

La producción de vainilla (*Vanilla planifolia*) en México: análisis y pronóstico

Samuel Luis-Rojas¹
Benito Ramírez-Valverde^{2§}
Maximino Díaz-Bautista³
José Pizano-Calderón¹
Carmen Rodríguez-López⁴

¹Posgrado en Economía, Posgrado en Recursos Genéticos y Productividad-*Campus* Montecillo-Colegio de Postgraduados. Carretera México-Texcoco km 36.5, Montecillo, Texcoco, Estado de México. CP. 56230. (luis.samuel@colpos.mx; pizano.jose@colpos.mx). ²Posgrado en Estrategias para el Desarrollo Agrícola Regional-*Campus* Puebla-Colegio de Postgraduados. Carretera Federal México-Puebla km 125.5, Santiago Momoxpan, Puebla. CP. 73475. (bramirez@colpos.mx). ³Universidad Intercultural del Estado de Puebla-Cuerpo Académico UIEP-CA-2 'Estudios del Patrimonio Biocultural'. Lipuntahuaca, Huehuetla, Puebla. CP. 56230. (maximino.diaz@uiep.edu.mx). ⁴Centro Gastronómico de Alto Rendimiento para la Nutrición Nacional (CGARNN). Calle Pinos núm. 1, Col. Guadalupe Victoria, Ciudad Serdán, Puebla. CP. 75520. (mc.colpos@hotmail.com).

§Autor para correspondencia: bramirez@colpos.mx.

Resumen

La vainilla (*Vanilla planifolia*) es uno de los aromáticos más demandados en la industria culinaria y refresquera, y entre los cultivos lícitos es de los más rentables, solo después del azafrán. En 2017 la producción en México fue 505.32 toneladas, ubicándolo como el cuarto productor mundial después de Madagascar, Indonesia y China. El objetivo de este estudio fue desarrollar un modelo de pronóstico para la producción anual de vainilla en México (PVAINI). Los datos fueron del periodo 2000 a 2016 y se utilizó la metodología Box-Jenkins de procesos autorregresivos integrados de medias móviles (ARIMA). Los parámetros de los modelos se estimaron con el método de máxima verosimilitud con el paquete de cómputo Statistical Analysis System. Un modelo basado en la serie PVAINI se adaptó para el periodo 2000 a 2016 y se validó con los datos de los años 2016 y 2017. El mejor modelo estimado fue ARIMA (1, 1, 1) e indicó que los PVAINI se explican con la producción ocurrida 4 años previos. Los valores pronosticados para 2017 fueron cercanos a los valores observados. El modelo ARIMA representó con cierta precisión el PVAINI en el próximo año y proporcionó información para planificar y tomar decisiones para los próximos seis años.

Palabras clave: metodología Box-Jenkins, modelos ARIMA, orquídea, planeación agrícola, predicción.

Recibido: noviembre de 2019

Aceptado: enero de 2020

Introducción

La vainilla (*Vanilla planifolia*) es una orquídea aromática originaria de México, cuya especie habita en regiones tropicales de América, aunque su principal zona de producción se ubica en la zona conocida como Totonacapan, que abarca 20 municipios del estado de Veracruz y 19 del estado de Puebla (Damirón, 2004; OMPI, 2009; Díaz *et al.*, 2017). Del fruto beneficiado de esta especie se obtiene la vainillina, el saborizante y aromatizante más popular y demandado en el mundo, ampliamente utilizado en la industria refresquera, vinícola y en el arte culinaria, ubicando a este cultivo como el más rentable, solo después del azafrán, con una demanda estimada por especialistas de la industria de saborizantes de alrededor de 14 000 t para el año 2012 (Lamas, 2012).

Sin embargo, a pesar de su importancia, a nivel mundial sólo se producen cerca de 100 toneladas de extractos de vainilla natural, el resto corresponde a sustitutos sintéticos con base en eugenol, cumarina y otros subproductos aromáticos de fermentaciones (Retes *et al.*, 2015). La producción nacional de vainilla pasó de 257 t en el año 2003 a 494.69 t en 2018, y registró una tasa de crecimiento medio anual (TCMA) de 4.78%, superior a la de la población (1.9%) (SIAP, 2017). Aun cuando México es el centro de origen y diversidad de la vainilla y cuenta con los agroecosistemas idóneos para su crecimiento y desarrollo, aunado a los amplios conocimientos técnicos y tradicionales, ampliamente desarrollados por la cultura Totonaca y Náhuatl (Díaz *et al.*, 2017).

La producción de frutos en verde obtenida en 2017 fue de 514.94 toneladas, ubicando al país como el cuarto productor mundial, después de Madagascar, Indonesia y China, que en el mismo año produjeron 3 227, 2 402 y 662 t respectivamente (FAO, 2017). Por consiguiente, Madagascar es el país que domina el mercado mundial en volumen de producción, haciendo que los demás países productores tengan que adaptarse a las circunstancias que marca dicho país. Las principales entidades productoras de esta orquídea en México son: Veracruz (76.08%), Puebla (11.55%), Oaxaca (10.5%), San Luis Potosí (1.87%) (SIAP, 2018).

En los tres principales estados se localizan 11 municipios con mayor número de productores de vainilla, entre los que destacan: Papantla, Cazonces de Herrera, Gutiérrez Zamora, Tecolutla, Tuxpan y Tihuatlán, todos ellos pertenecientes a la Zona del Totonacapan en el Estado de Veracruz, y por parte del Estado de Puebla, el municipio de Venustiano Carranza, Pantepec y Ayototxco de Guerrero, asimismo destaca la producción de los municipios de San Juan Bautista Valle Nacional y San Miguel Soyaltepec en Oaxaca (SIAP, 2018).

El destino de la vainilla beneficiada es el mercado internacional 90% aproximadamente 10% restante se distribuye en el territorio nacional en el mercado de artesanías y en las industrias de extractos (CESPV, 2007). Hay factores importantes que influyen en la toma de decisiones de los integrantes del sector agropecuario, como las predicciones y estimaciones de producción que establecen las agencias gubernamentales, por lo que deben ser eficientes y no sesgadas, Sanders y Manfredo (2002). Por esta razón es conveniente, tanto para los productores como para los diseñadores de política pública agropecuaria, disponer de medios que ayuden a conocer el comportamiento del mercado y que permitan el acceso a información actualizada que posibilite predecir el rumbo del mercado (Myers *et al.*, 2010).

Dada la variabilidad que presenta la oferta de la vainilla y su nivel de crecimiento en una producción se ve influida por factores biológicos y climáticos, Allen (1994), es relevante conocer con profundidad el comportamiento de esta oferta. Para este fin es fundamental la selección y estimación de un modelo estadístico con adecuada capacidad predictiva que permita representar el proceso generador de información (PGI). Los métodos de pronósticos se utilizan para realizar enunciados sobre el valor futuro de una variable bajo estudio ‘ya que tener indicios de lo que va a suceder, permite reducir la incertidumbre’ (Montemayor, 2013).

Según Evans (2003), desde los años sesenta se han realizado bastantes análisis económicos con la ayuda de modelos de regresión lineal multivariada, que buscan obtener pronósticos aunque en muchos casos con resultados particularmente decepcionantes, debido a que se parte de una relación causal establecida por la teoría o la lógica y se asumen que los valores futuros de la variable bajo estudio se pueden determinar proyectando los valores de variables de influencia que podrían ser o no controladas por el investigador o en algunos casos resulta imposible contar con la información de variables explicativas de importancia (Montemayor, 2013), por lo que para lograr mejores predicciones, la mayoría de economistas y pronosticadores prácticos, han recurrido al uso de métodos estadísticos que no involucran teoría económica como los modelos de series de tiempo univariados (Evans, 2003), pero con la desventaja de que la construcción de estos se basa en el uso del comportamiento de la misma variable para realizar la predicción (Asteriou y Hall, 2007).

Bajo este criterio Luis *et al.* (2019) utilizó una serie temporal univariada de precios mensuales de huevo blanco pagado al productor desde los años 2000 a 2016 y mediante la metodología Box Jenkins demostró que la serie se comporta como un modelo estacional Seasonal Autoregressive Integrate Movil Average (SARIMA) de coeficientes $(0, 1, 1)X(1, 0, 1)_s=12$ y con los resultados del modelo pudieron realizar estimaciones respecto al precio pagado al productor, hasta diciembre de 2019.

Sin embargo y a diferencia de las variables económicas necesarias para la construcción de modelos econométricos, en el estudio del comportamiento de las producciones agropecuarias en el tiempo la fundamentación teórica del modelo no es indispensable, lo cual incrementa la aplicabilidad de un modelo de series de tiempo (Barreras *et al.*, 2014). En la presente investigación se utilizó la metodología de procesos autorregresivos integrados de medias móviles (ARIMA). Estos modelos son más apropiados para predicciones a largo plazo que para corto plazo que presentan patrones estacionales; sin embargo, éstas se analizan sobre una serie estacionaria y se necesitan como mínimo 50 datos (Box y Jenkins, 1976), de ahí que su predicción sea más confiable.

El modelo ARIMA es útil en situaciones en las que los datos de series temporales muestran fluctuaciones periódicas de estacionalidad que se repiten con casi la misma intensidad cada año (Box y Jenkins, 1976). Esta característica hace que este modelo sea adecuado para estudios de productos agropecuarios y por lo tanto para la producción nacional de vainilla (García *et al.*, 2003) plantearon que el proceso biológico de la producción y los factores climatológicos, provocan que la mayor parte de la producción de los productos agrícolas salga en determinados meses del año, originando una marcada estacionalidad de la producción y, en consecuencia, de los precios.

Para el caso de la vainilla en México, la cosecha, está marcada por la Ley de Fomento y protección de la vainilla expedida el 13 de septiembre de 1941 y vigente hasta la fecha, pues establece que la cosecha se realiza a partir del 15 de noviembre cada año y ésta no dura más de un mes (Hernández,

2014). El objetivo de este estudio fue desarrollar un modelo de series de tiempo para pronosticar la producción de vainilla verde en México (PVAINI) con base en la producción de vainilla verde disponible entre 1961 y 2016. Con los datos recopilados de los años 2015 y 2016 se validó el modelo. Dada la incertidumbre del mercado, el pronóstico de la PVAINI proporciona información útil para facilitar la planificación en la toma de decisiones de los diferentes actores que forman parte de este sector. La hipótesis planteada fue que la producción de vainilla en verde en México en el futuro es explicada por la producción ocurrida 4 años antes.

Materiales y métodos

Para conocer el comportamiento de la PVAINI y realizar pronósticos se utilizó una serie histórica de producción de vainilla verde, proporcionados por el Sistema de Información Agroalimentaria y Pesquera (SIAP, 2019) y por la base de datos de la Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO, 2019). La serie de tiempo de la PVAINI se dividió en dos partes: los datos de 1961 a 2014 expresados en toneladas métricas (t), que se utilizaron para desarrollar el modelo de series temporales; asimismo, para validar el modelo se utilizaron los datos de la producción de 2015 y 2016.

Si se asume que $Y' = (Y_1, Y_2, \dots, Y_n)$ es una serie de tiempo, entonces un modelo ARIMA es denotado como (p, d, q) (Guerrero, 1991) y está dado por:

$$\phi_p(B) \nabla^d Y_t = \theta_q(B) a_t$$

Donde: $B =$ es el operador de retraso ($B^r Y_t = Y_{t-r}$); $Y_t =$ es la serie de tiempo de la variable Y en el tiempo t ; $\phi_p(B) =$ es el polinomio autorregresivo (AR), de orden 'p', se desarrolla así, $\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p$; $\theta_q(B) =$ es el polinomio de promedios móviles (MA), de orden 'q', donde: $\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q$; $\nabla^d = (Y_t - Y_{t-r})^{D=1}$ es el número de diferencias que se necesitan para estacionar la serie se y se simboliza 'd'; $a_t =$ se refiere a los términos de error aleatorio o también llamado ruido blanco que se supone que son variables aleatorias independientemente distribuidas de forma idéntica muestradas de una distribución preferentemente con media igual a cero y varianza $a_t \sim N(0, \delta^2)$ (Box y Jenkins, 1976).

En el análisis de series de tiempo, según Box y Jenkins (1976), la variable a_t se conoce comúnmente como ruido blanco, y se interpreta como un efecto exógeno que el modelo no explica. Teniendo en cuenta la serie temporal PVAINI, este ruido blanco puede ser un efecto de variables climáticas, enfermedades, el tipo de cambio peso-dólar, la situación de riesgo país, catástrofes naturales, etc. Se utilizó el procedimiento PROC ARIMA del software Statistical Analysis System (SAS) versión 9.4. SAS (2014), para estimar el modelo ARIMA a la serie PVAINI del periodo 1961-2014 utilizando la metodología Box y Jenkins para la construcción y ajuste del modelo de pronóstico (identificación, estimación de parámetros, verificación de supuestos y predicción) (Box y Jenkins, 1976).

La idoneidad de cada modelo se verificó mediante gráficas de los residuales contra el tiempo. Según Ljung y Box (1978), si el modelo es adecuado se esperaría que los residuales estén distribuidos de forma aleatoria alrededor de cero sin mostrar ningún patrón o tendencia; en caso de

que existiera algún tipo de comportamiento irregular, esto sería un motivo para suponer que los residuales no tienen una distribución normal y no son puramente aleatorios. En este trabajo, se utilizó la elección del modelo de forma parsimoniosa sugerida por Burnham y Anderson (2002) mediante el criterio de información akaike (AIC) y el criterio bayesiano de shwartz (SBC) los cuales comparan la bondad de ajuste de los diferentes modelos.

Ambos criterios se basan en la utilización de la suma de los cuadrados de los errores, buscando minimizarla a partir de diversas combinaciones de p y q . Los valores de AIC y el SBC más bajos indican un mejor ajuste en el modelo.

Resultados y discusión

La tasa de crecimiento media anual (TCMA) de la serie PVAINI para el periodo 1961-2017, fue de 1.43% lo cual indica una ligera tendencia al alza, sin embargo; también se observa que para el periodo 1961-1979, la producción de vainilla decayó a una TCMA de -16.9%, dicho comportamiento de la producción se debió que para esas épocas se destinaron menos hectáreas de tierra a este cultivo, al pasar de terrenos agrícolas productores de esta orquídea a terrenos para la producción de ganado, para la producción petrolera, casos climatológicos extremos y al desánimo en el precio internacional de la vainilla, ocasionaron pérdidas en la producción de vainilla.

Según Damirón (2004), en 1980 se estimó que existían al menos 800 productores organizados y 700 sin registro que produjeron 30 toneladas de vainilla, en este contexto, se esperaba que la producción de esta orquídea se estancara; sin embargo, para el periodo 2012-2017 se presenta una tasa de crecimiento (5.72%) y se esperaba que la producción se incremente en el corto plazo.

Los picos más altos de la serie temporal corresponden a los años de 1993, 1998 y 2007 y se explican por buenas condiciones de temporal y a la especulación del precio ocurrido tres años antes, mientras que la producción más baja de vainilla en México, corresponde al periodo 1978-1980, explicados en el párrafo anterior, respecto a la caída en la producción de 2002 se explican por sequías extremas ocurridas en 1996 y 1997 y la depresión tropical número 11 ocurrida en octubre de 1999, mientras que la caída en la producción en 2011, estuvo influenciada en gran medida por exceso de lluvias (los huracanes Stan en 2005 y Dean en 2007) que afectaron la zona del Totonacapan y siniestraron gran parte de los cultivos (Figura 2a).

En la metodología Box-Jenkins, es necesario que las series estudiadas sean estacionarias en media y varianza, la primera se obtiene con al menos una diferencia, mientras que para estabilizar la varianza se aplica una transformación de la familia Cox-Box, éstas además de estabilizar la varianza, mejoran la aproximación a la distribución normal del proceso $\{Z_t\}$, donde λ es el parámetro de transformación. La prueba de Cox-Box arrojó un $\lambda=0.5$, por lo que la serie PVAINI se debió transformar en logaritmos naturales para tener varianza constante.

La magnitud de este coeficiente concuerda con la teoría de las series temporales, planteada por (Cox y Box, 1964) para la aplicación de cualquiera de los modelos de la familia de los ARMA. Entonces $Y' = (Y_1, Y_2, \dots, Y_n)$ es el vector de los logaritmos naturales, renombrando ahora a la serie como LPVAINI. Se observa que la serie Y_1, Y_2, \dots, Y_n frente al tiempo todavía hay cierta tendencia, pero a través de la primera diferencia (∇) se obtiene una serie estacionaria. Por lo tanto, se considera $d=1$ (Figura 2b).

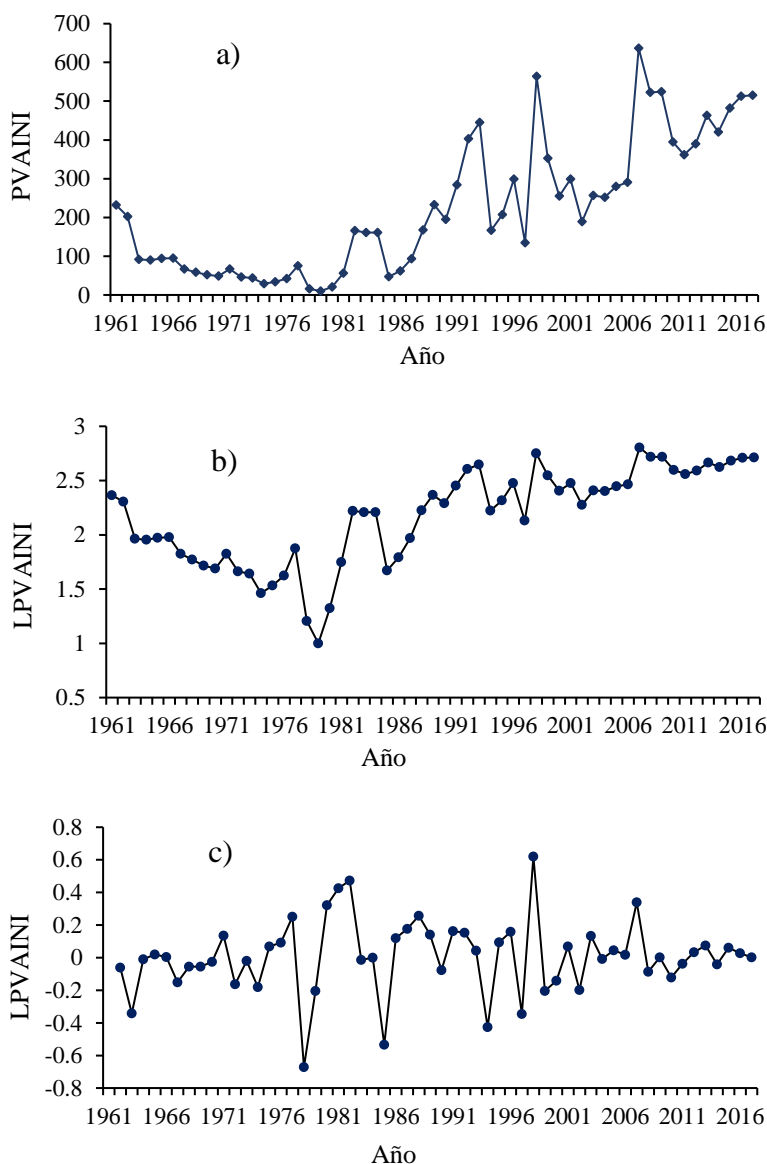


Figura 2. Diagnósticos gráficos de la serie original (PVAINI): a) comportamiento original de la serie PVAINI en (t), 1961-2017; b) comportamiento de la serie PVAINI transformada a logaritmos naturales (LPVAINI); y c) serie LPVAINI diferenciada, transformada con logaritmos naturales y sin aparente tendencia. Elaborado con datos de SIAP (2017); FAO (2017).

La serie PVAINI diferenciada y transformada a logaritmos naturales se ha renombrado LPVAINI (1), y se intuye a priori que ya es estacionaria en media (no existe tendencia) y en varianza, situación que concuerda con lo establecido con Box y Jenkins (1976) puesto que la serie a estudiar debe tener una variabilidad constante a través del tiempo y sin tendencia porque en la metodología trabajada es condición necesaria para obtener mejores pronósticos (Figura 2c). Se procedió a verificar la estacionariedad o no estacionariedad de la serie de tiempo de manera estadística, planteada por Dickey y Fuller (1981) mediante la prueba de raíz unitaria de Dickey Fuller Aumentada (ADF) la cual consiste en incluir rezagos de la primera diferencia de Y_t en la regresión de prueba para admitir la posible existencia de autocorrelación serial.

Para ello se plantea la siguiente hipótesis, H_0 : La serie no es estacionaria ($\rho = 1$) y tiene raíz unitaria vs H_a : la serie es estacionaria ($\rho \neq 1$) y no tiene raíz unitaria. Regla de decisión: se rechaza H_0 si $|\tau \text{ calculado}| \geq |\tau \text{ de tablas}|$ (Pankratz 1983). Dado que el valor absoluto de la tau calculada (7.81) es mayor que la tau de tablas, se rechaza la hipótesis nula y se concluye que la serie PVAINI no tiene raíz unitaria, por lo tanto, es estacionaria; es decir, tiene media y varianza constante en el tiempo (Cuadro 1).

Cuadro 1. Prueba de Dickey Fuller aumentada (ADF) para la serie de precios promedios mensuales de Vainilla en México (PVAINI).

Tipo	Retardos	Rho	Pr < Rho	Tau	Pr < Tau	F	Pr > F
Media cero	0	-58.3396	<0.0001	-7.81	<0.0001		
	1	-72.5996	<0.0001	-6.07	<0.0001		
	2	-197.244	0.0001	-5.82	<0.0001		
Media simple	0	-58.3927	0.0005	-7.75	0.0001	30.03	0.001
	1	-73.0092	0.0005	-6.04	0.0001	18.25	0.001
	2	-202.766	0.0001	-5.81	0.0001	16.88	0.001

Elaborado con los resultados de salida de las transformaciones de PVAINI, PROC ARIMA SAS (2014).

La función de autocorrelación estimada (ACF) y la función de autocorrelación parcial (PACF) de la serie $\nabla LPVAINI$, muestra una ACF alta en el rezago = 13. Con respecto a la FACP se presentan picos altos en las correlaciones para los rezagos = 3 y 4, por lo que se intuye que un proceso ARIMA describe y explica el comportamiento de los PVAINI (Figura 3).

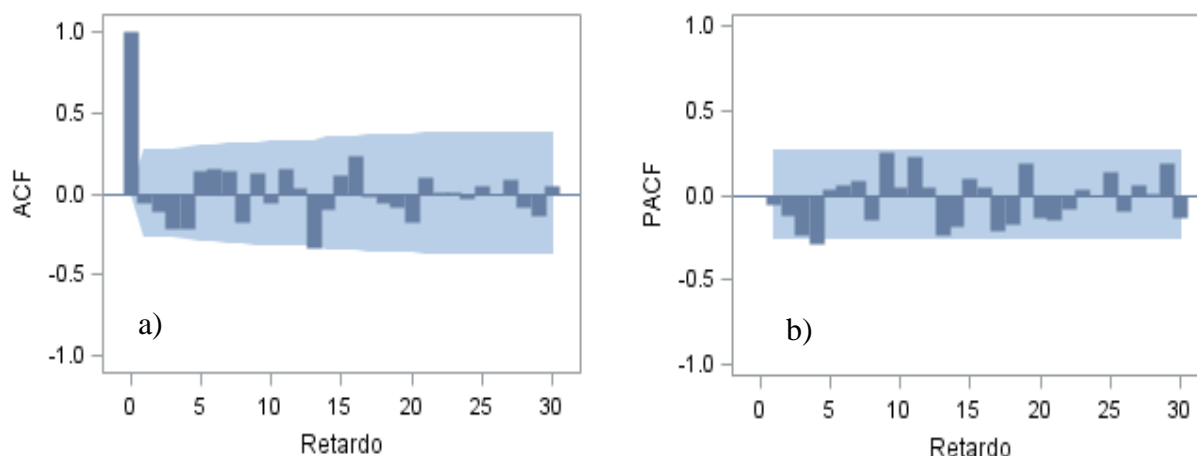


Figura 3. Función de autocorrelación ACF (a) y función de autocorrelación parcial PACF (b), estimadas para la serie $Y_t = (1-B)LPVAINI$. Elaborado con los resultados de salida de PROC ARIMA (SAS, 2014).

La etapa de identificación del modelo es una inspección visual de la ACF y la PACF graficadas, aunque pueden ocultar información de interés para la elección del modelo adecuado, de acuerdo a lo planteado por Gujarati y Porter (2009) por lo que se propusieron 10 modelos. De estos se identificaron los cuatro mejores que cumplen con la significancia de parámetros y ruido blanco. Se eligieron coeficientes AR, MA toda vez que la ACF y PACF presentan simultáneamente

coeficientes distintos de cero, asimismo, se estimaron considerando diferentes elecciones de p y q , así como los valores de SBC, AIC y la varianza δ_ε^2 para los cuatro mejores modelos ARIMA ajustados a la serie LPVAINI.

Se considera que la primera diferencia en la parte no estacional suele ser suficiente ($d=1$) se estableció d en 1 en todos los modelos. El modelo con el valor SBC y AIC más bajo para este conjunto de datos, fue el ARIMA (1, 1, 1) (Cuadro 2).

Cuadro 2. Valores de AR, MA, SBC y AIC de los modelos ARIMA identificados (p, 1, q) y estimadores para δ_ε^2 .

Mejores modelos ARIMA [†]	Rezagos		Coeficientes				SBC ^p	AIC ^α	δ_ε^2
	AR	MA	AR1, 1 [¶] (p)	AR1, 2 ^{¶¶} (p1)	MA1, 1 [§] (q)	MA1, 2 ^{§§} (q1)			
1.- (0, 1, 1)		13	-	-	0.37968	-	81.26018	79.23483	0.22836
2.- (1, 1, 1)	4	13	-0.27144	-	0.44899	-	81.21013	77.15943	0.21166
3.- (2, 1, 0)	3 y 4	13	-0.23933	-0.28862	0.46485	-	81.67734	75.60129	0.20075
4.- (1, 1, 2)	4	4 y 13	-0.07547		0.23633	0.43376	84.01513	77.93908	0.21044

[†]ARIMA= proceso autorregresivo integrado de media móvil; [¶]AR= coeficiente autorregresivo de orden (p), ^{¶¶}AR= coeficiente autorregresivo de orden (P); [§]MA= coeficiente de media móvil de orden (q); ^{§§}MA= coeficiente de media móvil de orden (Q); ^pSBC= criterio bayesiano de shwartz; ^αAIC= criterio de información akaike. Elaborado con resultados de las estimaciones de los modelos ARIMA, SAS (2014).

Utilizando el proceso PROC ARIMA (SAS 2014), se calcularon los parámetros AR1,1 (ϕ_4) y el componente de promedio móvil MA1,1 (θ_{13}), mediante máxima verosimilitud, debido a que este método parte que sus estimadores son asintóticamente óptimos y si el tamaño de la serie es grande, se puede considerar que son centrados o insesgados, eficientes y su distribución es normal (Montemayor, 2013). Para la interpretación y significancia de los estimadores (Box *et al.*, 1994), plantean que para que se considere un modelo moderado, se deben incluir aquellos parámetros cuya estadística t absoluta sea mayor que 2 o los p -values sean menores a 0.05. Por lo tanto, se considera al modelo ARIMA (1, 1, 1) como moderado, ya que se ajusta suficientemente a los datos antiguos sin usar ningún parámetro innecesario (Cuadro 3).

Cuadro 3. Estimación del modelo para la serie de tiempo LPVAINI(1) por máxima verosimilitud.

Parámetro	Estimación	Error estándar	Valor t	Aprox Pr > t	Retardo
MA1, 1	0.44899	0.15276	2.94	0.0033	13
AR1, 1	-0.27144	0.12845	-2.11	0.0346	4

Elaborado con los resultados de salida de PROC ARIMA, SAS (2014).

Después de estimar los parámetros del modelo se validó mediante el análisis de residuos (Yafee y McGee, 1999). Los residuos estandarizados, se analizaron a través de un histograma, el gráfico de ACF respectivo y los valores de p para las pruebas de ruido blanco. Se aclara de que los residuos estandarizados estimados, a partir de este modelo deberían comportarse como una secuencia independiente e idénticamente distribuida, con una media de cero y una varianza constante (Figura 4a). La distribución de los residuos se aproximó a una distribución normal, lo que indica una gran afinidad de los datos, sin importar la magnitud de éstos (Figura 4b).

La ACF de los residuos sugiere que las autocorrelaciones están dentro de la banda de confianza; es decir, son cercanas a cero. Este resultado indica que los residuos no se desviaron significativamente de un proceso de ruido blanco cero, y son puramente aleatorios, por lo que ya no hay información de dependencia de unos datos con otros a través del tiempo (Figura 4c). Dado el alto p-value, asociado con las estadísticas de Chi cuadrada, no existe razón para rechazar la hipótesis nula de que los residuos son ruido blanco (p value mayor o igual a 0.05). Por lo tanto, el modelo ARIMA (1, 1, 1) se ajusta al comportamiento de los datos de LPVAINI (Figura 4d).

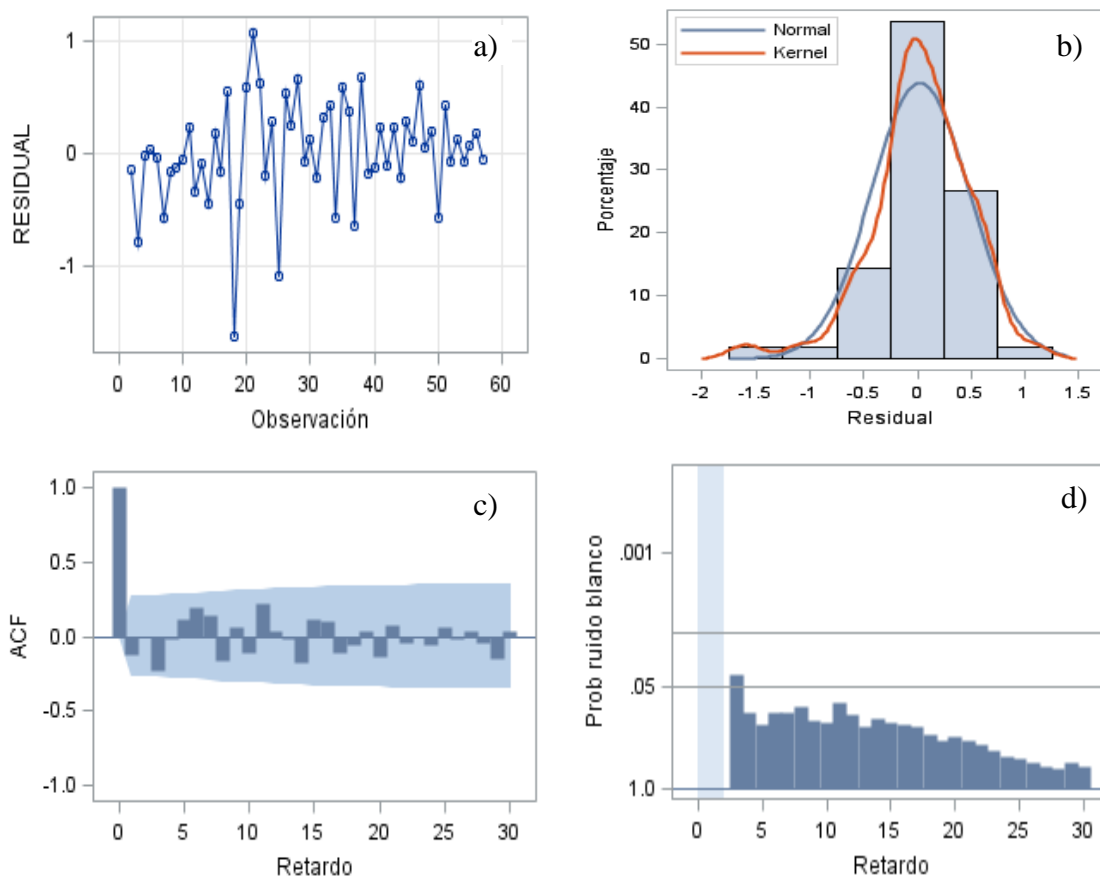


Figura 4. Diagnósticos gráficos para evaluar el ajuste del modelo ARIMA (1, 1, 1): (a) Residuos estandarizados; (b) Histograma de los residuos estandarizados; (c) ACF de los residuales; y (d) Valores p para las pruebas de ruido blanco. Elaborado con los resultados de la mejor estimación del modelo ARIMA (1, 1, 1).

Los valores pronosticados fuera de muestra para los años 2017 a 2024 teniendo en cuenta el modelo ARIMA (1, 1, 1) se compararon con los de la serie PVAINI, las predicciones son buenas, habiéndose obtenido un error porcentual absoluto medio (MAPE) de 8.32 que indica que nuestro pronóstico está errado en un $\pm 8.32\%$. Sin embargo, no se debe basar el pronóstico únicamente en este indicador, también se debe examinar la prueba de Chi cuadrada para conocer si la serie presenta residuales que se comportan como ruido blanco.

Este modelo predice valores de PVAINI muy próximos a los observados, porque todas se ubican dentro de la banda de confianza estimada al $\pm 95\%$, bajo este criterio de destaca que, si no se toman medidas urgentes por parte de los actores que influyen en las políticas públicas, la producción de

vainilla en México decaerá de manera acelerada en los próximos 6 años a una TCMA de -5.05%; es decir, la producción estimada para 2019 será de 464 a 340 t en 2024. Asimismo, se estimó el intervalo de confianza al 95% de los valores fuera de la serie y se encontró que para el año 2019, el rango de ocurrencia será de 188.3 t como mínimo y 1 143 t como máximo, mientras que para el año 2024, con el mismo intervalo de confianza, el rango se ubica en 40.31 t como mínimo y 2 868.2 t como máximo (Figura 5).

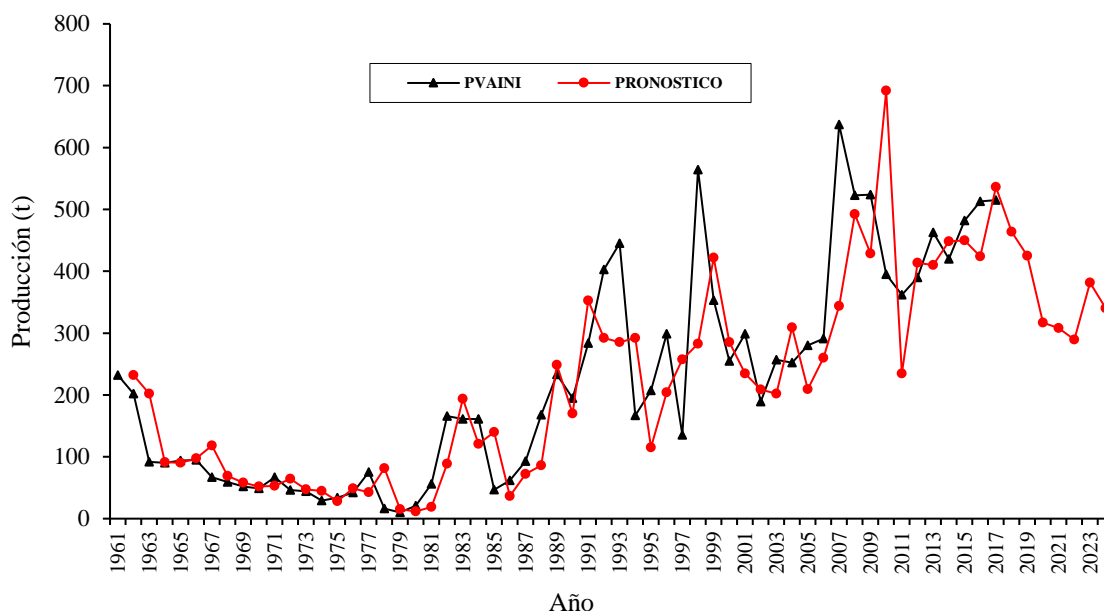


Figura 5. Producción observada y pronosticada de vainilla en México, 2017 (toneladas métricas), obtenida a través del modelo ARIMA (1, 1, 1). Elaborado con los resultados del modelo ARIMA (1, 1, 1) del SIAP (2017) y FAO (2017).

Este modelo se usa para predecir fuera de la muestra la PVAINI y fue capaz de hacerlo con relativa precisión para un año posterior. Sin embargo, Chu (1978) plantea que estas estimaciones pueden no ser creíbles para predecir los precios en el mediano y largo plazo, por lo que es preciso señalar, que no obstante, el principal inconveniente que presenta este tipo de modelos, radica en su propio fundamento, ya que al estar basados en hechos pasados y aprender de la propia historia para realizar las predicciones, éstas serán acertadas en la medida en que los factores que determinan la evolución sigan actuando del mismo modo y no sean alterados de forma notoria.

Dado que este producto es importante por su alto valor comercial y aprecio en la alta cocina gourmet, además de que su cultivo requiere de condiciones edafoclimáticas propias de ciertas regiones tropicales de México, es importante que los actores que influyen en el diseño de políticas agroalimentarias, planteen de manera urgente el rescate, promoción e incremento de la superficie sembrada de vainilla, así como el impulso y ejecución de leyes que sancionen el robo del fruto en plantación, intermediarismo y especulación puesto que de no ser así, no solo decaerá la producción sino también el ánimo por seguir sembrando la vainilla.

Conclusiones

Es posible modelar los PVAINI en México utilizando la metodología propuesta por Box y Jenkins, dado que el comportamiento de estas series no es completamente aleatorio y se describen como series de tiempo con una alta probabilidad de éxito en la modelación a través de dicha metodología. El mejor modelo que representó el comportamiento de la PVAINI fue el ARIMA (1, 1, 1) al demostrar que la producción de vainilla en verde en un año dado es estimada por la producción de vainilla que ocurrió 4 años antes. Las estimaciones de los parámetros y los pronósticos fueron apropiados estadísticamente y significativos, aun cuando la serie estudiada presenta variabilidad alta.

Las predicciones del modelo en el corto plazo difieren en $\pm 8.32\%$ de los datos observados, minimizándose el error aleatorio, encontrándose que en el corto plazo la producción de vainilla en México decaerá a una tasa de 5.05% anual. Los PVAINI estimados proporcionan información útil para planificar, incentivar y rescatar de manera urgente a la vainilla en México, a través del diseño y aplicación de políticas públicas de índole agroalimentario, toda vez que es un producto originario de nuestro país y rentable para mejorar el ingreso de los productores, porque de no ser así, la producción de vainilla en México en el mediano plazo tiende a disminuir de manera acelerada.

Recomendaciones

Se sugiere que para estudios futuros donde se utilice la metodología Box-Jenkins, se introduzca al modelo nuevas variables que expliquen la perturbación aleatoria y que estén asociadas de manera directa con el incremento de la PVAINI, aunque su relación no es instantánea ni estática, se considera la dimensión temporal de las observaciones, permitiendo medir y evaluar las respuestas dinámicas y la forma en que se transmiten los efectos entre las variables (Box y Jenkins, 1976). Las variables para considerar serían los precios promedios anuales de saborizantes artificiales de vainilla y la producción anual de los países de Madagascar e Indonesia, con lo que se esperarían estimaciones más precisas, puesto que al incluir variables de causalidad unidireccional (variables independientes) desde la entrada hacia la variable respuesta (dependiente) el MAPE del pronóstico debiera disminuir (Chigona *et al.*, 2007). Para ello se plantea el uso futuro de los modelos multivariantes de series de tiempo también llamados modelos de función de transferencia (TFM) a diferencia del modelo univariante ARIMA (1, 1, 1) estimado en este trabajo para la serie PVAINI. Según Keller (1987) estos modelos permiten simular y estudiar cómo ciertos 'escenarios' definidos por posibles evoluciones de la variable explicativa, afectan a la variable respuesta, de ahí que la predicción sea más confiable que los métodos que utilizan el modelo de regresión clásica.

Literatura citada

- Allen, P. G. 1994. Economic forecasting in agriculture. *Inter. J. Forecasting*. 10(1):81-135.
- Asteriou, D. and Hall, S. G. 2007. *Applied econometrics: A modern approach*. Palgrave MacMillan. New York, USA. 397 p.
- Barreras, S. A.; Sánchez, L. E.; Figueroa, S. F.; Olivas, V. J. A. y Pérez, L. C. 2014. Uso de un modelo univariado de series de tiempo, en el corto plazo del comportamiento de la producción de carne de bovino en Baja California, México. *Rev. Veterinaria México*. 1(1):1-9.

- Box, G. E. P. and Jenkins, G. 1976. Time series analysis: forecasting and control. Ed. Holden-day. San Francisco, CA. 575 p.
- Box, G. E. P.; Jenkins, G. M. and Reinsel, G. C. 1994. Time series analysis: forecasting and control. 3rd (Ed.). Ed. Prentice Hall. Englewood Cliffs, NJ. 712 p.
- Burnham, K. P. and Anderson, D. R. 2002. Model selection and multimodel inference: a practical information-theoretic approach. 2th (Ed.). Ed. Springer. New York. 515 p.
- CESPV. 2007. Comité Estatal del Sistema Producto Vainilla del Estado de Puebla. Estudio de oportunidades de mercado internacional para la vainilla. (Ed.). SAGARPA-FIRCO. 238 p.
- Chigona, M.; Gaetan, C. and Masarotto, G. 2007. Automatic identification of seasonal transfer function models by means of iterative stepwise and genetic algorithms. *In: J Time Ser Anal*, (29)1:37-50.
- Chu, K. Y. 1978. Short-run forecasting of commodity prices: an application of autoregressive moving average models. *IMF Staff Papers*. 25(1):90-111.
- Cox, D. R. and Box, G. E. P. 1964. An analysis of transformations. *J. Royal Statistical Society*. 26(2):211-252.
- Damirón, V. R. 2004. La vainilla y su cultivo. Dirección General de Agricultura y Fitosanitaria. Gobierno del Estado, Xalapa, Veracruz. México. 50 p.
- Dickey, D. A. and Fuller, W. A. 1981. Likelihood ratio statistics for autoregressive time series with a unit root. *Econ. J. Econ. Soc.* 49(4):1057-1072.
- Evans, M. 2003. Practical business forecasting. Malden Massachusetts: Blackwell Publishers. Oxford. 483 p.
- FAO. 2017. Organización de las Naciones Unidas para la Agricultura y la Alimentación. Descargar datos. Roma Italia. <http://faostat.fao.org> Información en línea.
- García, M. R.; García, S. J. A. y García, S. R. C. 2003. Teoría del mercado de productos agrícolas. Instituto de Socioeconomía, Estadística e Informática-Programa de Posgrado en Economía. Colegio de Postgraduados. Montecillo, Texcoco, Estado de México. 382 p.
- Guerrero, G. V. M. 1991. Análisis estadístico de series de tiempo económicas. Ed. UAM-Iztapalapa. México. 309 p.
- Gujarati, D. N. and Porter, C. D. 2009. Econometría. 5^a (Ed.), Ed. Mc Graw Hill, México. 921 p.
- Hernández, H. J. 2014. Técnicas implementadas para el cultivo de la vainilla en México. *In: 1^{er} Seminario internacional de la vainilla, promoviendo la investigación, la extensión y la producción de vainilla en Mesoamérica*. Araya, F. C.; Cordero, S. R.; Paniagua, V. A y Azofeifa, B. J. A. (Eds.). 1^{er}. Ed. Editorial Universidad Nacional. Heredia. Costa Rica. 81-92. pp.
- Keller, A. 1987. Modeling and forecasting primary production rates using Box-Jenkins transfer function Models. *Canadian J. Fisheries Aquatic Sci.* 44(5):1045-1052.
- Lamas, N. M. A. 2012. Producción de vainilla en invernaderos. *El economista*. <https://www.economista.com.mx/opinion/Produccion-de-vainilla-en-invernaderos-20120117-0010.html>.
- Ljung, G. M. and Box, G. E. P. 1978. On a measure of lack of fit in time series models. *Econométrica*. 65(2):297-303.
- Luis, R. S.; García, S. R. C.; García, M. R.; Arana, C. O. A. y González, E. A. 2019. Metodología Box - Jenkins para pronosticar los precios de huevo blanco pagados al productor en México. *Agrociencia*. 53(6):911-925.
- Montemayor, G. J. E. 2013. Métodos de pronóstico para los negocios. Ed. Digital. Tecnológico de Monterrey. Monterrey, NL. México. 27 p.

- Myers, R. J.; Sexton, R. J. and Tomek, W.G. 2010. A century of research on agricultural markets. *Amer. J. Agric Econ.* 92(2):376-402.
- OMPI. 2009. Organización Mundial de la Propiedad Intelectual. Registro Internacional de la denominación de origen: Vainilla de Papantla. Ginebra, Suiza, 3 septiembre-2009. 6 p.
- Pankratz, A. 1983. *Forecasting with univariate Box-Jenkins models concepts and cases*. Ed., John Wiley & Sons. United States of America. 576 p.
- Retes, M. R. F.; Torres, M. M. T. y Lugardo, B. M. T. 2015. Ventajas económicas para la industria de alimentos y bebidas en México con el uso de la vainillina obtenida del nejayote. *Rev. Custos e agronegocio on line.* 11(3):86-105.
- Sanders, D. R. and Manfredo, M. R. 2002. USDA production forecasts for pork, beef and broilers: An evaluation. *J. Agric. Res. Econ.* 27(1):114-127.
- SAS. 2014. *Statistical Analysis System. System for Windows V. 9.4*. Sas Institute Inc. Cary, NC. USA.
- SIAP. 2017. *Sistema de Información Agroalimentaria y Consulta. Servicio de Información Agroalimentaria y Pesquera. SADER. México.* <https://www.gob.mx/siap/prensa/sistema-de-informacion-agroalimentaria-de-consulta-siacon>.
- Yafee, R. A. and McGee, M. 1999. *Introduction to time series analysis and forecasting*. Ed., Academic Press, Inc. Brooklyn, New York. 83-89 pp.