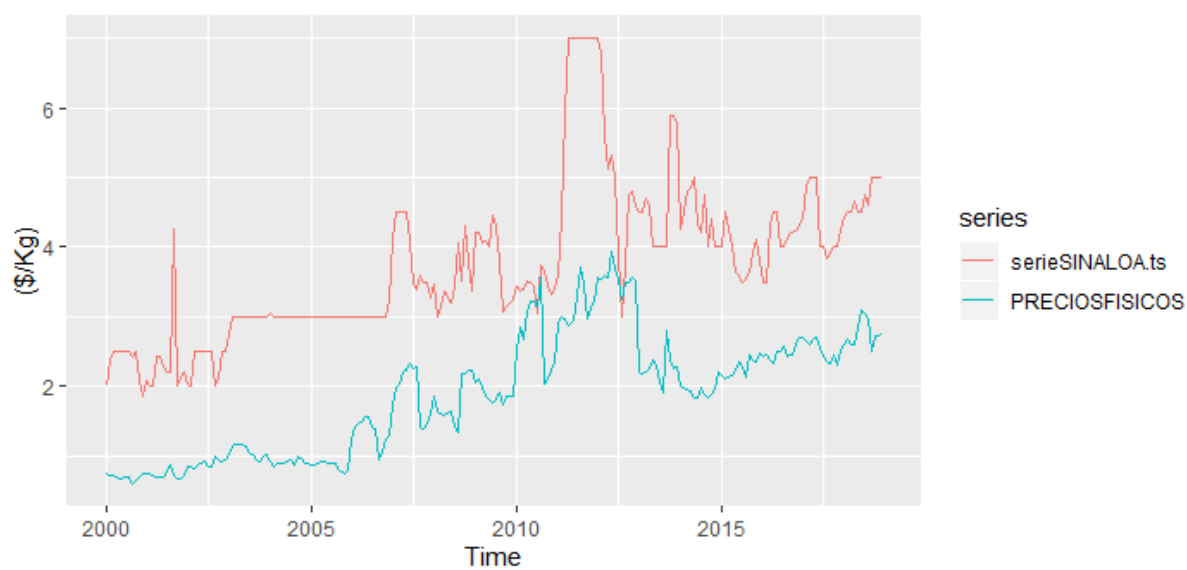


Figura 27. Precio de maíz en Sinaloa Vs precio de maíz en físico EU

Fuente: Elaboración propia con datos de USDA y SNIIM

La metodología de los modelos VAR considera una parte importante del proceso la elección del rezago adecuado, evitando tener un modelo sobre parametrizado con estimaciones de coeficientes no significativos (Enders, 2015), por lo cual, como primer paso se realizó la elección del número adecuado de rezagos del modelo VAR (el valor de  $p$ ) siguiendo los siguientes criterios: el Criterio de Información de Akaike (AIC), el Criterio de Información de Schwarz (SC) y el Criterio de Hannan Quinn (HQ) y con base en la significancia estadísticas de los rezagos dentro del modelo. Para cada modelo se estableció como rezago óptimo  $p = 2$ .

Para determinar la existencia de relaciones de cointegración entre las variables en niveles se realizó la prueba de cointegración de Johansen con base en el estadístico del máximo eigenvalor. Dicha prueba indica que existe al menos una relación de cointegración en cada uno de los siguientes modelos.

Cuadro 10. Contraste de cointegración de Johansen de modelos VEC

MODELO	$H_0$ : Existen $r$ vectores de cointegración	Valor del estadístico de prueba <sup>1</sup>	Valores Críticos			Eigenvalores (Lambda)
			10pct	5pct	1pct	
PMÉXICO PFÍSICO	$*r \leq 1$	1.05	6.50	8.18	11.65	0.09071
	$r \leq 0$	20.92	12.91	14.90	19.19	0.00047
PMICHOACÁN PFÍSICO	$*r \leq 1$	1.64	6.50	8.18	11.65	0.14603
	$r \leq 0$	34.73	12.91	14.90	19.19	0.0074
PJALISCO PFÍSICO	$*r \leq 1$	0.94	6.50	8.18	11.65	0.08451
	$r \leq 0$	19.43	12.91	14.90	19.19	0.00424

<sup>1</sup>Estadístico: Traza.  $r$ : Indica el número de relaciones de cointegración que se está probando.

\* Indica no rechazo de la hipótesis nula.

Fuente: Elaboración propia



En la práctica es aconsejable estimar el modelo VEC por máxima verosimilitud para lograr la máxima eficiencia estadística posible (Guerrero, 2007). Habiendo determinado mediante el contraste de Johansen la existencia de al menos una relación de cointegración entre las series de cada modelo, se estimaron los parámetros del modelo bivariado VEC mediante el método de máxima verosimilitud con base en la metodología establecida en Lütkepohl (2005).

Los modelos vectoriales de corrección del error (VEC) con dos variables, se definieron como se muestra en el siguiente sistema matricial, donde se tiene interés únicamente en predecir el precio de maíz en los estados de México, Michoacán y Jalisco.

El modelo para el precio en el Estado de México puede definirse como:

$$\begin{bmatrix} \Delta PMEX \\ \Delta PFIS \end{bmatrix} = \begin{bmatrix} \alpha PMEX \\ \alpha PFIS \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} \Delta PMEX \\ \Delta PFIS \end{bmatrix} - \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{21} & \Pi_{22} \end{bmatrix} \begin{bmatrix} \hat{\varepsilon}_{t-1}^{PMEX} \\ \hat{\varepsilon}_{t-1}^{PFIS} \end{bmatrix} + \begin{bmatrix} \epsilon_{PMEX} \\ \epsilon_{PFIS} \end{bmatrix}$$

Donde:

$\Delta$ : Es el operador de diferencias

$PMEX$ : Es el precio de maíz en el estado de México

$PFIS$ : Es el precio de maíz en físico en Estados Unidos

$\alpha$ : Es parámetro de intercepto a estimar

$\beta$ : Parámetros a estimar

$\Pi$ : es el vector cointegrante

$\varepsilon$ : son errores resultantes del modelo de regresión:  $PMEX = \alpha + \beta PFIS + \varepsilon$  rezagados un periodo

$\epsilon$ : Son los términos de error estimados

Los modelos para las series del Estado de Michoacán y Jalisco se formularon de manera similar.

Cabe destacar que la serie de Sinaloa (PSINALOA) es estacionaria originalmente por lo que se estableció un modelo VAR con dos variables: PSINALOA Vs  $\Delta$ PFISICO.

Con base en la estadística de Ljung-Box no se rechaza la hipótesis de no autocorrelación de los errores.

Cuadro 11. Prueba Ljung-Box a los residuales de modelos VEC

MODELO AJUSTADO VEC	Ljung-Box test $H_0$ : No existe autocorrelación
PMEX PFIS	X-squared = 1.9158e-05, df = 1, p-value = 0.9965
PMICH PFISICO	X-squared = 0.010036, df = 1, p-value = 0.9202
PJALISCO PFISICO	X-squared = 0.029634, df = 1, p-value = 0.8633
MODELO AJUSTADO VAR	
PSINALOA $\Delta$ PFISICO	X-squared = 0.010869, df = 1, p-value = 0.917

Fuente: Elaboración propia.  $\Delta$  indica la primera diferencia de la serie

## 6.5 Predicción del precio de maíz mediante modelos VEC

Se realizó un comparativo de los valores observados y predichos en los siguientes seis meses y se obtuvieron los siguientes valores:

Cuadro 12. Precios observados y predichos obtenidos a partir de un modelo VEC, serie México

	PREDICHOS	OBSERVADOS
Jul 2018	5.265407	5
Ago 2018	5.32044	5
Sept 2018	5.364324	4.75
Oct 2018	5.40364	4.75
Nov 2018	5.438702	4.3
Dic 2018	5.470321	6.3

Fuente: Elaboración propia con datos de SNIIM

Cuadro 13. Precios observados y predichos obtenidos a partir de un modelo VEC, Serie Michoacán

	PREDICHOS	OBSERVADOS
Jul 2018	5.266788	5
Ago 2018	5.294587	5
Sept 2018	5.319006	4.75
Oct 2018	5.34082	4.75
Nov 2018	5.360861	5
Dic 2018	5.379727	5

Fuente: Elaboración propia con datos de SNIIM

Cuadro 14. Precios observados y predichos obtenidos a partir de un modelo VEC, Serie Jalisco

	PREDICHOS	OBSERVADOS
Jul 2018	5.318384	5.3
Ago 2018	5.346429	5.34
Sept 2018	5.373724	5.52
Oct 2018	5.399235	5.8
Nov 2018	5.423192	5.2
Dic 2018	5.445876	5.3

Fuente: Elaboración propia con datos de SNIIM

Cuadro 15. Precios observados y predichos obtenidos a partir de un modelo VAR, Serie Sinaloa

	PREDICHOS	OBSERVADOS
Jul 2018	4.528049	4.75
Ago 2018	4.477373	4.6
Sept 2018	4.434453	5
Oct 2018	4.394323	5
Nov 2018	4.357027	5
Dic 2018	4.322354	5

Fuente: Elaboración propia con datos de SNIIM

En las siguientes gráficas se presentan los pronósticos de las series analizadas con base en los modelos VEC.

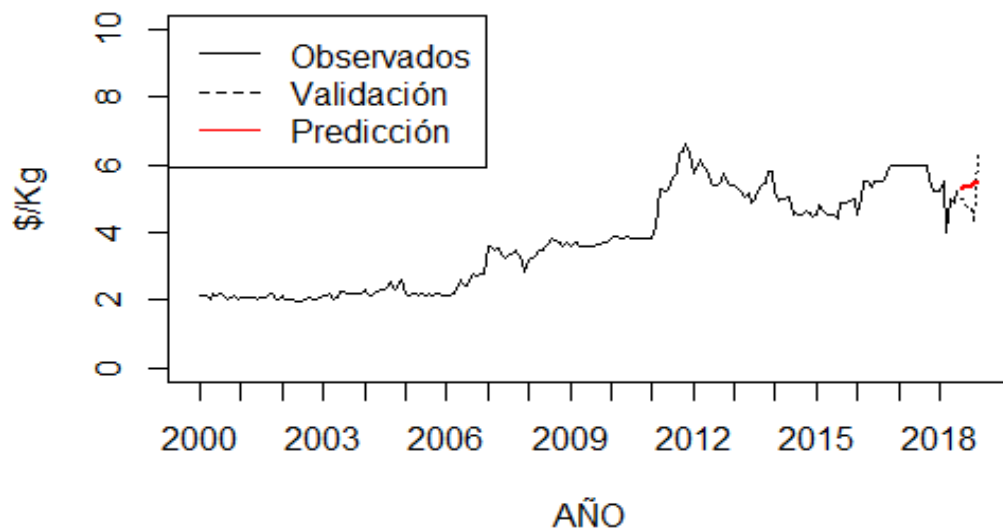


Figura 28. Predicción fuera de la muestra, serie México, modelo VEC

Fuente: Elaboración propia con datos de SNIIM

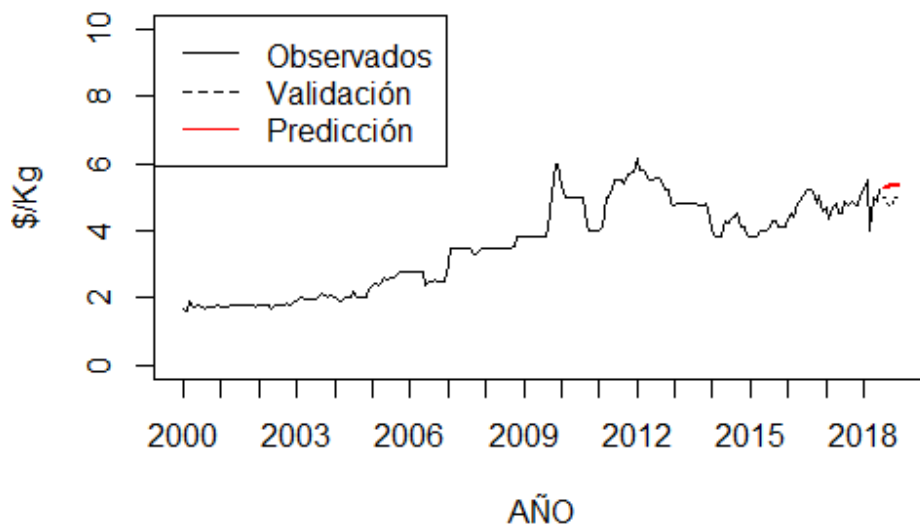


Figura 29. Predicción fuera de la muestra, serie Michoacán, modelo VEC

Fuente: Elaboración propia con datos de SNIIM

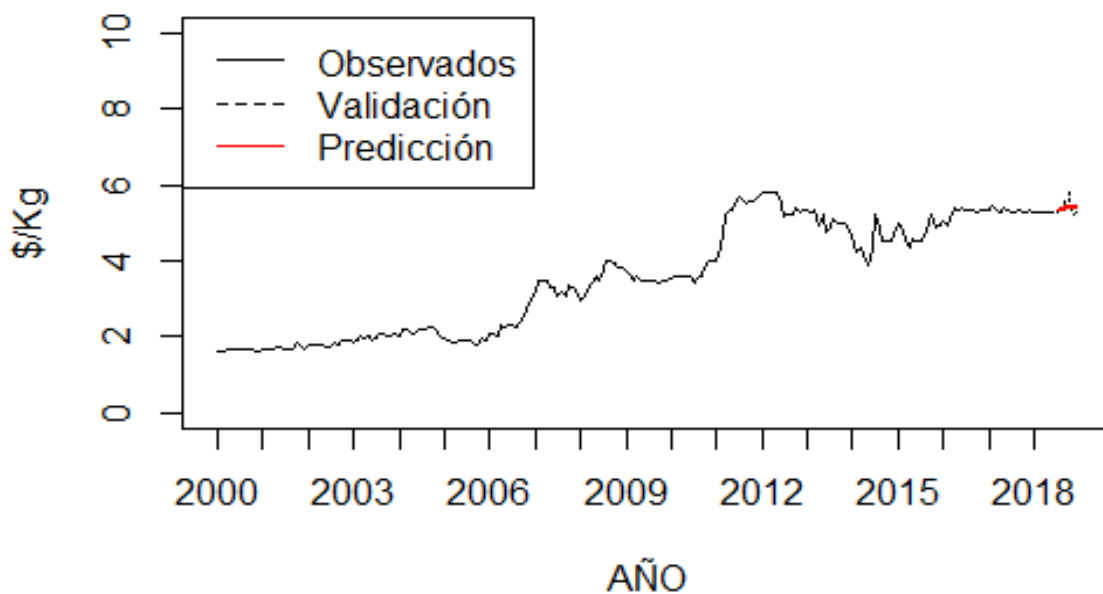


Figura 30. Predicción fuera de la muestra, serie Jalisco, modelo VEC

Fuente: Elaboración propia con datos de SNIIM

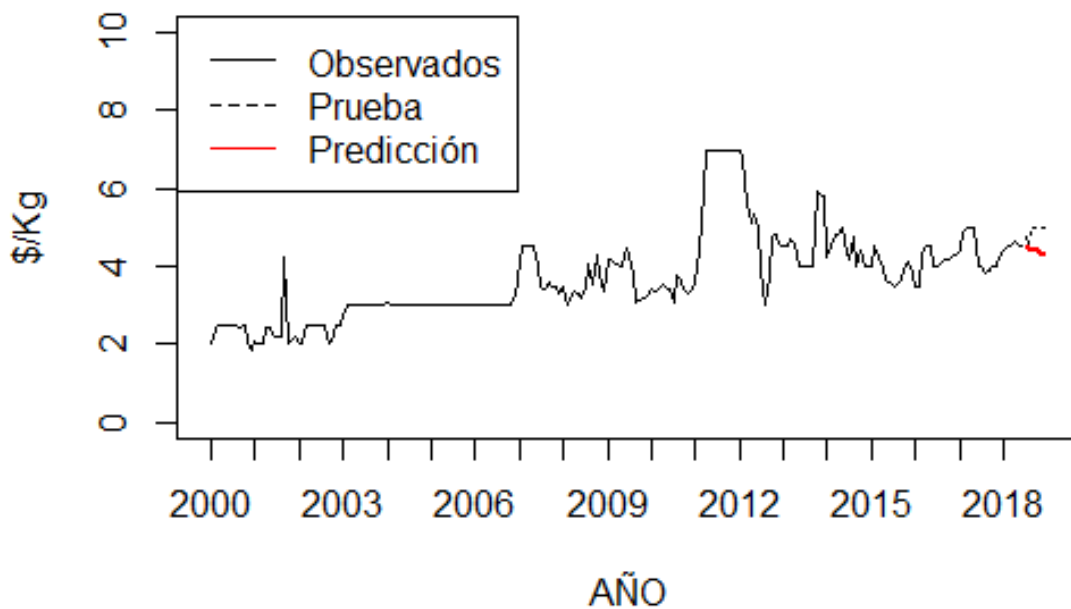


Figura 31. Predicción fuera de la muestra, serie Sinaloa, modelo bivariado VAR(1)

Fuente: Elaboración propia con datos de SNIIM

## 6.6 Modelos vectoriales autorregresivos (VAR)

Existen metodologías para la modelación de series de tiempo univariadas y multivariadas. En este caso se considera beneficioso el uso de modelos multivariados que tomen en cuenta la mayor cantidad de información disponible para la realización de pronósticos.

Uno de los objetivos de los modelos multivariados en series de tiempo es establecer las relaciones e interacciones entre variables, por lo tanto, se tuvo como objetivo mostrar la influencia que ejercen las variables: precio de maíz en físico (FOB) y precio de futuros sobre los precios de maíz en México en cuyo caso, el uso de estas variables proporciona una mejora en los pronósticos comparados con los obtenidos mediante modelos univariados ARIMA.

En un sentido estricto, en un modelo VAR todas las variables deben ser estacionarias en forma conjunta (Gujarati y Porter, 2010), atendiendo esto, se establecieron los modelos VAR con las series estacionarias:  $\Delta PMéxico$   $\Delta PFísico$   $PFuturo$ ,  $\Delta PMichoacán$   $\Delta PFísico$   $PFuturo$ ,  $\Delta PJalisco$   $\Delta PFísico$   $PFuturo$ ,  $PSinaloa$   $\Delta PFísico$   $PFuturo$ . Donde las series: precio de México, precio de Michoacán, precio de Jalisco y precio físico son no estacionarias y fueron transformadas<sup>22</sup> y las series: precio de Sinaloa y precio futuro son estacionarias  $I(0)$ .

---

<sup>22</sup>  $\Delta$  Denota que la serie ha sido diferenciada una vez

La estimación de los parámetros de los modelos VAR en forma matricial se puede expresar de la siguiente manera:

$$\begin{bmatrix} \Delta PMEX \\ \Delta PFIS \\ PFUT \end{bmatrix} = \begin{bmatrix} \alpha \Delta PMEX \\ \alpha \Delta PFIS \\ \alpha PFUT \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} \begin{bmatrix} \Delta PMEX \\ \Delta PFIS \\ PFUT \end{bmatrix}_{t-1} + \dots + \begin{bmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \\ \beta_{31} & \beta_{32} & \beta_{33} \end{bmatrix} \begin{bmatrix} \Delta PMEX \\ \Delta PFIS \\ PFUT \end{bmatrix}_{t-p} + \begin{bmatrix} \varepsilon_{\Delta PMEX} \\ \varepsilon_{\Delta PFIS} \\ \varepsilon_{PFUT} \end{bmatrix}$$

Donde:

*PMEX*: Es el precio de maíz en el estado de México

*PFIS*: Es el precio de maíz en físico en Estados Unidos

*PFUT*: Es el precio de futuros.

$\Delta$ : Es el operador de diferencias

$\alpha$ : Es el intercepto

$\beta$ : Parámetros a estimar

La selección del rezago adecuado para cada modelo VAR con tres variables se realizó mediante un procedimiento equivalente al seguido en los modelos VEC. Se construyeron modelos con  $p = 1$  y  $p = 2$  rezagos para el precio de maíz en México, Jalisco y Sinaloa y  $p = 1$  y  $p = 3$  para el precio de maíz en Michoacán.

Posteriormente se procedió a la estimación de los parámetros de los modelos. El método usual para la estimación de los parámetros en los modelos VAR es Mínimos Cuadrados Ordinarios (MCO) aplicado a cada ecuación por separado (Gujarati y Porter, 2010).



La estabilidad dinámica de un modelo VAR consiste en que ante una perturbación o choque aleatorio las variables regresen a su trayectoria de equilibrio de largo plazo; de lo contrario el modelo es inestable con un comportamiento explosivo (Loría, 2007). Siempre que las raíces características del sistema sean menores a uno en valor absoluto. En este caso las raíces del polinomio característico obtenidas en cada uno de los modelos planteados son menores a uno, por lo que se considera que modelos estimados son estables.

El  $p$  valor  $\geq 0.05$  asociado a la prueba de Ljung-Box indica no se rechaza la hipótesis de no autocorrelación de los errores en los modelos VAR (Cuadro 16). Se verificó que los residuales exhiban un comportamiento similar a ruido blanco mediante sus funciones de autocorrelación simple y parcial, asimismo, se realizaron pruebas de raíz unitaria para verificar que los errores sean estacionarios.

En los modelos VAR no se acostumbra a analizar los coeficientes de regresión estimados ni sus significancias estadísticas; tampoco la bondad del ajuste ( $R^2$  ajustado) de las ecuaciones individuales. Es común que se verifique la ausencia de correlación serial de los residuos de las ecuaciones individuales del modelo y la distribución normal multivariada de éstos. Algunos investigadores efectúan pruebas adicionales, como la estabilidad del modelo, la significancia conjunta de las variables consideradas, su dirección de causalidad, la cointegración de los residuos de las regresiones individuales y la descomposición de la varianza del error de pronóstico (Arias y Torres, 2004).

Cuadro 16. Prueba Ljung-Box a los residuales de los modelos VAR

SERIE	MODELO	Ljung-Box test
		$H_0$ : No existe autocorrelación
$\Delta$ PMEXICO $\Delta$ PFISICO PFUTURO	VAR (1)	X-squared = 0.019114, df = 1 p-value = 0.89
$\Delta$ PMEXICO $\Delta$ PFISICO PFUTURO	VAR (2)	X-squared = 0.19274, df = 1 p-value = 0.6606
$\Delta$ PMICHOACAN $\Delta$ PFISICO PFUTURO	VAR (1)	X-squared = 0.008472, df = 1 p-value = 0.9267
$\Delta$ PMICHOACAN $\Delta$ PFISICO PFUTURO	VAR (3)	X-squared = 0.00040389, df = 1 p-value = 0.984
$\Delta$ PJALISCO $\Delta$ PFISICO PFUTURO	VAR (1)	X-squared = 0.035861, df = 1 p-value = 0.8498
$\Delta$ PJALISCO $\Delta$ PFISICO PFUTURO	VAR (2)	X-squared = 0.0047847, df = 1 p-value = 0.9449
PSINALOA $\Delta$ PFISICO PFUTURO	VAR (1)	X-squared = 0.00048973, df = 1 p-value = 0.9823
PSINALOA $\Delta$ PFISICO PFUTURO	VAR (2)	X-squared = 0.00078182, df = 1 p-value = 0.9777

Fuente: Elaboración propia.

Los pronósticos de los modelos VAR para los meses de julio a diciembre de 2018 se muestran a continuación:

Cuadro 17. Precios observados y predichos obtenidos a partir de los modelos VAR, serie México

	PREDICHOS VAR(1)	PREDICHOS VAR(2)	OBSERVADOS
Jul 2018	5.239002	5.283500	5
Ago 2018	5.266470	5.312557	5
Sept 2018	5.289488	5.322914	4.75
Oct 2018	5.313013	5.332864	4.75
Nov 2018	5.336441	5.343970	4.3
Dic 2018	5.359846	5.355617	6.3

Fuente: Elaboración propia.

Cuadro 18. Precios observados y predichos obtenidos a partir de los modelos VAR, serie Michoacán

	PREDICHOS VAR(1)	PREDICHOS VAR(2)	OBSERVADOS
Jul 2018	5.249819	5.078685	5
Ago 2018	5.265032	5.129401	5
Sept 2018	5.280436	5.069171	4.75
Oct 2018	5.295738	5.088467	4.75
Nov 2018	5.311031	5.070352	5
Dic 2018	5.326306	5.072713	5

Fuente: Elaboración propia.

Cuadro 19. Precios observados y predichos obtenidos a partir de los modelos VAR, serie Jalisco

	PREDICHOS VAR(1)	PREDICHOS VAR(2)	OBSERVADOS
Jul 2018	5.316005	5.307016	5.3
Ago 2018	5.337683	5.319324	5.34
Sept 2018	5.359665	5.336745	5.52
Oct 2018	5.381652	5.354584	5.8
Nov 2018	5.403609	5.371191	5.2
Dic 2018	5.425533	5.387496	5.3

Fuente: Elaboración propia.

Cuadro 20. Precios observados y predichos obtenidos a partir de los modelos VAR, serie Sinaloa

	PREDICHOS VAR(1)	PREDICHOS VAR(2)	OBSERVADOS
Jul 2018	4.639459	4.619207	4.75
Ago 2018	4.689733	4.651813	4.6
Sept 2018	4.736690	4.662214	5
Oct 2018	4.777827	4.674476	5
Nov 2018	4.813927	4.682537	5
Dic 2018	4.845571	4.685028	5

Fuente: Elaboración propia.

En las siguientes gráficas se presentan los pronósticos de las series analizadas con base en los modelos VAR.



Figura 32. Predicción fuera de la muestra, serie México, modelo VAR(2)

Fuente: Elaboración propia con datos de SNIIM

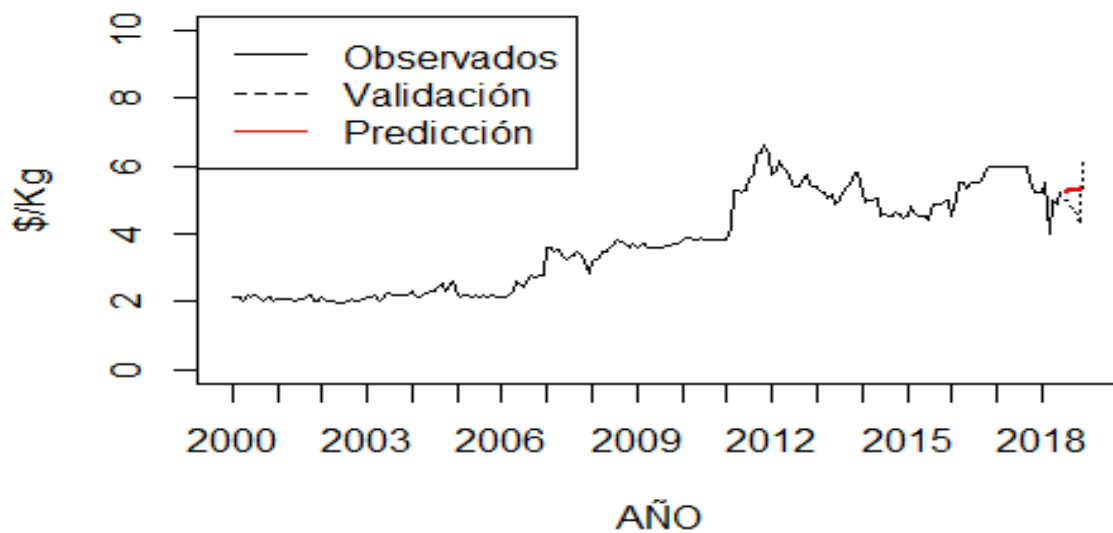


Figura 33. Predicción fuera de la muestra, serie México, modelo VAR(1)

Fuente: Elaboración propia con datos de SNII

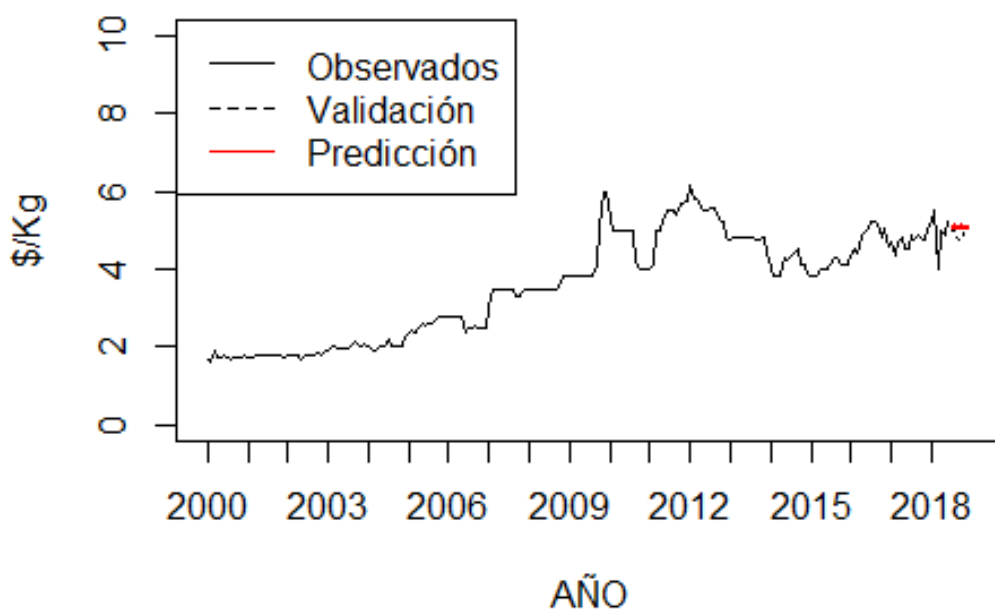


Figura 34. Predicción fuera de la muestra, serie Michoacán, modelo VAR(3)

Fuente: Elaboración propia con datos de SNIIM

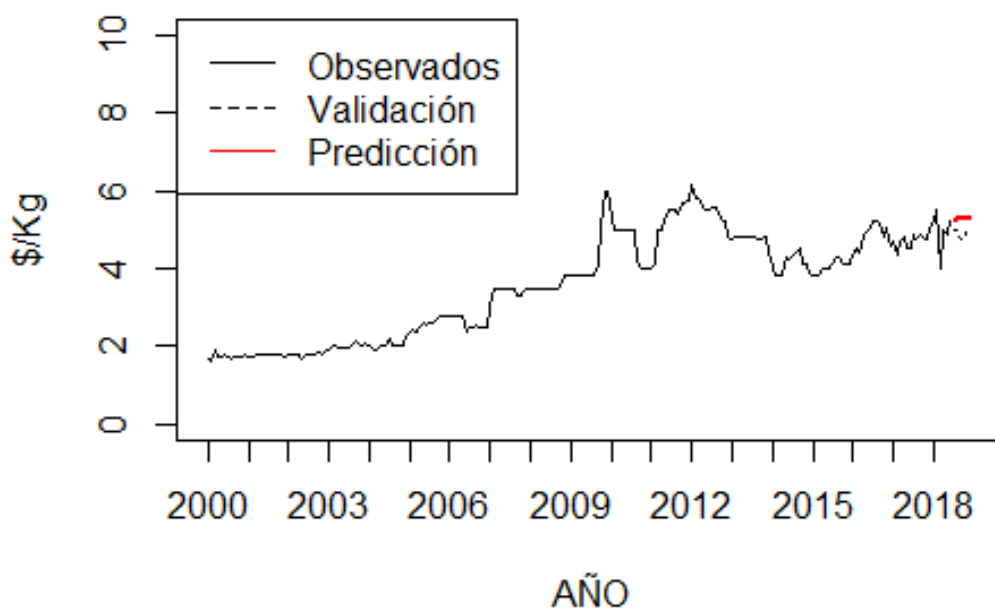


Figura 35. Predicción fuera de la muestra, serie Michoacán, modelo VAR(1)

Fuente: Elaboración propia con datos de SNIIM

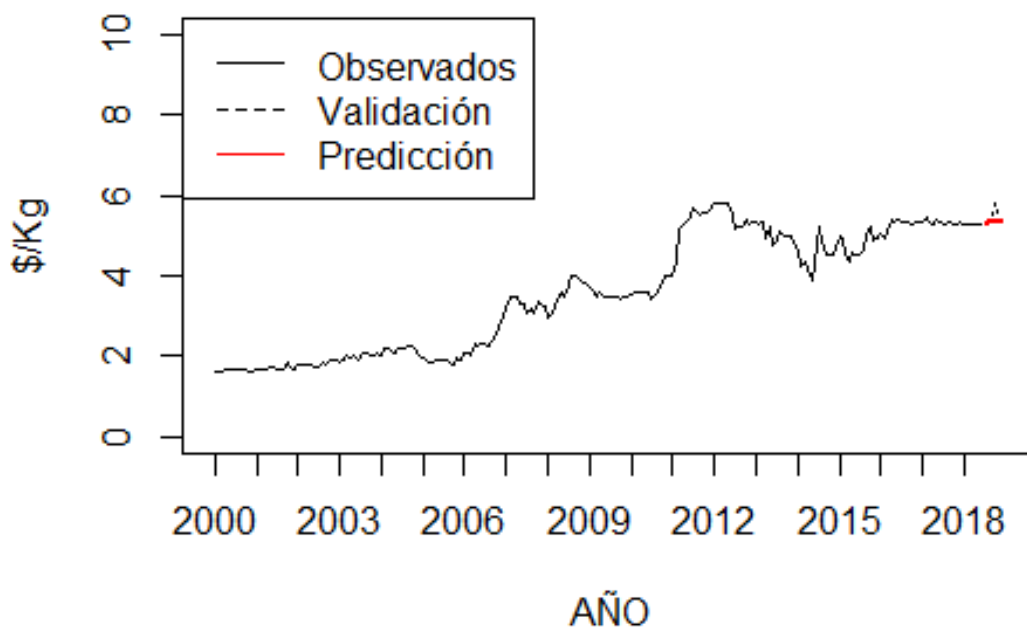


Figura 36. Predicción fuera de la muestra, serie Jalisco, modelo VAR(2)

Fuente: Elaboración propia con datos de SNIIM

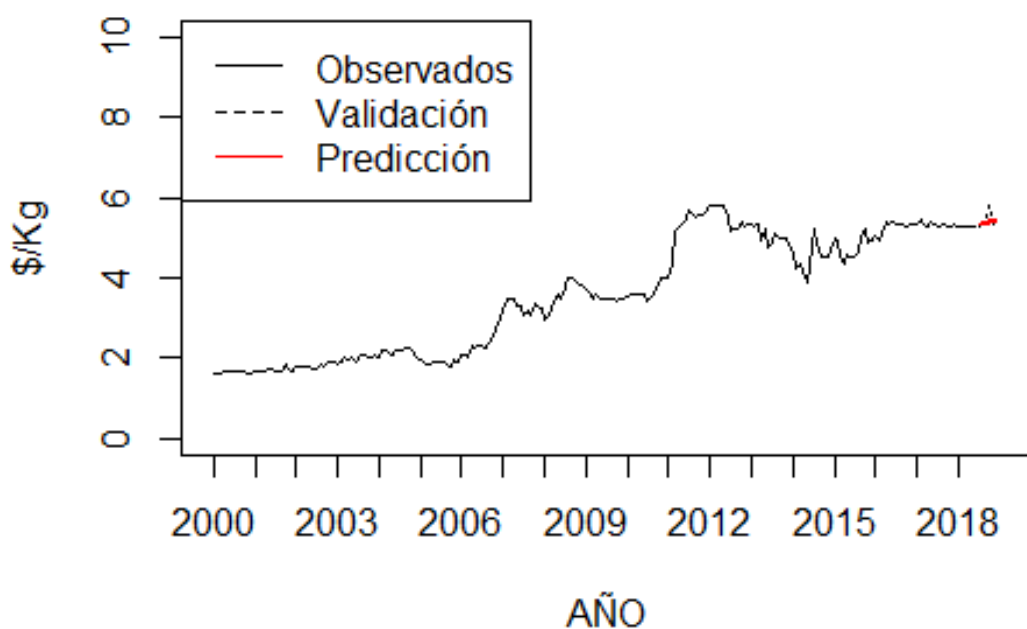


Figura 37. Predicción fuera de la muestra, serie Jalisco, modelo VAR(1)

Fuente: Elaboración propia con datos de SNIIM

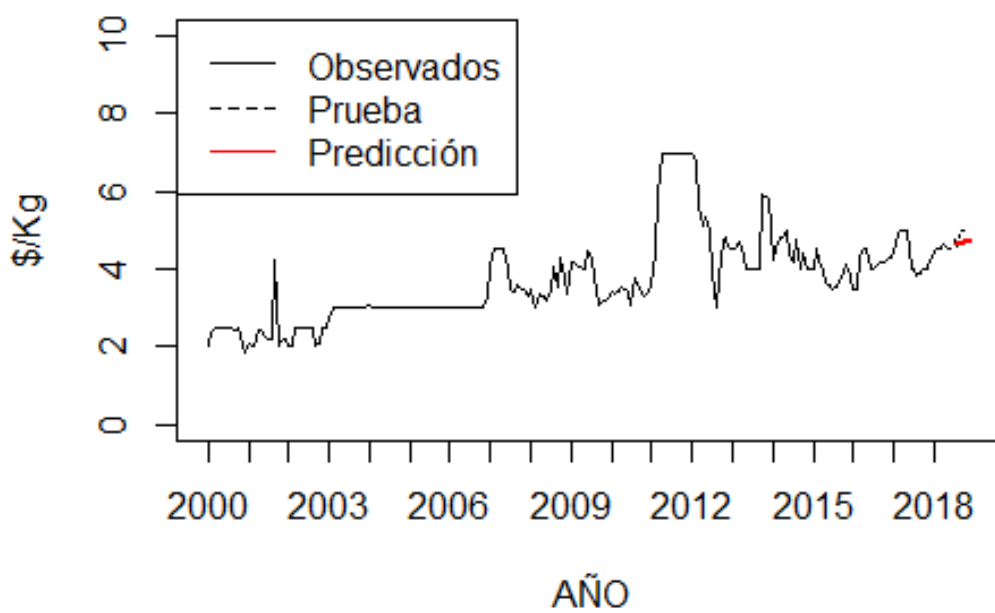


Figura 38. Predicción fuera de la muestra, serie Sinaloa, modelo VAR(2)

Fuente: Elaboración propia con datos de SNIIM

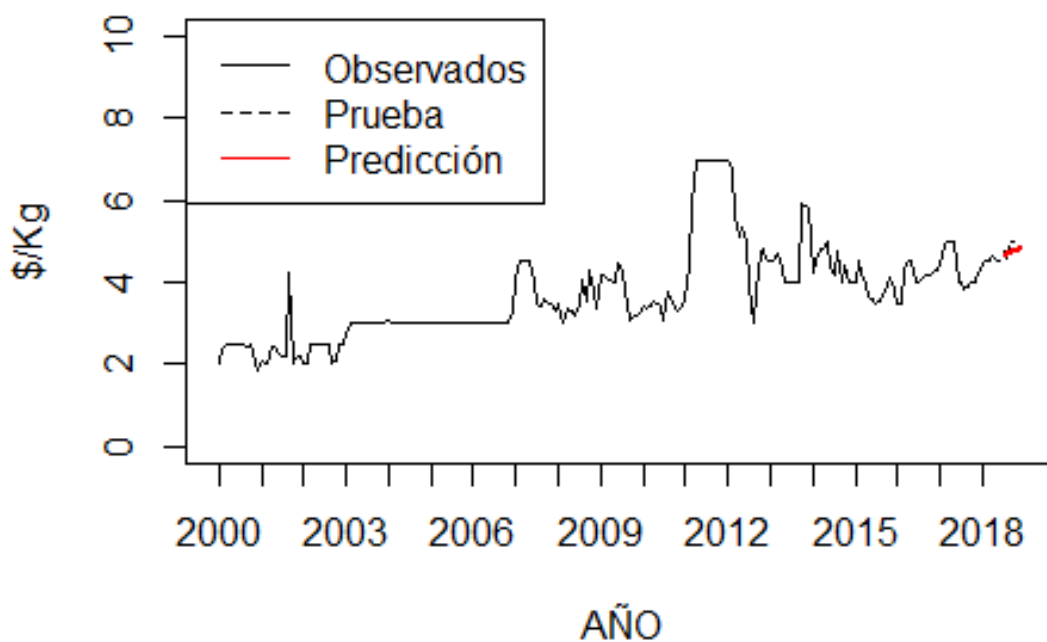


Figura 39. Predicción fuera de la muestra, serie Sinaloa, modelo VAR(1)

Fuente: Elaboración propia con datos de SNIIM



## 6.7 Evaluación de la capacidad predictiva de los modelos

Para determinar la capacidad de pronóstico de los modelos se utilizaron tres medidas: la U de Theil, la raíz del error cuadrático medio (RECM) y el error porcentual absoluto medio (MAPE) (Cuadro 21). Nótese que la U de Theil es una medida que oscila entre 0 y 1. Sí el coeficiente de Theil es igual 0, significa que el pronóstico es perfecto (exacto) y sí el coeficiente de Theil es igual a 1 no se justifica el uso del método bajo estudio.

En la serie de precios del estado de México, los modelos con mejor desempeño en la predicción fuera de la muestra fueron los modelos univariados ARIMA, donde se obtuvo el menor error porcentual absoluto medio (MAPE), que indica que la predicción está errada en un 11.32%. Estos mismos modelos obtuvieron la menor U de Theil igual a 0.06 y la menor raíz del error cuadrático medio (RECM) igual a 0.56, en comparación con los modelos multivariados VEC y VAR.

En la serie de precios del estado de Michoacán el modelo VAR(3) tuvo un error porcentual absoluto medio (MAPE) igual a 3.47, lo que significa que el pronóstico está errado en 3.47% , además, este mismo modelo obtuvo la menor U de Theil igual a 0.02 y la menor raíz del error cuadrático medio (RECM) igual a 0.4 comparado con los modelos univariados ARIMA.

De forma similar, en la serie de Sinaloa se tuvo un mejor desempeño en los modelos multivariados, teniendo los menores valores de U de Theil, RECM y MAPE en el modelo VAR(1), con valores de 0.018, 0.34 y 3.36 respectivamente.

En estos últimos dos casos se pone en evidencia que la inclusión de variables del mercado internacional si mejoró notablemente los pronósticos.

Por otra parte, en la serie de precios correspondiente al precio de maíz en el estado de Jalisco, los criterios utilizados no son determinantes en la selección de un modelo en particular, sin embargo, los modelos VEC, MA(2) y VAR(2) mostraron mejores predicciones con respecto a los otros modelos probados.

Cuadro 21. Criterios de discriminación de modelos

SERIE	MODELO	U de THEIL	RECM	MAPE
$\Delta \text{Ln}(\text{PMÉXICO})$	MA (1)	0.06358093	0.6033882	11.53
$\Delta \text{Ln}(\text{PMÉXICO})$	ARIMA (1,0,1)	0.06313138*	0.5657197*	11.32*
PMEXICO PFISICO	VEC	0.0673582	0.8829734	13.01
$\Delta \text{PMEXICO } \Delta \text{PFISICO PFUTURO}$	VAR (1)	0.06470929	0.6957609	12.05
$\Delta \text{PMEXICO } \Delta \text{PFISICO PFUTURO}$	VAR (2)	0.0665215	0.7558402	12.58
$\Delta \text{Ln}(\text{PMICHOACÁN})$	AR (1)	0.04176587	1.000933	8.37
$\Delta \text{Ln}(\text{PMICHOACÁN})$	MA (1)	0.04192237	1.005392	8.41
PMICHOACAN PFISICO	VEC	0.04189153	1.005021	8.40
$\Delta \text{PMICHOACAN } \Delta \text{PFISICO PFUTURO}$	VAR (1)	0.03826351	0.909725	7.61
$\Delta \text{PMICHOACAN } \Delta \text{PFISICO PFUTURO}$	VAR (3)	0.02039022*	0.4118365*	3.47*
$\Delta \text{Ln}(\text{PJALISCO})$	MA (2)	0.01911703	0.0169519*	2.99
$\Delta \text{Ln}(\text{PJALISCO})$	MA (1)	0.01914904	0.02830746	2.96
PJALISCO PFISICO	VEC	0.01903459*	0.0625275	2.84
$\Delta \text{PJALISCO } \Delta \text{PFISICO PFUTURO}$	VAR (1)	0.01923547	0.09628652	2.79
$\Delta \text{PJALISCO } \Delta \text{PFISICO PFUTURO}$	VAR (2)	0.01969396	0.1566225	2.74*
$\Delta \text{Ln}(\text{PSINALOA})$	ARIMA(1,0,1)	0.02801955	0.5875598	4.83
$\Delta \text{Ln}(\text{PSINALOA})$	MA (1)	0.03769274	0.8033393	8.83
PSINALOA $\Delta \text{PFISICO}$	VAR (1)	0.05585223	1.157964	9.52
PSINALOA $\Delta \text{PFISICO PFUTURO}$	VAR (1)	0.01880703*	0.3457012*	3.46*
PSINALOA $\Delta \text{PFISICO PFUTURO}$	VAR (2)	0.0202024	0.5612291	4.96

Fuente: Elaboración propia. \*Modelo con mejor desempeño según los criterios

Estos resultados respaldan la hipótesis propuesta por (Araujo, 2011) quien sostiene que existe integración entre los mercados de maíz en México y los Estados Unidos, sin embargo, dicha integración varía entre los distintos mercados regionales en México. Asimismo, los resultados apoyan la hipótesis de (Godínez y Fuentes , 2008) y Ortiz y Montiel (2017) en el caso de la serie del Estado de México, donde se prueba que los precios de futuros no se encuentran relacionados con el precio de este estado.

## **7. CONCLUSIONES**

En tres de las series analizadas los modelos multivariados VAR y VEC tuvieron un mejor desempeño en la predicción fuera de la muestra por lo que se consideran mejores predictores.

La hipótesis de la influencia de los precios de futuros de maíz y precio de maíz en físico sobre los precios en México solo se cumple para tres de las series analizadas.

La influencia de los precios de Estados Unidos no es igual en los dos estados analizados, porque en la serie de precios del Estado de México el modelo con mejor desempeño en la predicción fue el ARIMA (1,0,1), esto indica que en este caso no resulto beneficioso el uso de modelos multivariados y muestra que los modelos ARIMA también son una buena herramienta para realizar pronósticos.

La metodología de modelos univariados o multivariados no es determinante en la precisión de las predicciones obtenidas en todas las series analizadas.

Los modelos propuestos proporcionan un predictor para los precios de maíz en México y constituyen una herramienta útil en la planeación y toma de decisiones referentes al proceso productivo y comercialización de maíz y productos relacionados. Se pretende que los resultados anteriores permitan hacer un análisis de las políticas ya implementadas referentes al mercado de maíz en México, para finalmente proponer nuevas acciones.

## 8. FUENTES CONSULTADAS

- Alberto, M. (2007). *Introducción al Análisis de Series Temporales*. España: Universidad Complutense de Madrid.
- Anderson, D. R., Sweeney, D. J., y Williams, T. A. (2008). *Estadística para administración y economía* (10th ed.). México: CENGAGE Learning.
- Araujo E., S. R. (2011). Análisis de transmisión de precios entre los mercados de maíz mexicanos y el mercado estadounidense: métodos lineales y no lineales. *Revista Española de Estudios Agrosociales y Pesqueros*, 229, 39–78.
- Arias C., E., y Torres G., C. (2004). Modelos VAR y VECM para el pronóstico de corto plazo de las importaciones de Costa Rica. *Banco Central de Costa Rica*.
- Barreras S., A., Sánchez L., E., Pérez L., C., y Figueroa S., F. (2013). Uso de un modelo univariado de series de tiempo para la predicción del comportamiento de la producción de carne de cerdo en Baja California, México. *Revista Científica*, 23, 403–409.
- BMV. (2012). Futuros de maíz en MEXDER. [http://www.mexder.com.mx/wb3/wb/MEX/MEX\\_Repositorio/\\_vtp/MEX/1ede\\_baners/\\_rid/21/\\_mto/3/Folleto\\_futuro\\_maiz\\_amarillo.pdf](http://www.mexder.com.mx/wb3/wb/MEX/MEX_Repositorio/_vtp/MEX/1ede_baners/_rid/21/_mto/3/Folleto_futuro_maiz_amarillo.pdf)
- BMV. (2018). Los Futuros, uno de los instrumentos financieros más utilizados – Hablemos de Bolsa. Hablemos de Bolsa. <https://blog.bmv.com.mx/2018/09/los-futuros-uno-de-los-instrumentos-financieros-mas-utilizados/>
- Botero B., S., y Cano C., J. A. (2008). Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de Colombia. *Cuadernos de Economía*, 27(48), 173–208. [http://www.scielo.org.co/scielo.php?script=sci\\_arttext&pid=S0121-47722008000100007&lng=en&nrm=iso&tlng=es](http://www.scielo.org.co/scielo.php?script=sci_arttext&pid=S0121-47722008000100007&lng=en&nrm=iso&tlng=es)
- Bowerman, B. L., O'Connell, R. T., y Koehler, A. B. (2007). *Pronósticos, series de tiempo y regresión un enfoque aplicado* (4ta ed.). México: CENGAGE Learning.
- Box, G., Jenkins, G., Reinsel, G., and Ljung, G. (2016). *Time Series Analysis Forecasting and Control* (5ta ed.). EU: Wiley.
- Castillo T., D. (2017). Los efectos de las exportaciones y los TES sobre el EMBI para el caso colombiano un análisis utilizando modelos VEC para el periodo 2010-2016. *CIFE* 30, 47–77. <https://doi.org/http://dx.doi.org/10.15332/s0124-3551.2017.0030.02>

- CEPAL, FAO, e IICA. (2011). Volatilidad de precios en los mercados agrícolas (2000-2010): implicaciones para América Latina y opciones de políticas. Perspectivas de la agricultura y del desarrollo rural en las Américas: una mirada hacia América Latina y el Caribe. [https://repositorio.cepal.org/bitstream/handle/11362/36867/1/BoletinCepalFaolIICA1\\_2011\\_es.pdf](https://repositorio.cepal.org/bitstream/handle/11362/36867/1/BoletinCepalFaolIICA1_2011_es.pdf)
- Contreras J., A., Atziry Z., C., Martínez F., J. L., y Sánchez P., D. (2016). Análisis de series de tiempo en el pronóstico de la demanda de almacenamiento de productos perecederos. *Estudios Gerenciales*, 18, 387–396. <https://doi.org/10.1016/j.estger.2016.11.002>
- Contreras, J., Espínola, R., Nogales, F. J., y Conejo, A. J. (2003). ARIMA models to predict next-day electricity prices. *IEEE Transactions on Power Systems*, 18, 1014-1030. <https://doi.org/10.1109/TPWRS.2002.804943>
- Cryer, D. J., and Sik, K. (2008). Time series analysis with applications in R (2da ed.). EU: Springer. <https://doi.org/10.1007/BF00746534>
- Cuthbertson, K. (2002). Quantitative financial economics: stocks, bonds and foreign exchange. EU: John Wiley & Sons.
- De la Uz, N. (2002). La hipótesis de Martingala en el mercado bursátil mexicano. *Estudios Económicos*, 17, 91–127. <https://about.jstor.org/terms>
- Delgadillo R., O., Ramírez M., P. P., Leos R., J. A., Salas G., J. M., y Valdez C., R. D. (2016). Pronósticos y series de tiempo de rendimientos de granos básicos en México. *Acta Universitaria Multidisciplinary Scientific Journal*, 26, 23–32. <https://doi.org/10.15174/au.2016.882>
- Enders, W. (2015). Applied econometrics time series (4ta ed.). EU: Wiley.
- FAO. (2010). Price Volatility in Agricultural Markets: Evidence, impact on food security and policy responses. <http://www.fao.org/3/am053e/am053e00.pdf>
- Fiess, N., and Lederman, D. (2004). Mexican Corn: The Effects of NAFTA. *The World Bank Group*.
- FIRA. (2015). Panorama Agroalimentario Maíz 2015. [https://www.gob.mx/cms/uploads/attachment/file/61952/Panorama\\_Agroalimentario\\_Ma\\_z\\_2015.pdf](https://www.gob.mx/cms/uploads/attachment/file/61952/Panorama_Agroalimentario_Ma_z_2015.pdf)
- FIRA. (2016). PANORAMA AGROALIMENTARIO 2016. [https://www.gob.mx/cms/uploads/attachment/file/200637/Panorama\\_Agroalimentario\\_Ma\\_z\\_2016.pdf](https://www.gob.mx/cms/uploads/attachment/file/200637/Panorama_Agroalimentario_Ma_z_2016.pdf)

- Galindo, L. M., y Cardero, M. E. (1997). Un modelo econométrico de vectores autorregresivos y cointegración de la economía mexicana, 1980-1996. *Economía Mexicana Nueva Época*, 6, 223–247.
- Galindo, L. M., y Catalán, H. (2003). Los premios Nóbel de economía 2003: Clive W. J. Granger y R. F. Engel. *Economía Informa*, 52, 30–39.
- Gil Z., M. M., y Maya O., C. (2008). Modelación de la volatilidad de los precios de la energía eléctrica en Colombia. *Revista Ingenierías Universidad de Medellín*, 7, 87–114.
- Godínez P., J. A. (2006). Causalidad del precio futuro de la Bolsa de Chicago sobre los precios físicos de maíz blanco en México. *Estudios Sociales*, 15, 203–223.
- Godínez P., J. A., y Fuentes F., N. A. (2008). Las condiciones económicas para operar un mercado de futuros de maíz blanco en México. *Investigación Económica*, 67, 15–37.
- Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, 48, 213–228. <https://doi.org/10.1111/j.1468-0084.1986.mp48003002.x>
- Guerrero, V. M. (2007). Pronósticos restringidos con modelos de series de tiempo múltiples y su aplicación para evaluar metas de política macroeconómica en México. *Estudios Económicos*, 22, 241–311.
- Gujarati, D. N., y Porter, D. C. (2010). *Econometría* (5ta ed.). México: McGrawHillEducacion.
- Gutiérrez R., J., y Vásquez E., D. M. (2008). Un análisis de cointegración para el riesgo de crédito. in reporte de estabilidad financiera. <http://repositorio.banrep.gov.co/handle/20.500.12134/2124>
- Hull, J. C. (2009). *Introducción a los mercados de futuros y opciones* (6ta ed.). México: Pearson Prentice Hall.
- Kaasra, I., and Boyd, M. (1996). Designing a Neural Network for Forecasting Financial Time Series. *Neurocomputing*, 10, 215–236.
- Lanteri, L. N. (2010). Modelos de VAR alternativos para pronósticos (VAR bayesianos y FAVAR): el caso de las explotaciones argentinas. *Economía*, 33, 42–64.
- Loría, E. (2007). *Econometría con aplicaciones*. México: Pearson Prentice Hall. <https://doi.org/10.1017/CBO9781107415324.004>

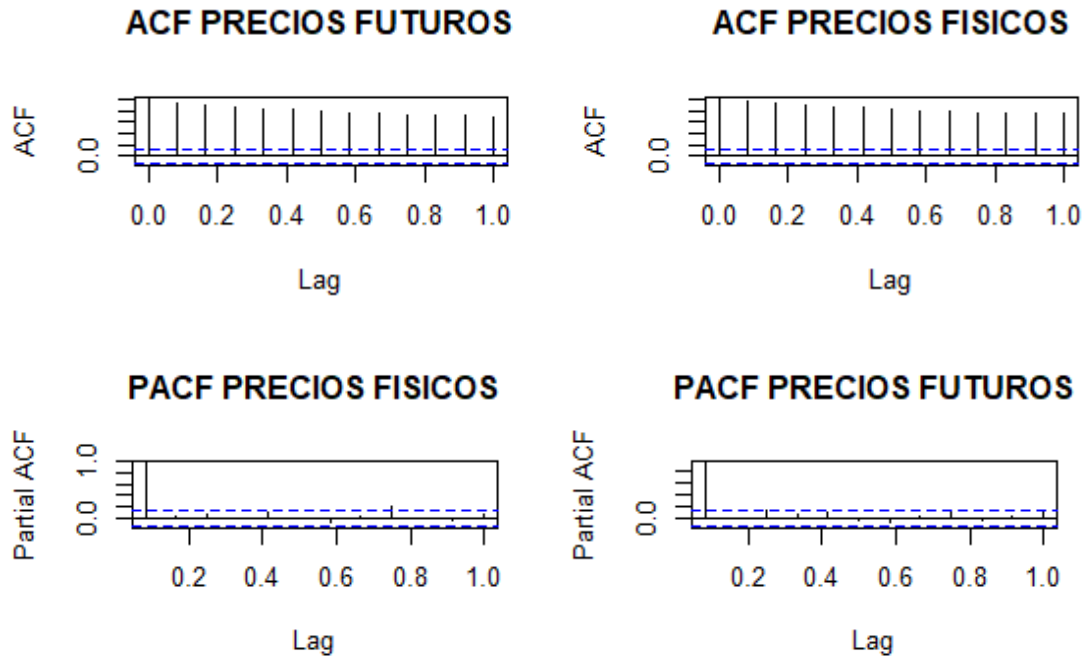
- Lütkepohl, H. (2005). New introduction to multiple time series analysis. EU: Springer.  
<https://doi.org/10.1007/978-3-540-27752-1>
- Makridakis, S., Wheelwright, S. C., and Hyndman, R. J. (1997). Forecasting, methods and applications (3ra ed.). Jhon Wiley and Sons.
- Maparu, T. S., and Mazumder, T. N. (2017). Transport infrastructure, economic development and urbanization in India (1990–2011): Is there any causal relationship? *Transportation Research Part A: Policy and Practice*, 100, 319–336.  
<https://doi.org/10.1016/j.tra.2017.04.033>
- Marroquín M., G., y Chalita T., L. E. (2010). Aplicación de la metodología Box-Jenkins para pronóstico de precios en jitomate. *Revista Mexicana de Ciencias Agrícolas*, 2, 573–577. <http://www.redalyc.org/articulo.oa?id=263119723008>
- Mercado P., D., Pedraza C., L., y Martínez G., E. (2015). Comparación de Redes Neuronales aplicadas a la predicción de Series de Tiempo. *Prospect*, 13, 88–95.  
<https://doi.org/10.15665/rp.v13i2.491>
- Morán C., D. M. (2014). Determinantes de la inflación en Ecuador un análisis econométrico utilizando modelos VAR. *Economía y Sociedad*, 31, 53–70.  
<http://revistas.uazuay.edu.ec/index.php/udaakadem/article/view/202>
- Morana, C. (2001). A semiparametric approach to short-term oil price forecasting. *Energy Economics*, 23, 325–338. [https://doi.org/10.1016/S0140-9883\(00\)00075-X](https://doi.org/10.1016/S0140-9883(00)00075-X)
- Nájera S., J. A. (2019). Relaciones en el comportamiento de los precios de las criptomonedas: un análisis econométrico a través de modelos VAR y VEC. *Estocástica: Finanzas y Riesgo*, 9, 33–61.
- Ortiz A., F., y Montiel G., A. N. (2017). Transmisión de precios futuros de maíz del Chicago Board of Trade al mercado spot mexicano. *Contaduría y Administración*, 62, 924–940. <https://doi.org/10.1016/j.cya.2016.01.004>
- Pankratz, A. (1983). Forecasting with univariate Box-Jenkins models concepts and cases. EU: Jhon Wiley and Sons.
- Peña, D. (2010). Análisis de series temporales. España: Alianza Editorial.
- Rodríguez P., H. Y., Agudelo G., S., Rincón, L., y Montañez, K. (2019). Inflación importada vía tipos de cambio y su aplicación al caso colombiano mediante el uso de modelos VEC para el periodo enero 2010 - diciembre 2018. *Comunicaciones en Estadística*, 12, 33–52.

- Rodríguez R., C. A. (2002). Análisis de la economía de Puerto Rico con un modelo de vectores autorregresivos y cointegración. *Revista de Ciencias Sociales*, 11, 91–110.
- SAGARPA. (2017). PLANEACIÓN AGRÍCOLA NACIONAL 2017-2030. [https://www.gob.mx/cms/uploads/attachment/file/256429/B\\_sico-Ma\\_z\\_Grano\\_Blanco\\_y\\_Amarillo.pdf](https://www.gob.mx/cms/uploads/attachment/file/256429/B_sico-Ma_z_Grano_Blanco_y_Amarillo.pdf)
- Sánchez L., E., Barreras S., A., Pérez L., C., Fernando F., S., y Olivas V., J. A. (2013). aplicación de un modelo ARIMA para pronosticar la producción de leche de bovino en baja california, México. *Tropical and Subtropical Agroecosystems*, 16, 315–324.
- SIAP. (2019). PANORAMA AGROALIMENTARIO 2019. [https://nube.siap.gob.mx/gobmx\\_publicaciones\\_siap/pag/2019/Atlas-Agroalimentario-2019](https://nube.siap.gob.mx/gobmx_publicaciones_siap/pag/2019/Atlas-Agroalimentario-2019)
- Velásquez H., J. D., y Franco C., C. J. (2010). Predicción de los precios de contratos de electricidad usando una red neuronal con arquitectura dinámica. *INNOVAR JOURNAL*, 20, 7–14.
- Zavaleta V., O. H. (2019). Impacto de la introducción del contrato futuro del maíz amarillo en la volatilidad del precio spot: evidencia del MexDer. *Contaduría y Administración*, 65, 1–33.



## 9. APÉNDICE

### Apéndice 1. ACF Y PACF de las series originales



### Apéndice 2. Verificación de ruido blanco en residuales en modelos univariados

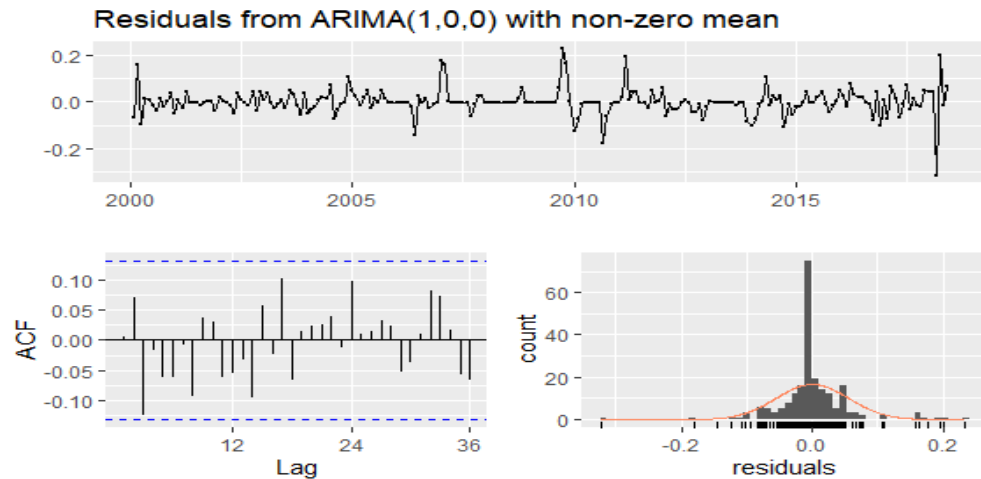


Figura. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie MICHOACÁN.

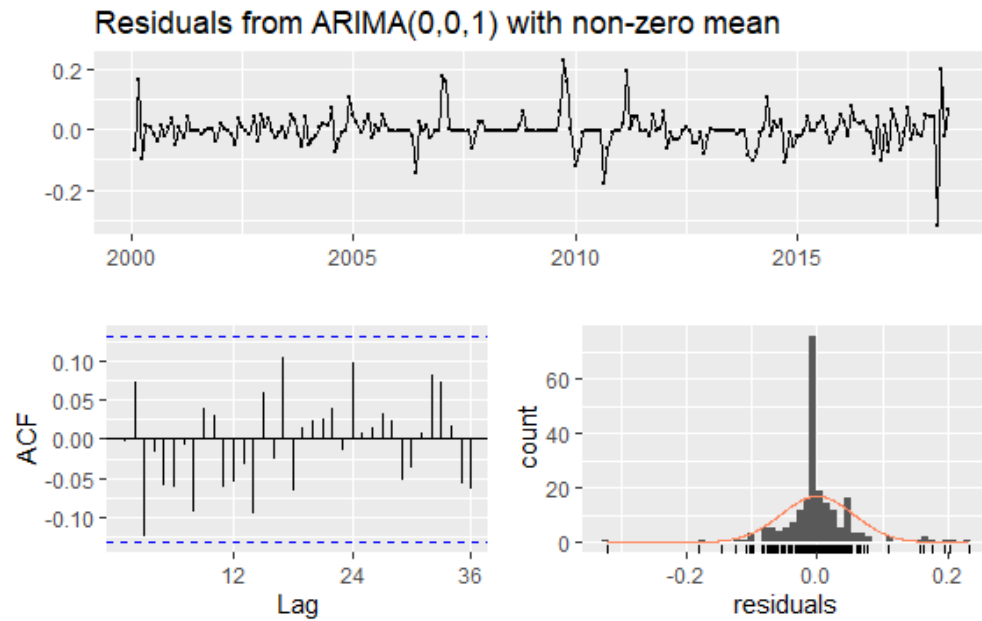
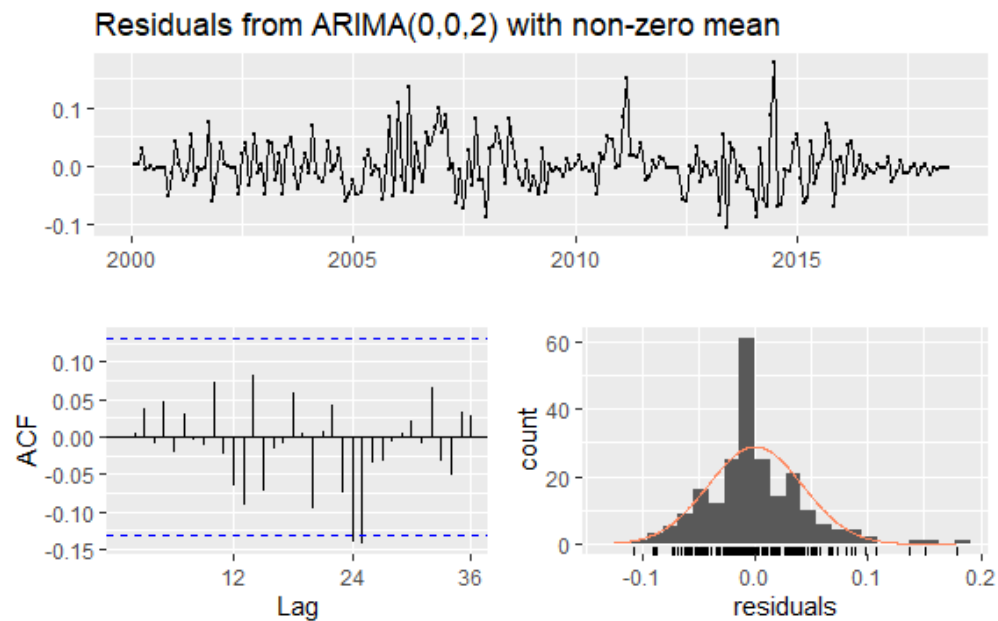


Figura. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie JALISCO.



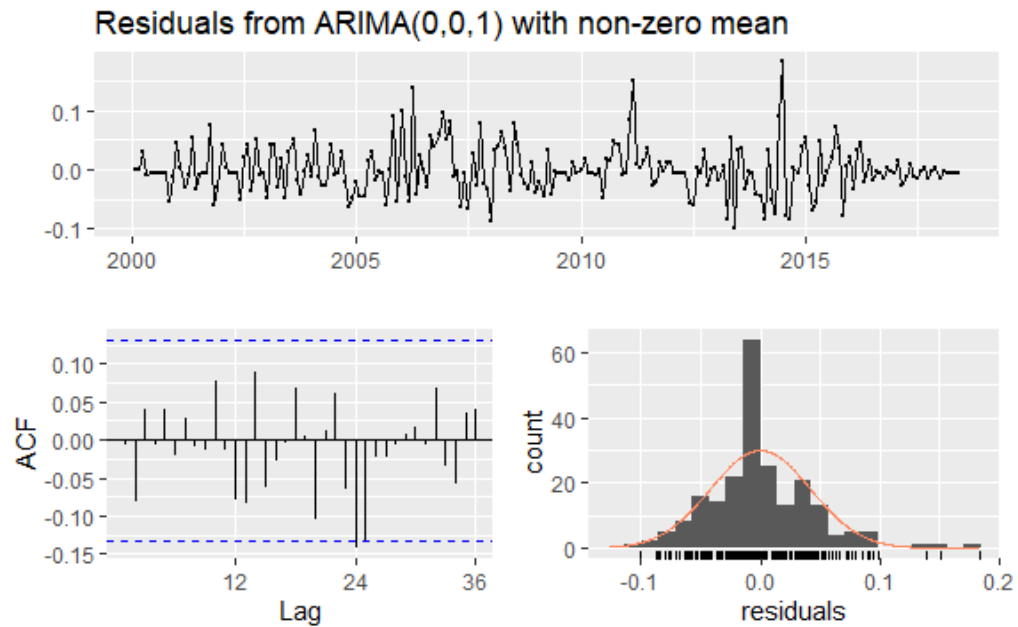


Figura. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie JALISCO.

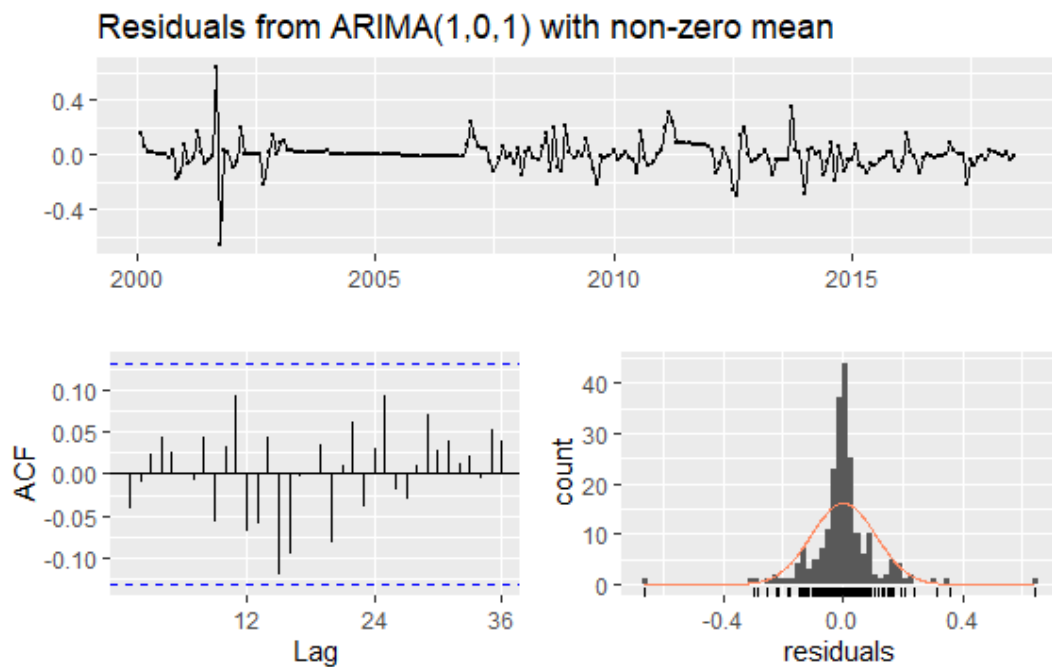


Figura. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie SINALOA

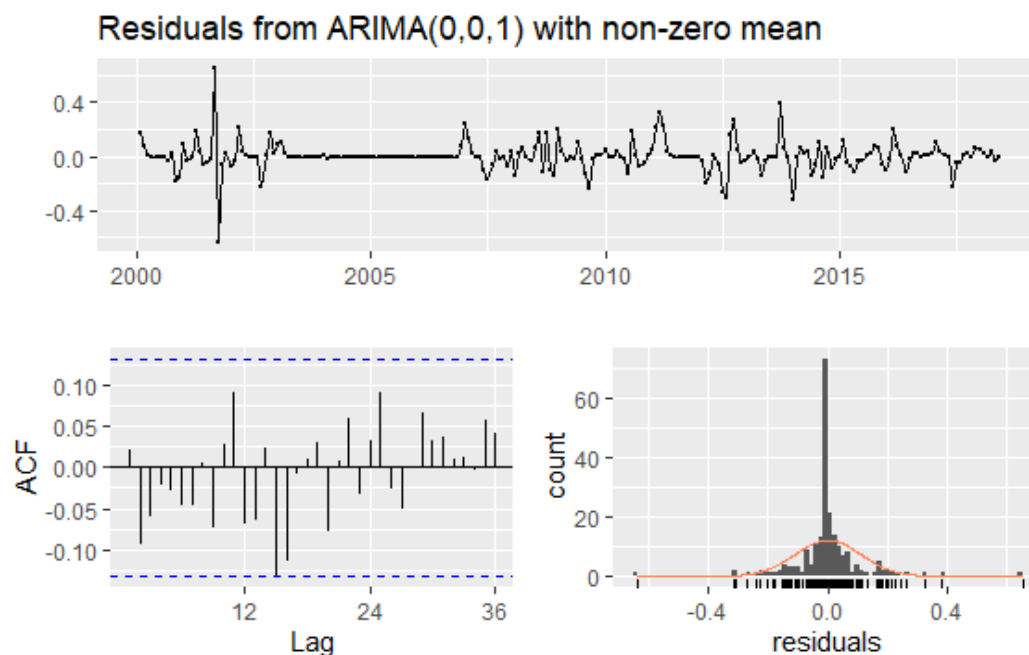


Figura. Gráfico de residuales, función de autocorrelación parcial y gráfico de distribución normal del modelo ajustado a la serie SINALOA

### Apéndice 3. Coeficientes estimados de modelos VEC

```

MODELO VEC: PMEXICO PFISICO
summary(VEC_MEX_FIS$rlm)
Response serieMEXICO.in.d :

Call:
lm(formula = serieMEXICO.in.d ~ ect1 + constant + serieMEXICO.in.d11 +
    PRECIOSFISICOS.in.d11 - 1, data = data.mat)

Residuals:
    Min       1Q   Median       3Q      Max
-1.46895 -0.06908  0.00176  0.07871  1.02755

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
ect1          -0.06472   0.02057  -3.146  0.00189 **
constant       0.04577   0.01758   2.603  0.00987 **
serieMEXICO.in.d11 -0.10205   0.06668  -1.531  0.12733
PRECIOSFISICOS.in.d11 -0.03119   0.06861  -0.455  0.64989
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2169 on 216 degrees of freedom
Multiple R-squared:  0.06188, Adjusted R-squared:  0.04451
F-statistic: 3.562 on 4 and 216 DF, p-value: 0.00775

Response PRECIOSFISICOS.in.d :

```

```
Call:
lm(formula = PRECIOSFISICOS.in.d ~ ect1 + constant + serieMEXICO.in.d
    11 +
    PRECIOSFISICOS.in.d11 - 1, data = data.mat)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-1.36233	-0.07612	-0.00810	0.06971	0.79745

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
ect1	0.065366	0.020983	3.115	0.00209	**
constant	-0.019752	0.017936	-1.101	0.27202	
serieMEXICO.in.d11	-0.002809	0.068014	-0.041	0.96709	
PRECIOSFISICOS.in.d11	0.011095	0.069993	0.159	0.87420	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2213 on 216 degrees of freedom  
Multiple R-squared: 0.04826, Adjusted R-squared: 0.03064  
F-statistic: 2.738 on 4 and 216 DF, p-value: 0.02973

```
summary(VEC_MICH_FIS$rlm)
Response serieMICHACAN.in.d :
```

```
Call:
lm(formula = serieMICHACAN.in.d ~ ect1 + constant + serieMICHACAN.i
    n.d11 +
    PRECIOSFISICOS.in.d11 - 1, data = data.mat)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-1.48588	-0.04966	-0.01163	0.05474	1.07144

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
ect1	-0.044625	0.032261	-1.383	0.1680	
constant	0.046663	0.026410	1.767	0.0787	.
serieMICHACAN.in.d11	0.003877	0.072173	0.054	0.9572	
PRECIOSFISICOS.in.d11	-0.058572	0.073537	-0.796	0.4266	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2319 on 216 degrees of freedom  
Multiple R-squared: 0.01556, Adjusted R-squared: -0.002673  
F-statistic: 0.8534 on 4 and 216 DF, p-value: 0.4929

Response PRECIOSFISICOS.in.d :

```
Call:
lm(formula = PRECIOSFISICOS.in.d ~ ect1 + constant + serieMICHACAN.i
    n.d11 +
    PRECIOSFISICOS.in.d11 - 1, data = data.mat)
```

Residuals:

	Min	1Q	Median	3Q	Max
	-1.36343	-0.08228	0.00720	0.07490	0.70326

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
ect1	0.16372	0.02929	5.589	6.84e-08	***
constant	-0.09776	0.02398	-4.077	6.43e-05	***
serieMICHOCAN.in.d11	-0.02756	0.06553	-0.421	0.675	
PRECIOSFISICOS.in.d11	0.04315	0.06677	0.646	0.519	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2106 on 216 degrees of freedom  
Multiple R-squared: 0.1379, Adjusted R-squared: 0.122  
F-statistic: 8.639 on 4 and 216 DF, p-value: 1.74e-06

[summary\(VEC\\_JAL\\_FIS\\$rlm\)](#)  
Response serieJALISCO.in.d :

Call:  
lm(formula = serieJALISCO.in.d ~ ect1 + constant + serieJALISCO.in.d11 +  
PRECIOSFISICOS.in.d11 - 1, data = data.mat)

Residuals:

Min	1Q	Median	3Q	Max
-0.47547	-0.07265	-0.01263	0.06866	0.88743

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
ect1	-0.02424	0.01584	-1.531	0.1273	
constant	0.02074	0.01140	1.820	0.0701	.
serieJALISCO.in.d11	0.12227	0.06798	1.799	0.0735	.
PRECIOSFISICOS.in.d11	-0.03515	0.04981	-0.706	0.4812	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1586 on 216 degrees of freedom  
Multiple R-squared: 0.03317, Adjusted R-squared: 0.01527  
F-statistic: 1.853 on 4 and 216 DF, p-value: 0.1199

Response PRECIOSFISICOS.in.d :

Call:  
lm(formula = PRECIOSFISICOS.in.d ~ ect1 + constant + serieJALISCO.in.d11 +  
PRECIOSFISICOS.in.d11 - 1, data = data.mat)

Residuals:

Min	1Q	Median	3Q	Max
-1.29187	-0.07969	-0.01234	0.06935	0.73317

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )	
ect1	0.08793	0.02161	4.069	6.63e-05	***
constant	-0.01234	0.01555	-0.793	0.428	
serieJALISCO.in.d11	0.11967	0.09276	1.290	0.198	
PRECIOSFISICOS.in.d11	0.01535	0.06797	0.226	0.822	

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2164 on 216 degrees of freedom  
Multiple R-squared: 0.0894, Adjusted R-squared: 0.07254  
F-statistic: 5.302 on 4 and 216 DF, p-value: 0.0004314

summary(VAR\_SINALOA\_FIS)

VAR Estimation Results:

=====

Endogenous variables: SINA, diff.PRECIOFISICOS.in.  
Deterministic variables: const  
Sample size: 220  
Log Likelihood: -107.506  
Roots of the characteristic polynomial:  
0.9238 0.05471  
Call:  
VAR(y = SINALOA\_FIS, p = 1)

Estimation results for equation SINA:

=====

SINA = SINA.l1 + diff.PRECIOFISICOS.in..l1 + const

	Estimate	Std. Error	t value	Pr(> t )	
SINA.l1	0.92436	0.02564	36.047	< 2e-16	***
diff.PRECIOFISICOS.in..l1	0.28363	0.12904	2.198	0.02901	*
const	0.27879	0.09973	2.795	0.00565	**

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4288 on 217 degrees of freedom  
Multiple R-Squared: 0.8573, Adjusted R-squared: 0.856  
F-statistic: 651.9 on 2 and 217 DF, p-value: < 2.2e-16

Estimation results for equation diff.PRECIOFISICOS.in.:

=====

diff.PRECIOFISICOS.in. = SINA.l1 + diff.PRECIOFISICOS.in..l1 + const

	Estimate	Std. Error	t value	Pr(> t )
SINA.l1	-0.001851	0.013498	-0.137	0.891
diff.PRECIOFISICOS.in..l1	-0.055243	0.067923	-0.813	0.417
const	0.018155	0.052493	0.346	0.730

Residual standard error: 0.2257 on 217 degrees of freedom  
Multiple R-Squared: 0.003122, Adjusted R-squared: -0.006066  
F-statistic: 0.3398 on 2 and 217 DF, p-value: 0.7123

Covariance matrix of residuals:

	SINA	diff.PRECIOFISICOS.in.
SINA	0.1838388	0.0005892
diff.PRECIOFISICOS.in.	0.0005892	0.0509336

Correlation matrix of residuals:

	SINA	diff.PRECIOFISICOS.in.
SINA	1.000000	0.006089
diff.PRECIOFISICOS.in.	0.006089	1.000000

#### Apéndice 4. Diagnóstico de residuales en modelos VAR

Diagram of fit and residuals for diff.serieMEXICO.in.

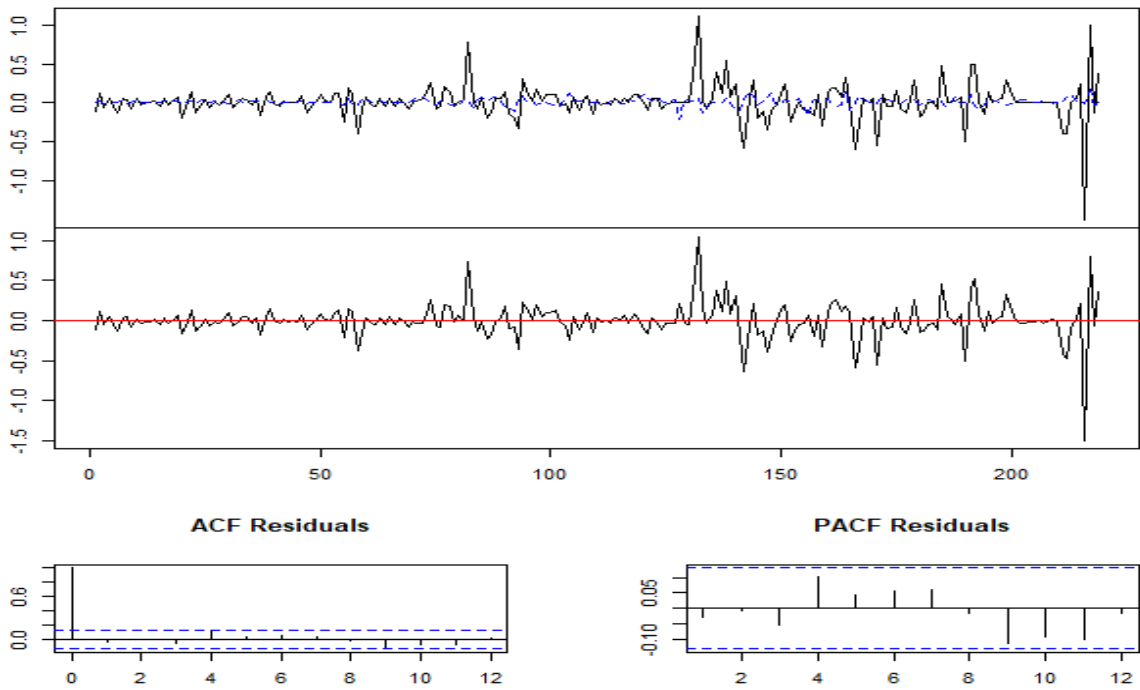


Figura. Análisis de residuales de modelos VAR para el precio de maíz en el Estado de México

VAR\_MEX\_FIS\_FUTP\_2

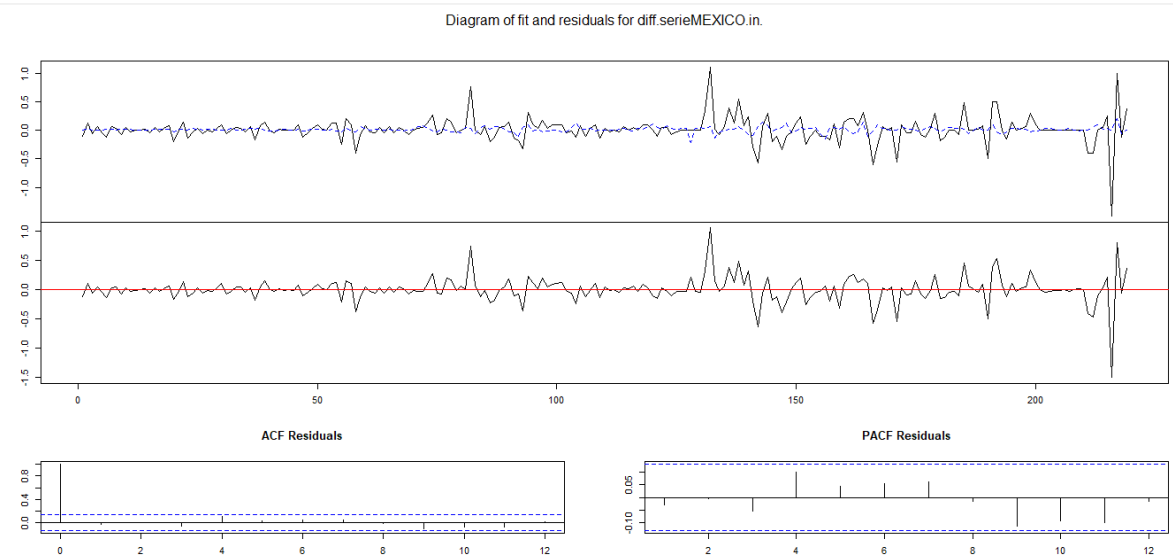


Figura. Análisis de residuales de modelos VAR para el precio de maíz en el estado de México



## VAR\_MICH\_FIS\_FUT\_P1

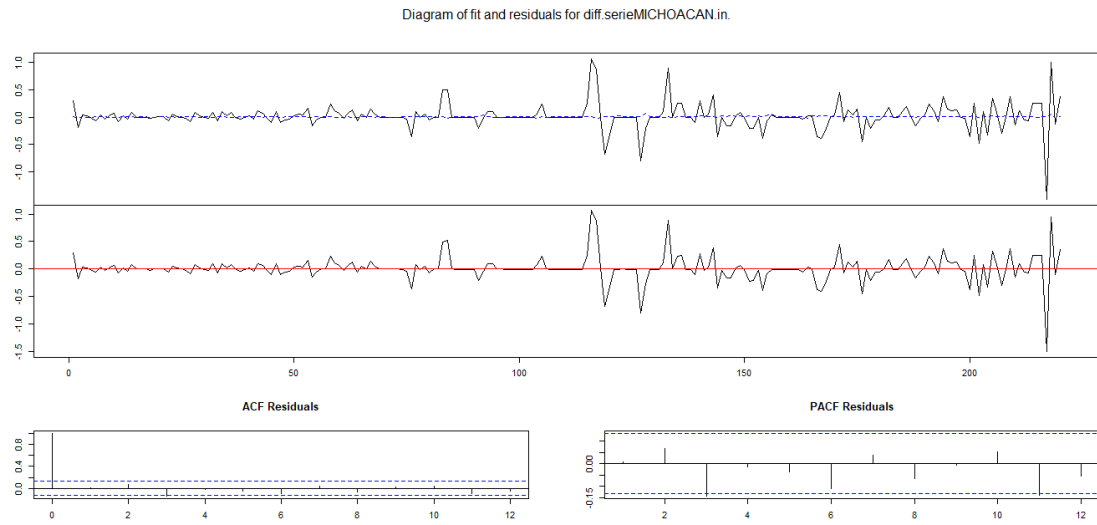


Figura. Análisis de residuales de modelos VAR para el precio de maíz en el Estado de Michoacán

## VAR\_JAL\_FIS\_FUT\_P1

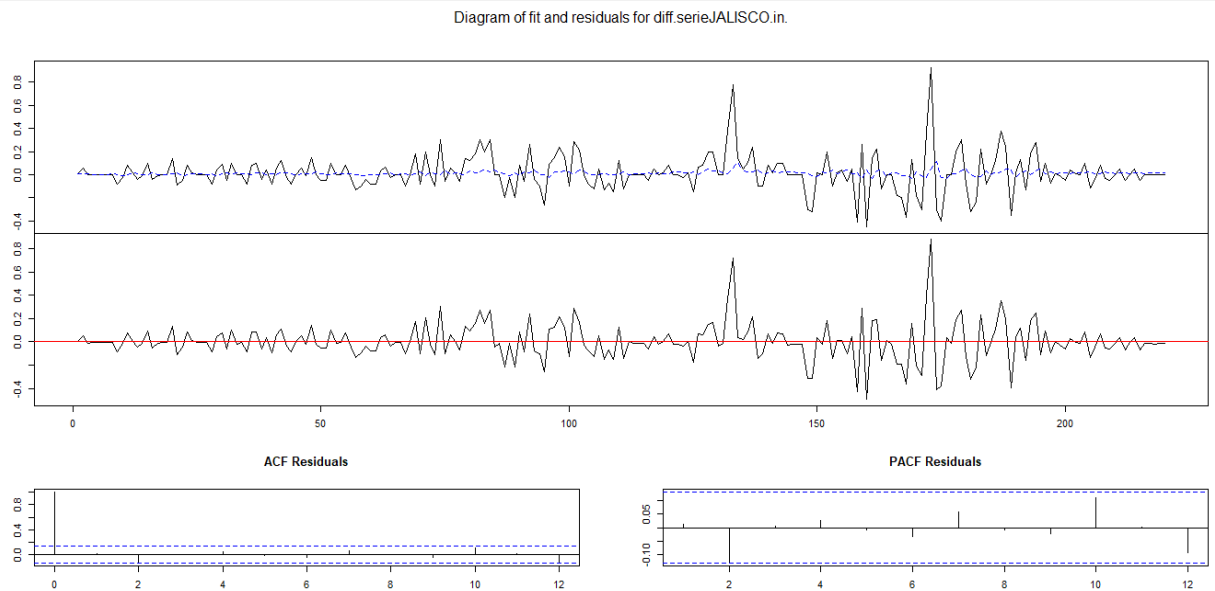


Figura. Análisis de residuales de modelos VAR para el precio de maíz en el estado de Jalisco

## VAR\_SINA\_FIS\_FUT\_P1

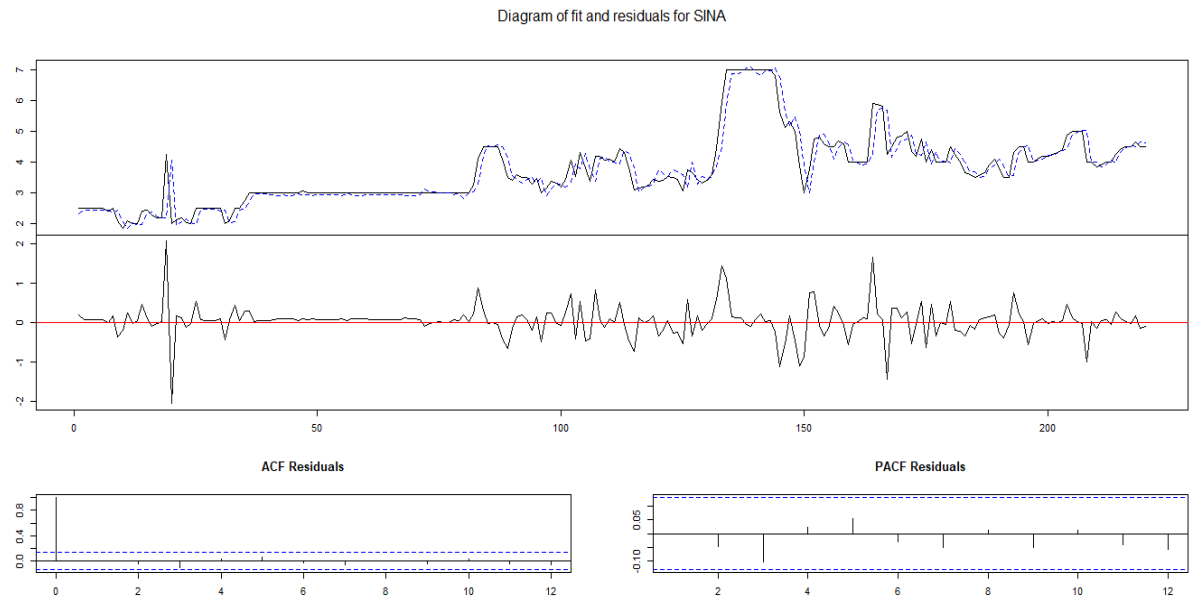


Figura. Análisis de residuales de modelos VAR para el precio de maíz en el estado de Sinaloa

## Apéndice 5. Coeficientes estimados en modelos VAR

[summary\(VAR\\_MEX\\_FIS\\_FUT\)#P=2](#)

VAR Estimation Results:

=====  
Endogenous variables: diff.serieMEXICO.in., diff.PRECIOFISICOS.in., PFUT

Deterministic variables: const

Sample size: 219

Log Likelihood: 167.811

Roots of the characteristic polynomial:

0.9641 0.3946 0.3946 0.3296 0.08457 0.01357

Call:

VAR(y = MEX\_FIS\_FUT, p = 2)

Estimation results for equation diff.serieMEXICO.in.:

=====  
diff.serieMEXICO.in. = diff.serieMEXICO.in..l1 + diff.PRECIOFISICOS.in..l1 + PFUT.l1 + diff.serieMEXICO.in..l2 + diff.PRECIOFISICOS.in..l2 + PFUT.l2 + const

	Estimate	Std. Error	t value	Pr(> t )
diff.serieMEXICO.in..l1	-0.12989	0.06820	-1.905	0.0582 .
diff.PRECIOFISICOS.in..l1	-0.02850	0.12595	-0.226	0.8212
PFUT.l1	0.07380	0.10626	0.695	0.4881
diff.serieMEXICO.in..l2	-0.04792	0.06831	-0.702	0.4837
diff.PRECIOFISICOS.in..l2	-0.15670	0.06790	2.308	0.0220 *
PFUT.l2	-0.07069	0.10561	-0.669	0.5040
const	0.00847	0.03637	0.233	0.8161

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2208 on 212 degrees of freedom  
Multiple R-Squared: 0.04225, Adjusted R-squared: 0.01515  
F-statistic: 1.559 on 6 and 212 DF, p-value: 0.1606

`summary(VAR_MICHOACAN_FIS_FUT)#P=3`

VAR Estimation Results:

=====

Endogenous variables: diff.serieMICHOACAN.in., diff.PRECIOFISICOS.in., PFUT

Deterministic variables: const

Sample size: 218

Log Likelihood: 179.211

Roots of the characteristic polynomial:

0.9745 0.5938 0.5938 0.5298 0.5298 0.5157 0.4611 0.4611 0.3753

Call:

VAR(y = MICHOACAN\_FIS\_FUT, p = 3)

Estimation results for equation diff.serieMICHOACAN.in.:

=====

diff.serieMICHOACAN.in. = diff.serieMICHOACAN.in..l1 + diff.PRECIOFISICOS.in..l1 + PFUT.l1 + diff.serieMICHOACAN.in..l2 + diff.PRECIOFISICOS.in..l2 + PFUT.l2 + diff.serieMICHOACAN.in..l3 + diff.PRECIOFISICOS.in..l3 + PFUT.l3 + const

	Estimate	Std. Error	t value	Pr(> t )
diff.serieMICHOACAN.in..l1	-0.01986	0.07009	-0.283	0.7772
diff.PRECIOFISICOS.in..l1	0.10591	0.13685	0.774	0.4399
PFUT.l1	-0.13420	0.11378	-1.179	0.2396
diff.serieMICHOACAN.in..l2	0.05617	0.06950	0.808	0.4199
diff.PRECIOFISICOS.in..l2	0.11301	0.13744	0.822	0.4119
PFUT.l2	0.10471	0.15367	0.681	0.4964
diff.serieMICHOACAN.in..l3	-0.15503	0.07254	-2.137	0.0337 *
diff.PRECIOFISICOS.in..l3	-0.03294	0.07250	-0.454	0.6501
PFUT.l3	0.01229	0.11546	0.106	0.9153
const	0.05061	0.03869	1.308	0.1923

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2304 on 208 degrees of freedom  
Multiple R-Squared: 0.04991, Adjusted R-squared: 0.008801  
F-statistic: 1.214 on 9 and 208 DF, p-value: 0.2878

`summary(VAR_JALISCO_FIS_FUT)#p=2`

VAR Estimation Results:

=====

Endogenous variables: diff.serieJALISCO.in., diff.PRECIOFISICOS.in., PFUT

Deterministic variables: const

Sample size: 219

Log Likelihood: 247.353

Roots of the characteristic polynomial:  
0.9654 0.4081 0.4081 0.2666 0.2666 0.172

Call:

VAR(y = JALISCO\_FIS\_FUT, p = 2)

Estimation results for equation diff.serieJALISCO.in.:

=====

diff.serieJALISCO.in. = diff.serieJALISCO.in..l1 + diff.PRECIOFISICOS.in..l1 + PFUT.l1 + diff.serieJALISCO.in..l2 + diff.PRECIOFISICOS.in..l2 + PFUT.l2 + const

	Estimate	Std. Error	t value	Pr(> t )
diff.serieJALISCO.in..l1	0.12352	0.06877	1.796	0.0739 .
diff.PRECIOFISICOS.in..l1	0.01260	0.09137	0.138	0.8904
PFUT.l1	-0.01774	0.07745	-0.229	0.8190
diff.serieJALISCO.in..l2	-0.12919	0.06868	-1.881	0.0614 .
diff.PRECIOFISICOS.in..l2	-0.02289	0.04910	-0.466	0.6416
PFUT.l2	0.01676	0.07699	0.218	0.8279
const	0.01899	0.02628	0.723	0.4707

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.1595 on 212 degrees of freedom

Multiple R-Squared: 0.03012, Adjusted R-squared: 0.002669

F-statistic: 1.097 on 6 and 212 DF, p-value: 0.3651

`summary(VAR_SINALOA_FIS_FUT)#p=2`

VAR Estimation Results:

=====

Endogenous variables: SINA, diff.PRECIOFISICOS.in., PFUT

Deterministic variables: const

Sample size: 219

Log Likelihood: 41.061

Roots of the characteristic polynomial:  
0.9674 0.8213 0.3226 0.3226 0.2758 0.05714

Call:

VAR(y = SINALOA\_FIS\_FUT, p = 2)

Estimation results for equation SINA:

=====

SINA = SINA.l1 + diff.PRECIOFISICOS.in..l1 + PFUT.l1 + SINA.l2 + diff.PRECIOFISICOS.in..l2 + PFUT.l2 + const

	Estimate	Std. Error	t value	Pr(> t )
SINA.l1	0.89669	0.07000	12.810	< 2e-16 ***
diff.PRECIOFISICOS.in..l1	0.19753	0.23127	0.854	0.39400

PFUT.l1	0.17018	0.19695	0.864	0.38853
SINA.l2	-0.03971	0.07015	-0.566	0.57191
diff.PRECIOFISICOS.in..l2	-0.04613	0.12693	-0.363	0.71667
PFUT.l2	-0.04904	0.19583	-0.250	0.80252
const	0.30224	0.09632	3.138	0.00194 **

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4018 on 212 degrees of freedom

Multiple R-Squared: 0.8763, Adjusted R-squared: 0.8728

F-statistic: 250.4 on 6 and 212 DF, p-value: < 2.2e-16

`summary(VAR_MEX_FIS_FUTP1)#P=1`

VAR Estimation Results:

=====

Endogenous variables: diff.serieMEXICO.in., diff.PRECIOFISICOS.in., PFUT

Deterministic variables: none

Sample size: 220

Log Likelihood: 156.47

Roots of the characteristic polynomial:

0.9987 0.127 0.04757

Call:

VAR(y = MEX\_FIS\_FUT, p = 1, type = "none")

Estimation results for equation diff.serieMEXICO.in.:

=====

diff.serieMEXICO.in. = diff.serieMEXICO.in..l1 + diff.PRECIOFISICOS.in..l1 + PFUT.l1

	Estimate	Std. Error	t value	Pr(> t )
diff.serieMEXICO.in..l1	-0.119777	0.067810	-1.766	0.0787
diff.PRECIOFISICOS.in..l1	0.030029	0.066906	0.449	0.6540
PFUT.l1	0.007822	0.006980	1.121	0.2636

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2212 on 217 degrees of freedom

Multiple R-Squared: 0.01988, Adjusted R-squared: 0.006326

F-statistic: 1.467 on 3 and 217 DF, p-value: 0.2245

`summary(VAR_MICHOACAN_FIS_FUTP1)##P=1`

VAR Estimation Results:

=====

Endogenous variables: diff.serieMICHOACAN.in., diff.PRECIOFISICOS.in., PFUT

Deterministic variables: none

Sample size: 220

Log Likelihood: 149.263

Roots of the characteristic polynomial:

0.9989 0.06561 0.06561

Call:

```
VAR(y = MICHOACAN_FIS_FUT, p = 1, type = "none")
```

```
Estimation results for equation diff.serieMICHOACAN.in.:
```

```
=====
diff.serieMICHOACAN.in. = diff.serieMICHOACAN.in..l1 + diff.PRECIOFISICOS.in..l1 + PFUT.l1
```

	Estimate	Std. Error	t value	Pr(> t )
diff.serieMICHOACAN.in..l1	-0.024100	0.068681	-0.351	0.726
diff.PRECIOFISICOS.in..l1	-0.030627	0.070911	-0.432	0.666
PFUT.l1	0.004683	0.007334	0.639	0.524

```
Residual standard error: 0.2328 on 217 degrees of freedom
Multiple R-Squared: 0.003183, Adjusted R-squared: -0.0106
F-statistic: 0.231 on 3 and 217 DF, p-value: 0.8748
```

```
summary(VAR_JALISCO_FIS_FUTP1)##P=1
```

```
VAR Estimation Results:
```

```
=====
Endogenous variables: diff.serieJALISCO.in., diff.PRECIOFISICOS.in., PFUT
```

```
Deterministic variables: none
```

```
Sample size: 220
```

```
Log Likelihood: 232.466
```

```
Roots of the characteristic polynomial:
```

```
0.9985 0.09033 0.0419
```

```
Call:
```

```
VAR(y = JALISCO_FIS_FUT, p = 1, type = "none")
```

```
Estimation results for equation diff.serieJALISCO.in.:
```

```
=====
diff.serieJALISCO.in. = diff.serieJALISCO.in..l1 + diff.PRECIOFISICOS.in..l1 + PFUT.l1
```

	Estimate	Std. Error	t value	Pr(> t )
diff.serieJALISCO.in..l1	0.110358	0.067655	1.631	0.104
diff.PRECIOFISICOS.in..l1	-0.016895	0.048227	-0.350	0.726
PFUT.l1	0.005899	0.005041	1.170	0.243

```
Residual standard error: 0.1593 on 217 degrees of freedom
Multiple R-Squared: 0.02017, Adjusted R-squared: 0.006623
F-statistic: 1.489 on 3 and 217 DF, p-value: 0.2185
```

```
summary(VAR_SINALOA_FIS_FUTP1)#P=1
```

```
VAR Estimation Results:
```

```
=====
Endogenous variables: SINA, diff.PRECIOFISICOS.in., PFUT
```

```
Deterministic variables: none
```

```
Sample size: 220
```

```
Log Likelihood: 29.453
```

```
Roots of the characteristic polynomial:
```

```
0.9995 0.8842 0.03824
```

```

Call:
VAR(y = SINALOA_FIS_FUT, p = 1, type = "none")

Estimation results for equation SINA:
=====
SINA = SINA.l1 + diff.PRECIOSFISICOS.in..l1 + PFUT.l1

              Estimate Std. Error t value Pr(>|t|)
SINA.l1        0.94456    0.02556  36.950  <2e-16 ***
diff.PRECIOSFISICOS.in..l1 0.27295    0.12586   2.169   0.0312 *
PFUT.l1        0.09661    0.04639   2.083   0.0385 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4068 on 217 degrees of freedom
Multiple R-Squared: 0.9893,    Adjusted R-squared: 0.9891
F-statistic: 6659 on 3 and 217 DF,  p-value: < 2.2e-16

```

#### Apéndice 6. Prueba de raíz unitaria a modelos VAR

SERIE	MODELO*	PRUEBA DFA
$\Delta$ PMEXICO $\Delta$ PFISICO PFUTURO	VAR (1)	-4.486 p-value = 0.01
$\Delta$ PMEXICO $\Delta$ PFISICO PFUTURO	VAR (2)	-4.3807 (p-value =0.01)
$\Delta$ PMICHOACAN $\Delta$ PFISICO PFUTURO	VAR (1)	-5.9634 p-value = 0.01
$\Delta$ PMICHOACAN $\Delta$ PFISICO PFUTURO	VAR (3)	-5.6023, p-value = 0.01
$\Delta$ PJALISCO $\Delta$ PFISICO PFUTURO	VAR (1)	-5.367 p-value = 0.01
$\Delta$ PJALISCO $\Delta$ PFISICO PFUTURO	VAR (2)	-5.0459, p-value = 0.01
PSINALOA $\Delta$ PFISICO PFUTURO	VAR (1)	-5.993 p-value = 0.01
PSINALOA $\Delta$ PFISICO PFUTURO	VAR (2)	-5.4162 p-value = 0.01