



Análisis de la metodología ARIMA en el pronóstico de datos a un caso de estudio de la industria de la construcción

Julia Patricia Melo Morín^{1*}, María de los Ángeles Ahumada Cervantes¹

¹Instituto Tecnológico Superior de Pánuco

*patricia.melo@itspanuco.edu.mx

RESUMEN

La predicción se utiliza para anunciar algo que sucederá en un futuro a partir de hechos científicos, hipótesis o indicios; anticipando como las variables económicas cuantitativas o cualitativas, se pueden comportar a partir de sus valores pasados y presentes. La predicción se utiliza para anunciar algo que sucederá en un futuro a partir de hechos científicos, hipótesis o indicios; anticipando como las variables económicas cuantitativas o cualitativas, se pueden comportar a partir de sus valores pasados y presentes.

Una metodología utilizada en la predicción de datos es la metodología ARIMA, que contiene una serie etapas para determinar los posibles valores a ocurrir a partir de datos históricos conocidos. El presente artículo tiene como objetivo describir paso a paso la metodología ARIMA a un caso de estudio de una serie temporal de 110 valores para el pronóstico de los datos de dos meses de los posibles valores del índice de precio y cotizaciones en la bolsa de valores de la industria CEMEX. Se demuestra paso a paso cada una de las etapas de la metodología con el software Minitab, describiendo el proceso que se tiene que verificar y validar para continuar con el pronóstico de los datos y las modificaciones que se tienen que realizar.

Palabras claves: Metodología ARIMA, series temporales, Minitab.

ABSTRACT

The prediction is used to announce something that will happen in the future based on scientific facts, hypotheses or clues; anticipating how quantitative or qualitative economic variables can behave based on their past and present values. The prediction is used to announce something that will happen in the future based on scientific facts, hypotheses or clues; anticipating how quantitative or qualitative economic variables can behave based on their past and present values.

One methodology used in data prediction is the ARIMA methodology, which contains a series of steps to determine the possible values to occur from known historical data. The objective of this article is to describe step by step the ARIMA methodology to a case study of a time series of 110 values for the forecast of two-month data of the possible values of the price index and quotes on the stock market of the CEMEX industry. Each of the stages of the methodology is demonstrated step by step with the Minitab software, describing the process that must be verified and validated to continue forecasting the data and the modifications that must be made.

Key words: ARIMA methodology, time series, Minitab.

INTRODUCCIÓN

Para las instituciones financieras, educativas y empresariales, es de gran interés conocer lo que pueda suceder en el futuro, para apoyar a la toma de decisiones, considerándose en el algo importante y constante. De acuerdo con Cortez (1992), "el futuro solo puede ser inferido a partir de un profundo estudio del pasado", indicando que las predicciones son de gran utilidad en la toma de decisiones.

El pronóstico es la predicción de algo futuro a partir de indicios pasados, que permite disminuir la incertidumbre a una decisión que las empresas pueden tomar y que impacta más adelante; son utilizados en temas científicos, tecnológicos o empresariales; el análisis de las series de tiempo es una de las formas tradicionales para llevar a cabo el pronóstico, considerándose una secuencia de datos u observaciones sobre una variable cuantitativa, en diferentes medidas de tiempo, ordenados cronológicamente.

Box y Jenkins, propusieron (1970), una metodología para identificar, estimar y diagnosticar modelos en series de tiempo; denominada ARIMA (Autoregressive Integrated Moving

Average), que se compone de los modelos autorregresivos (AR) y los de medias móviles (MA) una particularidad de los modelos ARMA. Para llevar a cabo la aplicación de la metodología ARIMA, se requiere de un conocimiento teórico de cada una de las fases de la misma y de su interpretación, así como de un software que permita realizar los cálculos requeridos, de forma rápida (Box, 1984).

Este artículo describe la aplicación de los modelos ARIMA como una herramienta para el pronóstico de datos para un tiempo dado a un caso de estudio de la industria de la construcción, utilizando la herramienta de software Minitab.

METODOLOGÍA

La metodología aplicada para llevar a cabo el pronóstico de los datos a partir de datos históricos utilizando series temporales es la metodología ARIMA.

La metodología de los modelos ARIMA de acuerdo con Box, G. y Jenkins (1990) citado por Cortez (2011), se presenta en la Figura 1.

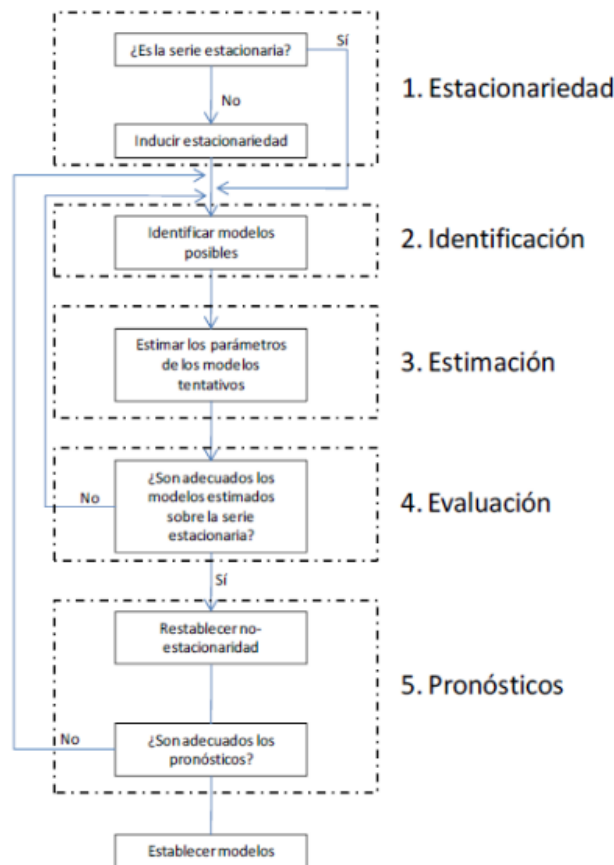


Figura 1. Metodología Box-Jenkins. Fuente Cortez (2011).

En series de datos con patrones estacionales, son útiles la metodología ARIMA, aunque es factible aplicarse a series no estacionarias. Las etapas o fases de la metodología son: Estacionariedad, Identificación, Estimación, Evaluación y Pronósticos, cada una de ellas, se describirán y ejemplificarán en la sección de resultados a un caso de estudio para realizar el pronóstico de los posibles valores del índice de precio y cotizaciones en la bolsa de valores de la industria CEMEX.

RESULTADOS

Se aplicará los pasos de la Metodología ARIMA para llevar a cabo la predicción a corto plazo, considerando el pronóstico de 2 meses siguientes (60 días) a la serie de datos de 110 valores previos a la predicción.

1) Fase de Estacionariedad

La metodología ARIMA, indica como primera fase la verificación de si la serie es o no estacionaria, considerándose como un modelo en una serie de tiempo, indica incrementos y disminuciones de los daos en forma regular. La estacionalidad es una forma de validar la precisión de los pronósticos.

Para llevar a cabo ese elemento se obtiene la auto correlación de los datos de la serie, previamente capturados en Minitab. El procedimiento es eligiendo el Menú Estadísticas, la opción Series de Tiempo y se selecciona la función de auto correlación, el resultado es mostrado en la Figura 2.

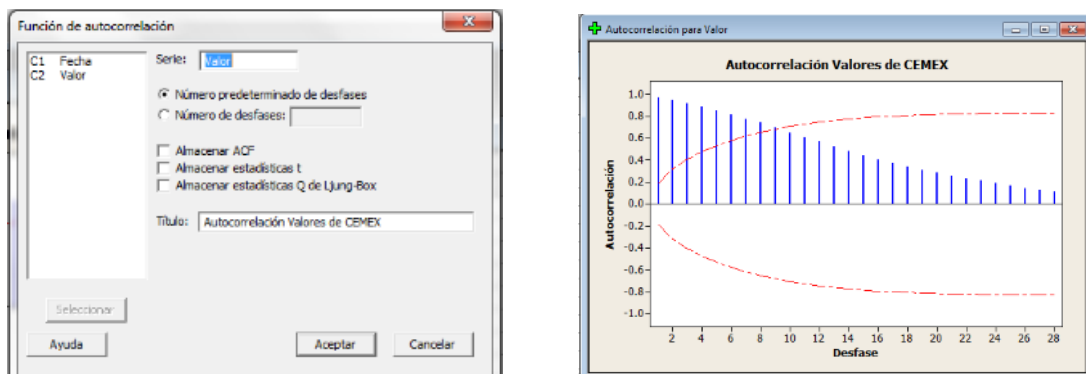


Figura 2. Auto correlación de los valores iniciales de la serie de CEMEX. Fuente: Ejecución en programa Minitab

Al observar la gráfica se verifica que la serie tiene tendencia, ya que los valores descienden con rapidez hasta llegar a 0, por lo que no es estacionaria.

Para convertirla en una serie estacionarias, se debe aplicar la diferenciación a los datos. El procedimiento en Minitab es seleccionando el *Menú Estadísticas, la opción Series de Tiempo y se selecciona la función Diferencias*. La *Figura 3* muestra el resultado de la diferenciación de la serie original.

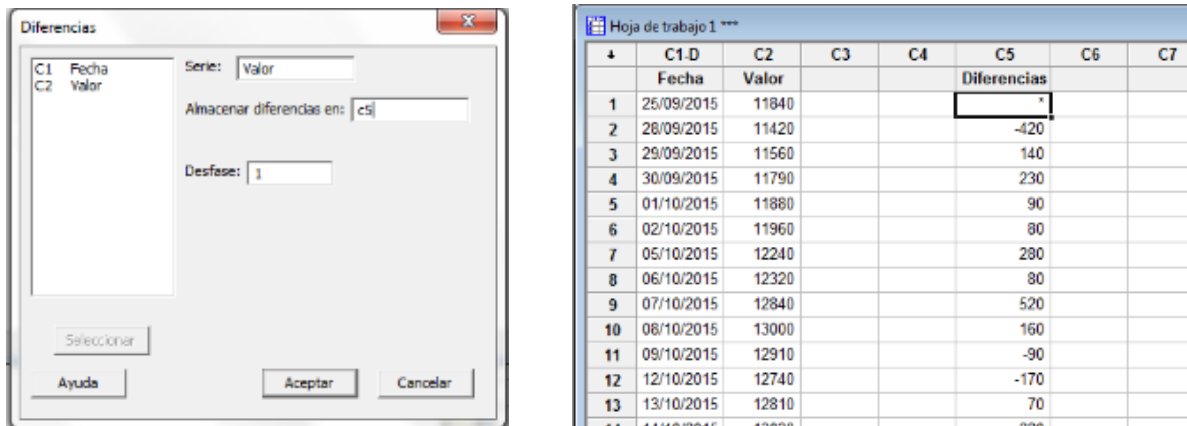


Figura 3. Aplicación de diferencias. Fuente: Ejecución en programa Minitab

Se obtiene nuevamente el grafico de auto correlación final y parcial, pero considerando la serie ya diferenciada. Al obtener la gráfica de Auto correlación se observa que la serie ya está estacionaria, después de aplicar la diferencia. La *Figura 4* muestra los histogramas de las auto correlación final y parcial de los valores diferenciados.

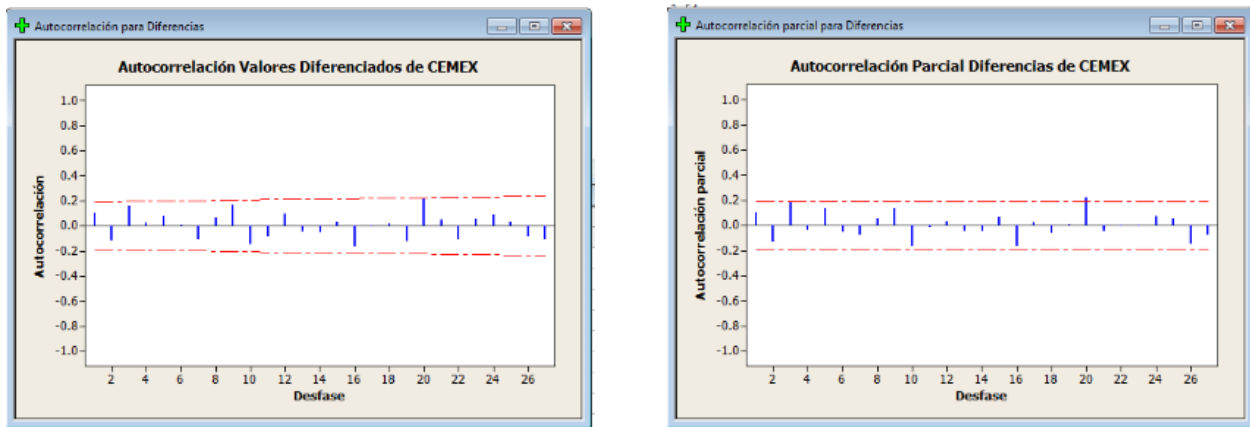


Figura 4. Histogramas de la serie con una diferencia. Fuente: Ejecución en programa Minitab

2) Fase de Identificación

Posterior a convertir si es necesario la serie a estacionaria, la fase 2 consiste en identificar los resultados de las funciones de auto correlación simple y parcial (FAC y FACP), considerando los modelos teóricos de la metodología ARIMA.

De acuerdo a los resultados de las gráficas de FAC y FACP, hay que identificar el modelo a utilizar, de acuerdo a los datos teórico del modelo, resultado en un modelo ARIMA (1, 1, 1), considerando para el modelo Autorregresivo un valor de $p = 1$, para el modelo de Medias móviles un valor de $q = 1$ y para las diferencias de 1 para d .

La **Figura 5**, indica el modelo correspondiente a partir de los resultados obtenidos

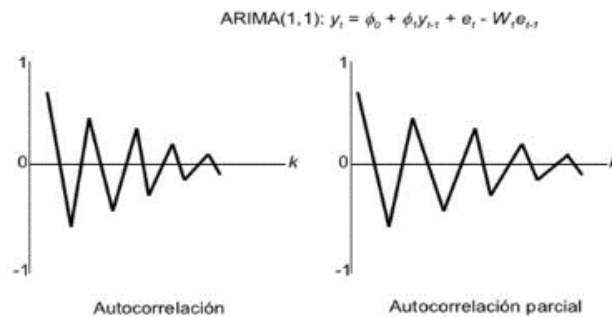


Figura 5. Modelo teórico ARMA (1,1).

3) Fase de Estimación

En la fase de estimación consiste en aplicar el modelo ARIMA con cada uno de los valores identificados en p , q y d . Los pasos en Minitab es eligiendo el *Menú Estadísticas*, la opción *Series de Tiempo* y se selecciona la función ARIMA, en el recuadro se establece los valores correspondientes a los parámetros identificados, tal como se muestra en la **Figura 6**.

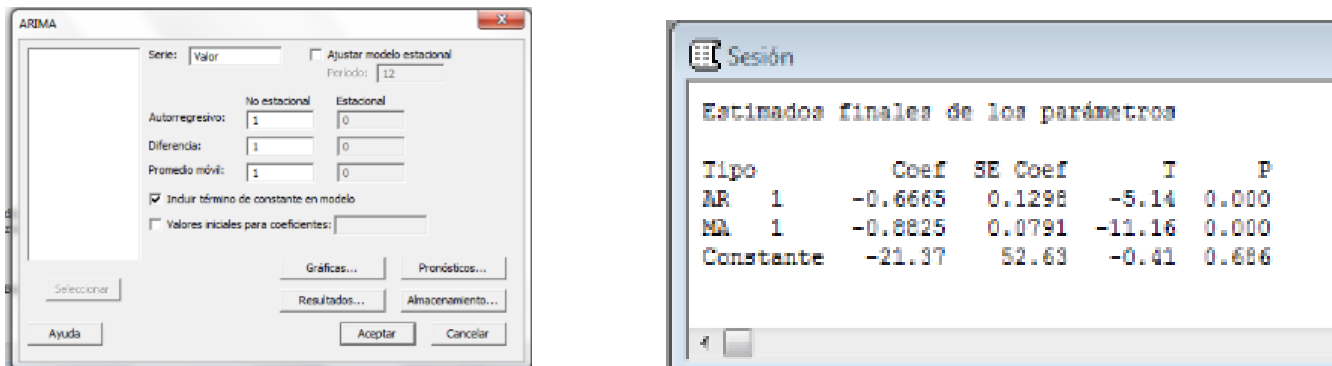


Figura 6. Aplicación del modelo ARIMA (1,1,1). Fuente: Ejecución en programa Minitab

Al analizar los coeficientes de auto correlación para verificar si los coeficientes deben permanecer en el modelo, se observa que los valores resultantes para el error estándar (T) son mayores a 2 en valor absoluto, por lo que los parámetros son correctos.

Así mismo se verifica si los coeficientes obtenidos cumplen las condiciones de estacionariedad e inversibilidad, por lo que las dos condiciones se cumplen. La **Tabla 1** indica las condiciones y los valores resultantes.

Tabla 1. Resultados obtenidos del modelo ARMA.

Modelo	Condición de estacionariedad	Condición de inversibilidad
ARMA(1,1)	$ \varphi_1 < 1$ $\varphi_1 = -0.6665$ $ -0.665 < 1$	$ w_1 < 1$ $w_1 = 0.8825$ $ 0.8825 < 1$

4) *Fase de Evaluación.*

Esta fase consiste en comprobar si el modelo es el correcto estadísticamente, analizando los residuales de las funciones de auto correlación.

Los pasos en Minitab son *Menú Estadísticas, la opción Series de Tiempo, se selecciona la función ARIMA*, en el cuadro se establece la *opción gráfica* seleccionando las opciones de ACF de residuos y PACF de residuos. Los histogramas resultantes se indican en la **Figura 7**.

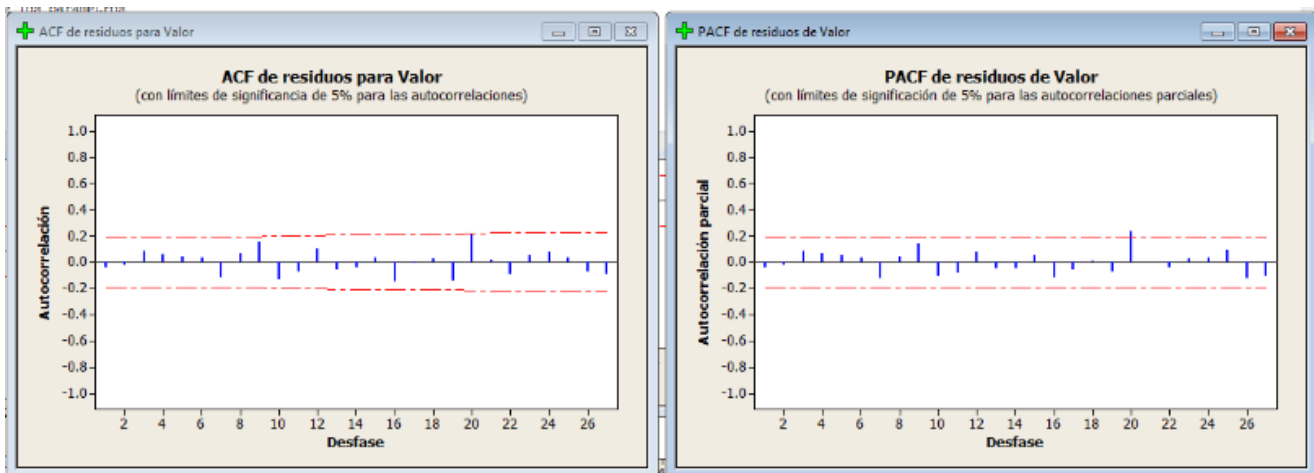
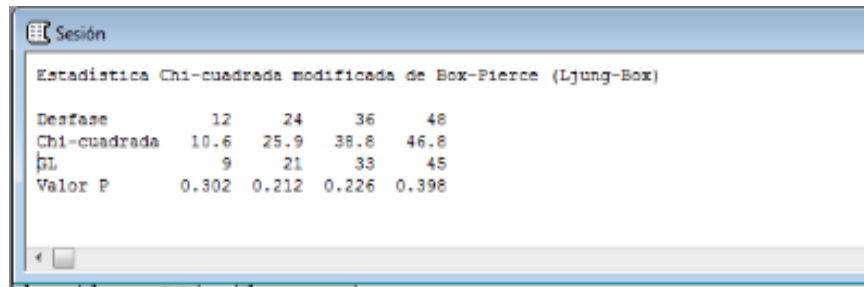


Figura 7. Residuales del modelo ARIMA (1,1,1). Fuente: Ejecución en programa Minitab

También se verifica con los resultados del valor P (probabilidad) de la estadística de Chi-cuadrada, verificando que cumplan con un 95% de confianza, por lo que los valores son suficientes a 0.05 por lo que el modelo es suficiente, como se indica en la **Figura 8**.



	12	24	36	48
Desfase				
Chi-cuadrada	10.6	25.9	38.8	46.8
PL	9	21	33	45
Valor P	0.302	0.212	0.226	0.398

Figura 8. Valor de la probabilidad. Fuente: Ejecución en programa Minitab

5) *Fase de Pronostico.*

Después de la validación se procede a generar los pronósticos de 60 días a partir de los datos propuestos. El procedimiento en Minitab es *Menú Estadísticas, la opción Series de Tiempo, se selecciona la función ARIMA, en el cuadro se establece la opción Pronósticos*, indicando la cantidad de elementos a pronosticar, así mismo en la opción de graficas se determina la gráfica de series de tiempo incluyendo pronósticos, tal como se indica en la **Figura 9**.

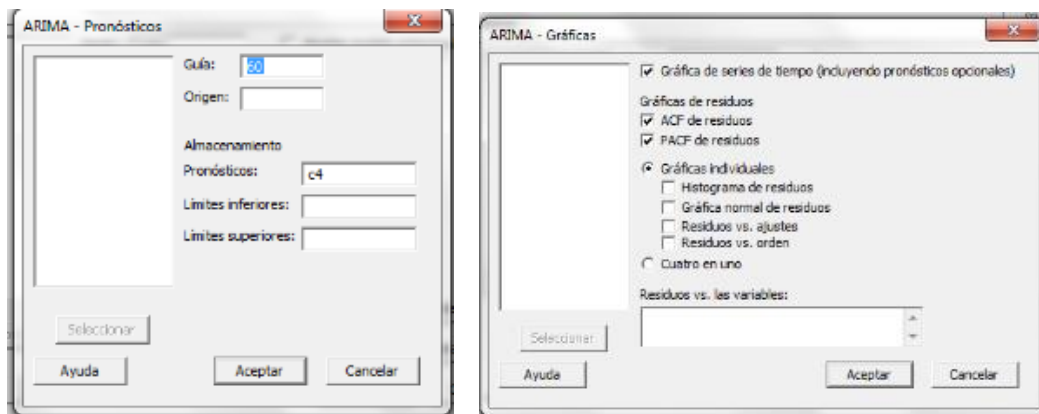


Figura 9. Configuración de pronósticos y gráficas. Fuente: Ejecución en programa Minitab

La **Figura 10** muestra los resultados de los pronósticos a partir del periodo 111, y la gráfica correspondiente.

Minitab - Sin título - [Sesión]

Archivo Editar Datos Calc Estadísticas Gráfica Editor Herram

Pronósticos del periodo 110

Período	Pronóstico	Límites 95%		Actual
		Inferior	Superior	
111	10534.4	9962.1	11106.6	
112	10470.1	9569.1	11371.1	
113	10491.6	9401.5	11581.6	
114	10455.9	9177.2	11734.6	
115	10458.3	9032.1	11884.5	
116	10435.3	8865.4	12005.3	
117	10429.3	8733.9	12124.7	
118	10412.0	8596.0	12228.0	
119	10402.1	8475.5	12328.8	
120	10387.3	8354.5	12420.2	
121	10375.8	8243.1	12508.6	
122	10362.1	8133.3	12591.0	
123	10349.9	8029.4	12670.4	
124	10336.7	7927.7	12745.7	
125	10324.1	7829.9	12818.3	
126	10311.1	7734.5	12887.7	
127	10298.4	7642.0	12954.9	
128	10285.5	7551.5	13019.5	
129	10272.7	7463.4	13082.1	
130	10259.9	7377.1	13142.7	
131	10247.1	7292.7	13201.5	
132	10234.3	7210.0	13258.6	
133	10221.4	7128.8	13314.1	
134	10208.6	7049.2	13368.1	
135	10195.8	6970.9	13420.7	
136	10183.0	6893.9	13472.0	
137	10170.2	6818.1	13522.2	
138	10157.3	6743.6	13571.1	
139	10144.5	6670.1	13619.0	
140	10131.7	6597.6	13665.8	
141	10118.9	6526.2	13711.6	
142	10106.0	6455.6	13756.5	

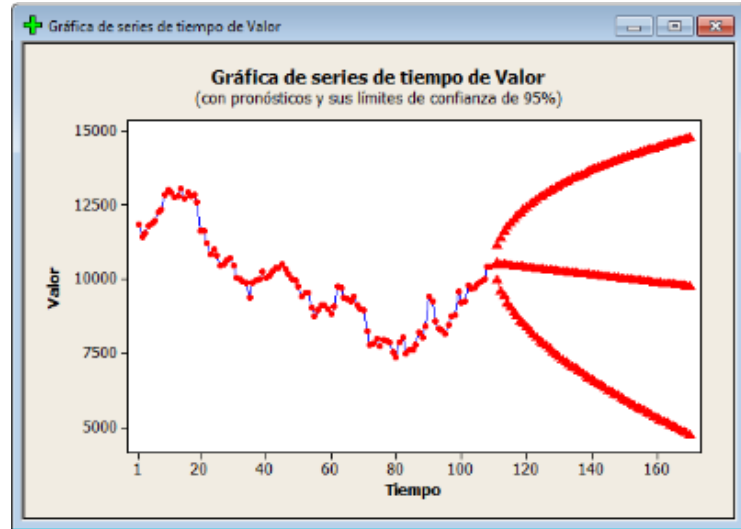


Figura 10. Pronósticos y gráfica. Fuente: Ejecución en programa Minitab

CONCLUSIONES

Si se cuenta con series de tiempo, la metodología ARIMA es una herramienta fácil y eficiente para la predicción de datos de cualquier variable productiva.

Utilizando el software Minitab es posible realizar el proceso complejo de modelar una serie de tiempo con el propósito de pronosticar nuevos datos dentro de un periodo de corto, mediano o largo plazo.

La metodología ARIMA permite a través de cada una de sus etapas, establecer si los datos históricos son útiles o no para la predicción del futuro. Con los resultados obtenidos en la demostración del caso de estudio, se puede identificar fácilmente nuevos datos que determinen el futuro, y así a la empresa le permita tomar decisiones oportunas.

El objetivo propuesto para este artículo se cumplió, generando un conjunto de nuevos valores para la serie temporal utilizada, así como ejemplificar paso a paso la metodología ARIMA y la secuencia de la misma con el software Minitab; considerando que hay literatura que describe de forma aislada cada una de las etapas o literatura en otro idioma.

LITERATURA DE APOYO

- Aguirre J. (1994). Introducción al tratamiento de series temporales. Madrid España. Ediciones Días de Santos.
- Box, G.E.P., y Jenkins, G.M., (1976). Time Series Analysis Forecasting and Control, 2nd edn. San Francisco, CA: Holden-Day. Box
- Cortez J. (1994). Técnicas Estadísticas de predicción aplicable en campo empresarial. Lima: Universidad del Pacifico. CDD: 03.01.02/C77
- Cortez J. (2011). Aplicación de series de tiempo en el monitoreo estructural. Tesis de grado de Ingeniería. Universidad Nacional Autónoma de México. México D.F.
- González M. (2009). Técnicas de predicción económica. Universidad del país Vasco. facultad de Ciencias Económicas y Empresariales. Consultado el 29 de febrero de 2020 en <https://addi.ehu.es/bitstream/10810/12493/1/05-09pil.pdf>
- Hanke J. y Reitsch A. (1996). Estadística para Negocios. McGraw Hill.
- Infobolsa (2016). Cotizaciones históricas Cemex, CPO. Consultado el 04 de marzo de 2020 en http://www.infobolsa.es/cotizacion-mx/historico-cemex____cpo
- Pankratz A. (1983). Forecasting with univariate Box-Jenkins model. E.U.A. John Wiley & Sons.
- Pinzon L. (2009). Taller series de tiempo con SPSS. Memorias VIII Encuentro Nacional de Educación Matemática y Estadística. Centro Virtual de Noticias.
- Ruiz, R. S. G., López-Herrera, F., & Aké, S. C. (2018). Mecanismos de transmisión de las condiciones macroeconómicas al mercado crediticio. Panorama Económica, 13 (26), 82-112.
- Sánchez L. (2013). Pronóstico de la producción de leche, mediante modelos ARIMA. Caso UBPC "Maniabo". Tesis de master en Bioestadística. Facultad de Matemática y Computación "Universidad de La Habana".
- Scientific Computing Associates, Corp. (1994). Manual Forecasting and time series analysis using the sca statistical system. Consultado el 29 de enero de 2020 en http://scausa.com/SCADocs/SCAFTS_V1.pdf