

Método de descomposición de series de tiempo aplicados al movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)

Cañaveral López; Carlos de Jesús
Universidad Autonoma de Chiapas
jesustww63@gmail.com

Resumen

El presente trabajo tiene por objeto realizar una síntesis preliminar del método clásico de descomposición de series de tiempo abordando bajo este enfoque la problemática de las series de tiempo tanto empíricos como formales. En este sentido se analizará el método de descomposición para el caso univariado a fin de encontrar el modelo que se ajuste adecuadamente a una muestra determinada de datos como es el del movimiento nacional de carga de altura general contenerizada de enero de 2005 a

marzo de 2014 en millones de toneladas. Esta metodología se aplica a series de tiempo tales como el producto interno bruto, los registros de empleo y desempleo, las observaciones del crecimiento de la población a través del tiempo o las cifras de carga general que genera el sistema portuario nacional de la Secretaría de Comunicaciones y Transportes desglosada como de carga general suelta y contenerizada, graneles, petróleo, vehículos automotores y cruceros entre otros.

Palabras clave: Series de tiempo, movimiento nacional, carga de altura
JEL: N76

Method of decomposition of time series applied to the national movement of containerized general height cargo (2005: 01-2014: 03)

Abstract

The purpose of this paper is to make a preliminary synthesis of the classical method of decomposition of time series, addressing the problem of both empirical and formal time series under this approach. In this sense, the decomposition method for the univariate case will be analyzed in order to find the model that adequately fits a given sample of data such as the national containerized general height cargo movement from

January 2005 to March 2014 in millions of tons. This methodology applies to time series such as gross domestic product, employment and unemployment records, observations of population growth over time or general load figures generated by the national port system of the Ministry of Communications and Transportation broken down as general cargo loose and containerized, bulk, oil, motor vehicles and cruises among others.

Palabras clave: Time series, national movement, height cargo
JEL: N76

1. Introducción

El presente trabajo tiene por objeto realizar una síntesis preliminar del método clásico de descomposición de series de tiempo abordando bajo este enfoque la problemática de las series de tiempo tanto empíricos como formales. En este sentido se analizará el método de descomposición para el caso univariado a fin de encontrar el modelo que se ajuste adecuadamente a una muestra determinada de datos como es el del movimiento nacional de carga de altura general contenerizada de enero de 2005 a marzo de 2014 en millones de toneladas. Esta metodología se aplica a series de tiempo tales como el producto interno bruto, los registros de empleo y desempleo, las observaciones del crecimiento de la población a través del tiempo o las cifras de carga general que genera el sistema portuario nacional de la Secretaría de Comunicaciones y Transportes desglosada como de carga general suelta y contenerizada, graneles, petróleo, vehículos automotores y cruceros entre otros; con el objetivo en un primer momento de encontrar el modelo adecuado que explique la realidad que estamos investigando, segundo estimarlo, tercero validarlo, y finalmente pronosticar la conducta de dicha variable, y de esa manera tener cierta perspectiva sobre el futuro que es incierto, o bien utilizarlo para análisis y evaluación de estructura del fenómeno, producción, ventas, márgenes de ganancias, costos, productividad, entre otros múltiples aspectos que puede aplicarse esta metodología.

2. Marco teórico

En general existen modelos ingenuos o producto de la experiencia y los formales que se fundamentan en una metodología científica, los primeros han sido utilizados de manera muy general que solucionan los problemas de pronóstico de corto plazo y por su naturaleza son de poco costo, entre los cuales se encuentran los simples promedios móviles, los promedios móviles centrados, los dobles promedios móviles, los promedios móviles ponderados, el modelo simple exponencial entre otros. Existen modelos más sofisticados como son los dos parámetros de Holt, la estacionalidad de Winter o una combinación de los mismos, y el método del Census X11.

En esta investigación aplicaremos el método de estimación de mínimos cuadrados ordinarios (método diseñado por Gauss desde los 1800's y ampliamente utilizado en la estadística y la econometría y otras áreas del conocimiento) para separar cada uno de los componentes de la series a analizar a fin de convertirlos en un proceso estacionario conocido como ruido blanco situación analizada por George Box y Gwilm Jenkis (1976), actualmente conocidos como modelos Autorregresivos, integrados y promedios móviles (ARIMA), y los desarrollos posteriores como son: los procesos autorregresivos con heterocedasticidad (ARCH) así como también los procesos autorregresivos generalizados con heterocedasticidad condicional (GARCH) (Engel, 1982:987-1008).

En nuestro caso analizaremos el método clásico de descomposición de una serie de tiempo, el cual consiste en diagnosticar si la serie tiene un patrón de conducta determinado con el objeto de conocer la parte sistemática de la serie y la no sistemática o aleatoria, según Wold (1938). Este autor plantea que todo proceso estocástico se compone de una parte sistemática (tendencia, estacionalidad y ciclo) y otro componente estocástico referida al error de la serie, mientras que Nerlove, Grether y Carvalho (1979) señalan que las series cronológicas se encuentran integradas por cuatro componentes: tendencia, ciclo, estacionalidad y movimiento irregular. En este sentido, confluyen diversos autores en un consenso de que cualquier series de tiempo puede tener dos componentes: la parte sistemática y la aleatoria, para ello es necesario contar con procedimientos que nos permitan extraer esos componentes dado que no son directamente observables algunos aspectos de las series y dependiendo del tipo de componente y la forma como se defina el modelo el método variarán de un caso a otro.

El modelo general de una serie de tiempo puede definirse de la siguiente manera:

$$X_t = T + E + C + \varepsilon \quad (1)$$

Dónde:

X_t = La variable a descomponer

Método de descomposición de series de tiempo aplicados al movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)

T = La tendencia

E = Estacionalidad

C = Es el ciclo

ε = error o parte aleatoria del modelo

El componente sistemático puede ser de dos tipos:

A) Aditivo

$$X_t = T + E + C + \varepsilon \quad (2)$$

B) Multiplicativo

$$X_t = TEC\varepsilon \quad (3)$$

En el caso b, con una transformación logarítmica podemos tratarla de manera aditiva y utilizar el método de mínimos cuadrados ordinarios para su estimación, de no ser así, es preciso estimarlo mediante métodos no lineales.

3. Metodología

En este artículo se desarrollará la investigación de la forma siguiente: 1) análisis de los datos; 2) definición del modelo formal; 3) Extracción de los componentes sistemáticos en caso de existan (tendencia, estacionalidad y ciclo); 4) Análisis del proceso estacionario de la serie y validación del modelo; 5) Pronóstico de la serie, y 6) Conclusiones.

4. Análisis y discusión de los resultados

En Makridakis, Wheelwright y Hyndman (1998) señalan que el pronóstico cuantitativo puede ser aplicado si existen las siguientes tres condiciones:

1. Información histórica disponible

2. Que los datos sean cuantificados en forma numérica, y

3. Que algunos de los patrones pasados de la series continúen en el futuro.

Este último refiere al supuesto de continuidad el cual tiene que aplicarse independientemente si los métodos son simples o complejos para realizar los pronósticos de series de tiempo tanto cuantitativos como cualitativos.

Por su parte, Pindick y Rubinfeld (1985) establecen que los modelos de series de tiempo proporcionan métodos sofisticados de extrapolación de series de tiempo, su aplicación depende si la variable analizada corresponde a conductas de patrones de tendencia de largo plazo o de corto plazo, los ciclos responden al desenvolvimiento histórico de una economía determinada, por ejemplo la conducta cíclica de una economía capitalista, de tipo financiero o de negocios o en todo caso corresponde a su conducta innata propio del fenómeno que se está investigando (probablemente puede ser físico o biológico entre otros), si la serie muestra alguna estructura o patrón o tiene alguna estacionalidad, es decir, en cada momento del tiempo, diaria, semanal, quincenal, mensual, bimestral, trimestral, etc., con el objeto de adecuar el mejor modelo que explique dicha realidad ya sea variable continua o discreta; en nuestro caso solo se abordará el caso discreto.

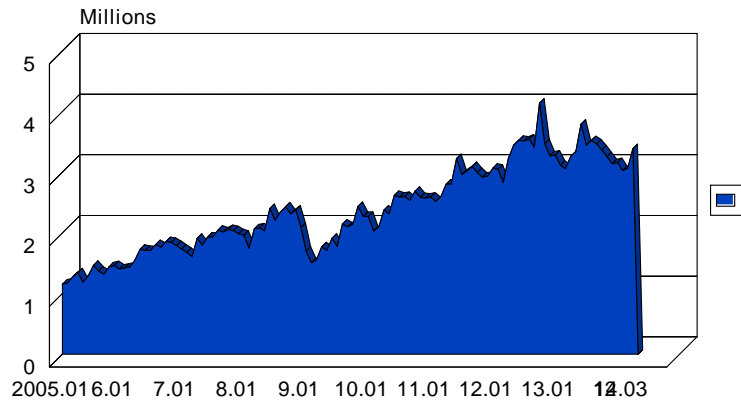
Veamos la siguiente gráfica que refiere al movimiento nacional de carga de altura contenerizada para el periodo que va de enero de 2005 a marzo de 2014, cifras en millones de toneladas procesadas del sistema portuario nacional de la Secretaría de Comunicaciones y Transportes.

Método de descomposición de series de tiempo aplicados al movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)

Gráfica no. 1 Movimiento Nacional de Carga de Altura

(Enero 2005 - Marzo 2014)

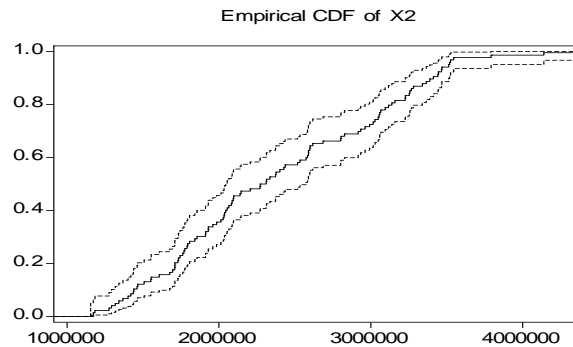
(millones de toneladas)



Como puede observarse, la gráfica no. 1 muestra una tendencia a lo largo del periodo señalado aunque presenta alta variación con un máximo de 4.139 millones de toneladas observándose una dramática caída del 1.154 millones respectivamente. Una mediana de 2.271 millones de toneladas mientras que la media del periodo alcanzó los 2.370 millones de toneladas de carga durante el periodo analizado. Así mismo registra un sesgo de su distribución hacia la derecha y una curtosis menor a 3 lo que refleja una anchura en los brazos más allá de lo normal.

En la siguiente gráfica podemos observar su distribución empírica acumulada, el cual cuando la muestra es grande ésta es muy próxima a función de densidad teórica sobre la recta real. Por ello, se dice que la función empírica acumulada converge a la función de densidad de manera uniforme para todos de los valores de la variable analizada.

Gráfica no. 2 Distribución Empírica Acumulada



Por otro lado, sí observamos la función de autocorrelación y la de autocorrelaciones parciales registra un patrón autorregresivo de orden (1). Es necesario transformarlo en un proceso estacionario para definir el modelo a estimar.

4.1 Definición formal del modelo

Como puede observarse el modelo en general refleja un componente de tendencia, otro referido a la estacionalidad, un tercero al ciclo y finalmente un error de estimación o residual. Lo que más adelante llamaremos serie estacionaria o ruido blanco con el que estimaremos el modelo y pronosticaremos la serie. Es importante señalar que cada estimación de cada componente es necesario realizar la prueba de significancia de cada factor mediante la prueba t, F y el R^2 o bondad de ajuste del modelo.

$$X_t = \beta_1 T + \beta_2 E + \beta_3 C + \beta_4 + \varepsilon \quad (4)$$

Dónde:

X_t = La serie a analizar

β_1 = El parámetro a estimar del componente de la tendencia

β_2 = El parámetro a estimar del componente de la estacionalidad

β_3 = El parámetro a estimar del componente del ciclo.

Método de descomposición de series de tiempo aplicados al movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)

ε = Es el componente aleatorio o residuo.

4.2 Extracción o sustracción de cada uno de los componentes: tendencia, estacionalidad y ciclo.

En este apartado se procederá de manera metodológica ir sustrayendo cada uno de los componentes sistemáticos de la serie a fin de convertirlo en un proceso estacionario o de ruido blanco, es decir, que no contenga ningún componente fijo y así construir un modelo que posteriormente sea usado para fines de pronóstico, y en segundo momento integrar la serie con cada uno de los componentes o factores sistemáticos extraídos anteriormente.

a. Componente de tendencia

El componente de tendencia puede definirse de manera lineal o ajustando un polinomio de segundo orden o de un orden superior, en nuestro caso solamente es relevante la tendencia lineal. Los α 's son valores de los parámetros estimados y los que se encuentran entre paréntesis refieren a las t 's de student calculadas que el software eviews (econometric views) que reporta respectivamente.

$$X_t = \alpha_0 + \alpha_1 T + \varepsilon \quad (5)$$

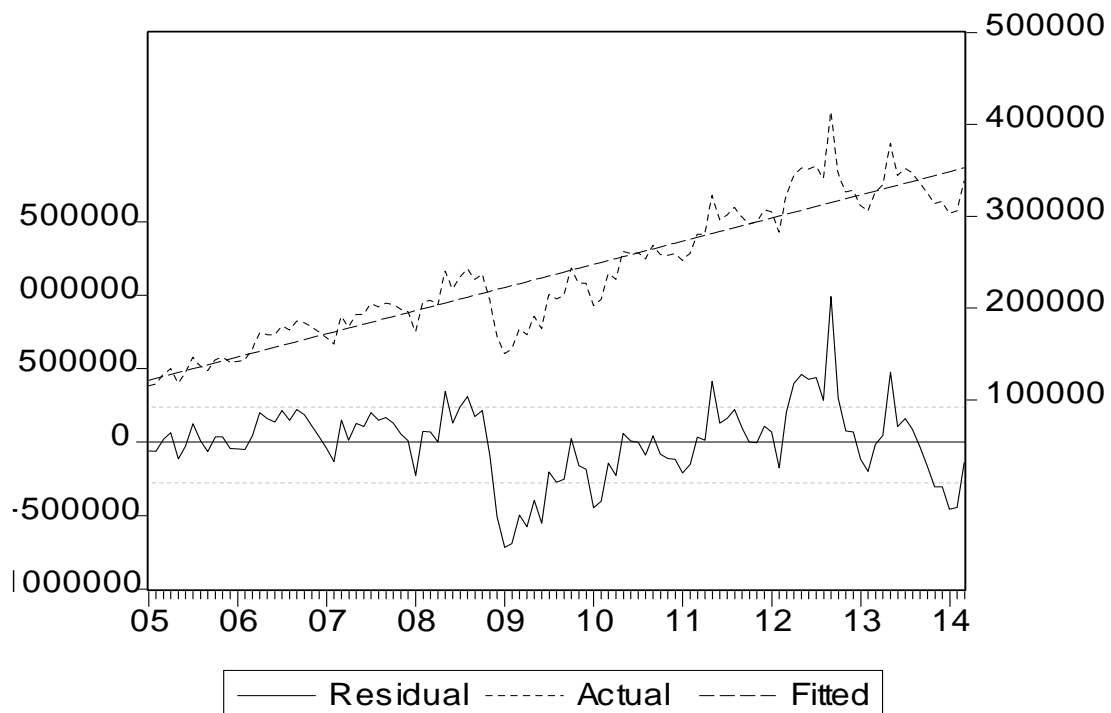
Dónde:

$$\alpha_0 = 1193234 \text{ (24.23205)}$$

$$\alpha_1 = 21018.86 \text{ (27.53973)}$$

$$R^2 = 0.874342$$

Grafica 3. Estimación de la tendencia del movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)



b. Componente de estacionalidad

La estacionalidad puede estimarse según el tipo de serie de tiempo que se evalúe, por ejemplo si los datos son semestrales se generaría dos variables binarias o mudas, cuatrimestrales cuatro variables binarias, trimestrales cuatro variables binarias, bimestrales seis variables binarias, y finalmente si es mensual se generarían 12 variables binarias y así sucesivamente. En nuestro caso, el modelo estimado solamente salieron significativas las variables binarias correspondiente a: δ_1 = Dummy de enero; δ_2 = Dummy de febrero, δ_5 = Dummy de mayo, y δ_9 = Dummy de septiembre.

$$X_t = \delta_1 D_{01} + \delta_2 D_{02} + \dots + \delta_{12} D_{12} + \varepsilon \quad (6)$$

Dónde

$$\delta_1 = -243871.4 \quad (-3.4518)$$

Método de descomposición de series de tiempo aplicados al movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)

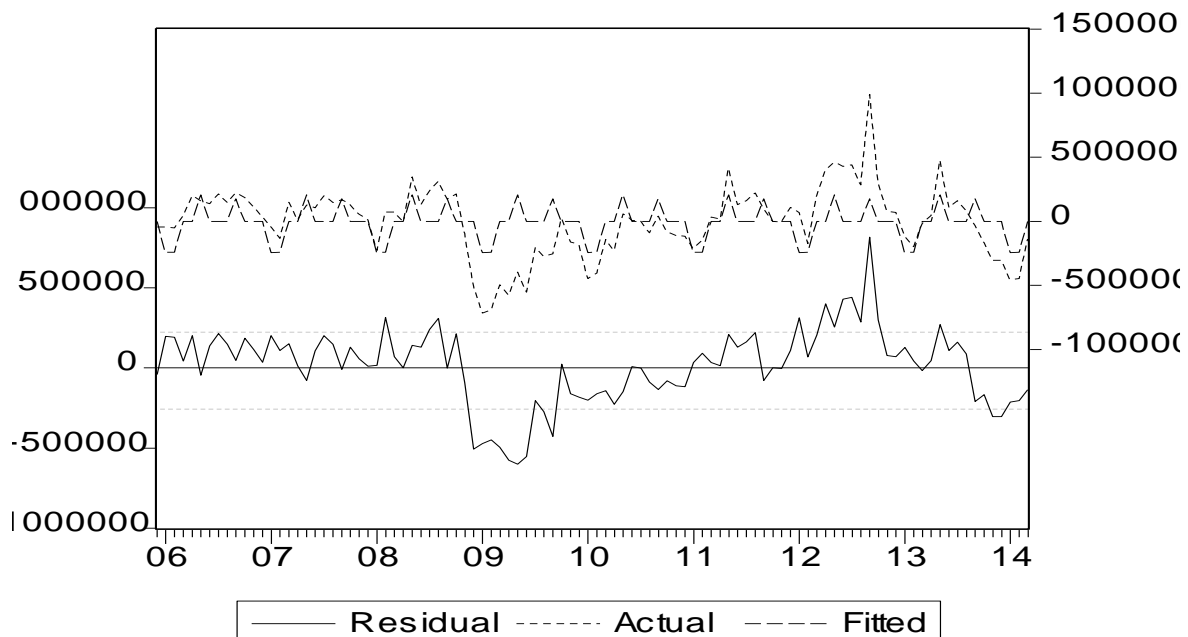
$$\delta_2 = -241425.5 \text{ (-3.014639)}$$

$$\delta_5 = 205930.8 \text{ (2.424360)}$$

$$\delta_9 = 175849.0 \text{ (2.070216)}$$

$$R^2 = 0.229065$$

Grafica 4. Estimación de la estacionalidad del movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)



c. Componente de ciclo

Brockwell y Davis (2002:111-136) en su capítulo de análisis espectral definen de manera concreta la forma de abordar el ciclo de cualquier serie de tiempo mediante la metodología de la función senoidal:

$$X_t = \theta_1 \cos(\omega t) + \theta_2 \sin(\omega t) \quad (7)$$

Dónde:

$\omega = (\pi/4)$ es la fase o frecuencia del periodo.

Θ_1 = Parámetro estimado para el caso del coseno.

Θ_2 = Parámetro estimado para el caso del seno.

Esta serie no contiene ciclo por ello solamente se hace mención de la metodología existente para el caso de que alguna serie tenga un componente sistemático como es el ciclo.

4.3 Análisis del proceso estacionario de la serie y validación del modelo

En este apartado se abordará lo que resta de la series una vez extraída los componentes sistemáticos a fin de analizar si la serie se comporta como un proceso estacionario o ruido blanco, en este sentido identificaremos el modelo con el criterio gráfico, así también utilizando la función de autocorrelación y las autocorrelaciones parciales. Uno de los criterios de identificación refiere si la serie se desvanece exponencialmente nos encontramos frente a un proceso autorregresivo y el orden de éste modelo, lo observamos con la autocorrelación parcial mediante la conducta de la misma, es decir, si cae de golpe y se mantiene dentro de las bandas de confianza el autorregresivo es de orden 1.

En caso de que su conducta sea diferente como un componente senoidal o estacional entonces refiere a un proceso promedio móvil y el orden depende de cómo en la gráfica se observe la conducta de la autocorrelación parcial: sí la serie cae verticalmente y se mantiene dentro de la banda de confianza entonces estamos frente aún promedio móvil de orden uno, en caso de sean dos o tres los que se encuentran fuera de la banda es el orden del promedio móvil. Veamos la conducta de la serie:

Como puede observarse el correleograma evalúa de forma visual la función de autocorrelación y las autocorrelaciones parciales del residuo de la regresión, en la medida que la autocorrelación se devanece exponencialmente la función de autocorrelación parcial cae de manera dramática, esto nos proporciona el criterio de elección del tipo de modelo que se tiene que estimar y es de un autorregresivo de orden 1, ya que

Método de descomposición de series de tiempo aplicados al movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)

posteriormente sigue una caminata aleatoria de la serie analizada, por ello, procedemos a estimar y pronosticar con este modelo doce periodos siguientes, para que posteriormente procedamos a integrar la serie. Por lo tanto, definimos el modelo de la forma siguiente:

$$X_t = \varphi_1 AR(1) \quad (8)$$

$$\varphi_1 = 0.756159 (0.066340)$$

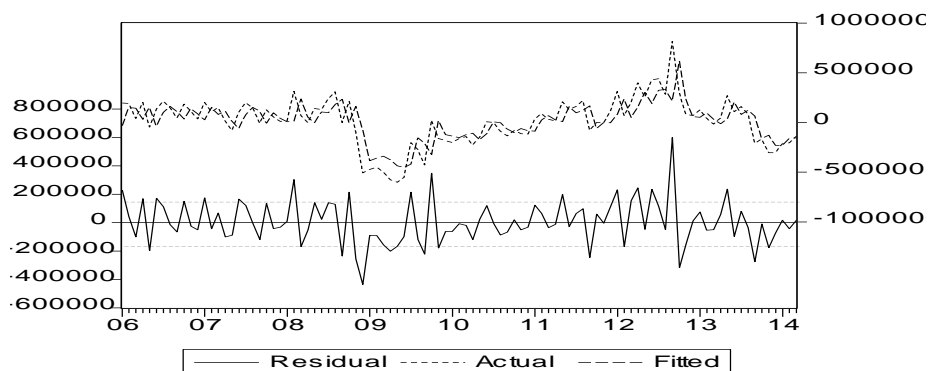
Dónde:

X_t = es la serie sin los componentes sistemáticos.

φ_1 = Es el parámetro estimado del autorregresivo de orden 1.

Nota. El dato que se encuentra en paréntesis corresponde al error estándar del parámetro estimado.

Grafica no. 5. Estimación del modelo AR(1) del movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)



Los valores que se encuentran entre paréntesis corresponden a los errores estándar del parámetro estimado en este caso de un autorregresivo de orden (1), este sentido la prueba de hipótesis sobre la significancia de la variable en el modelo comúnmente se realiza con la prueba t de student, sin embargo no es propio para este tipo de modelos

autorregresivos, es necesario modificarlo para poder aplicarse según los estándares o criterios econométricos; para ello, requiere aplicarse la prueba estadística correcta que actualmente se conoce como τ (tao) refiere a la prueba aumentada de raíz unitaria de Dickey – Fuller (1976) a fin de aceptar o rechazar la hipótesis nula de la significancia de la variable en el modelo, es decir si $\rho=1$, en sentido estricto si $|\rho|<1$ esto implica que si el modelo tiene raíz unitaria el parámetro estimado es significativo y verdadero, por ello, es necesario modificar la prueba t que comúnmente se utiliza para verificar la relevancia del componente del modelo, en caso de que sea así se concluye que el modelo estimado se ajusta a los estándares utilizados en la jerga econométrica.

La Prueba incrementada de Dickey – Fuller (1976-1979)

Los criterios de aceptación o rechazo en el caso de un modelo autorregresivo de orden 1 bajo del supuesto de que $\rho=1$ o en sentido estricto $|\rho|<1$ a fin de evaluar si el modelo tiene raíz unitaria lo calculamos según lo establece el artículo Distribution of the Estimators for Autoregressive Times Series with a Unit root en Dickey – Fuller (1979) de la siguiente manera:

$$N(\beta-1) = 99(0.756159 - 1) = -24.140259$$

$$\hat{\tau} = -24.140259/0.066340 = -363.886931$$

Posteriormente nos vamos a la tabla B5 de los valores críticos de la prueba Phillips Perron Z_p , y de la prueba de Dickey –Fuller basado en mínimos cuadrados ordinarios con coeficientes autorregresivos para el caso de orden 1, con $n= 100$ observaciones al 95% de confianza para el valor de la τ es de -7.9 en nuestro caso según el dato calculado rechazamos que el modelo estimado tenga raíz unitaria.

4.4 Pronosticar la parte aleatoria de la serie evaluada

El proceso de generación de datos conocidos como pronóstico se realiza de la siguiente manera, dado el modelo estimado:

Método de descomposición de series de tiempo aplicados al movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)

$$X_{t+1} = \varphi ar(t)$$

$$X_{t+2} = \varphi Ar(t+1)$$

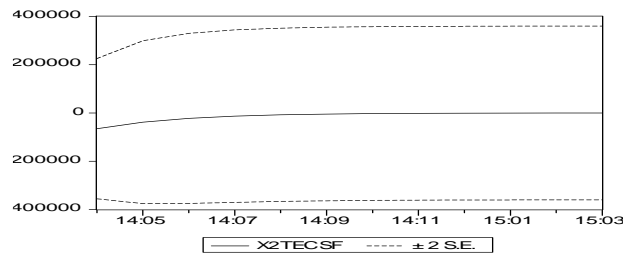
...

$$X_{t+12} = \varphi Ar(t+11) \tag{9}$$

Este modelo nos genera el pronóstico y las bandas de confianza:

$X_{f \pm 2}$ veces el error estándar de la regresión.

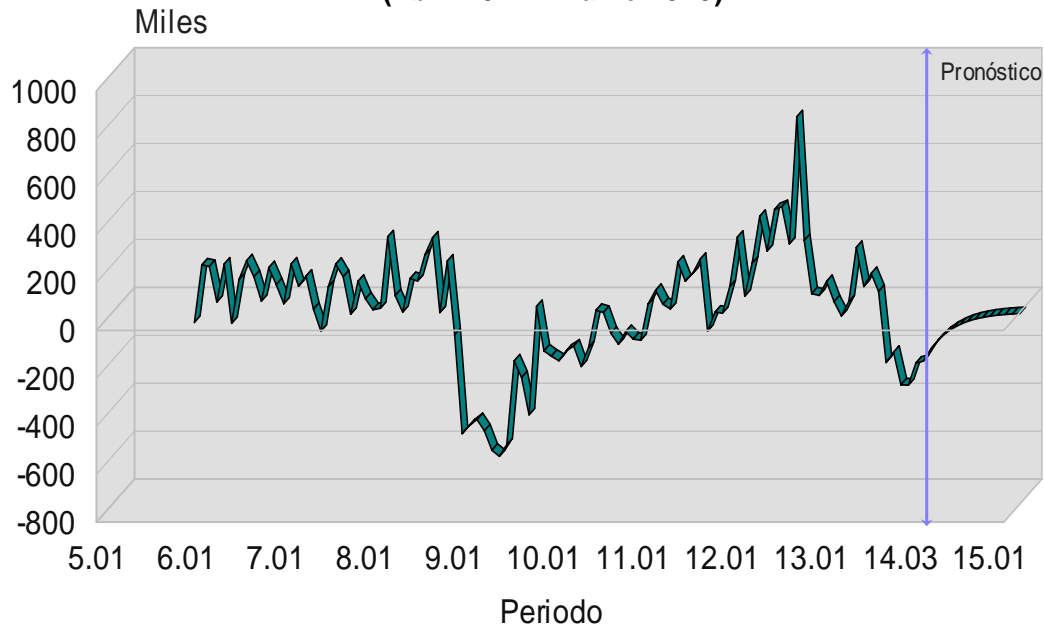
Grafica 6. Bandas de confianza del pronóstico del movimiento nacional de carga de altura general contenerizada (2005:01-2014:03)



Considerando la serie completa de 2005.01 a 2014.03 adicionalmente incorporando los doce meses de pronóstico así como la integración de los componentes sistemáticos referidos a la tendencia y la estacionalidad obtenemos para el siguiente año de la siguiente manera:

Pronóstico de la serie con un ar(1)

Pronóstico de la Serie sin tendencia y estacionalidad (Abril 2014 - Marzo 2015)



Fuente: Elaboración propia con datos de la sct.

5. Conclusiones

La aplicación metodológica del análisis de series de tiempo al movimiento nacional de carga de altura y cabotaje contenerizada durante el periodo de enero de 2005 a marzo de 2014, nos permite validar la metodología que es aplicable a cualquier series sea de la economía real o de la economía financiera, así mismo verificamos que el modelo no pasa la prueba t modificada con la metodología de Dickey – Fuller, es necesario seguir evaluando las distintas series de esta área a fin de validar de manera consistente esta herramienta a fin ampliar nuestro conocimiento en la aplicación econométrica en las diferentes áreas de la economía marítima.

6. Referencias

- Box, George and Gwilym Jenkins, (1976), *Time Series Analysis: forecasting and control*, Revised Edition, Holden Day.
- Brockwell, Peter J., and Richard Davis A., (2002). *Introduction to Time Series and Forecasting* Second Edition, Springer Verlag New York, Inc.
- Dickey, David A and Wayne A. (1978). Fuller Distribution of the Estimators for Autoregressive Time Series With a Unit Root, *Journal of the American Statistical Association*, (74) Number 366, Theory and Methods Section.
- Dickey, David A. (2005). Stationarity Issues in Time Series Models. *Paper 192-30, SUGI 30 Proceedings*, SAS Institute Inc., Proceeding of the Thirtieth Annual SAS User Group International Conference, Cary, NC: SAS Institute Inc.
- Fuller, Wayne A. (1996). *Introduction to Statistical Time Series*. Second Edition, John Wiley & Sons Inc.
- Gujarati, Damodar N. (1999). *Econometría*. Tercera Edición, McGraw Hill.
- Koutsoyiannis, A., (1979). *Theory of Econometrics an introductory exposition of econometric model*. Second Edition, The Macmillan Press LTD.
- Makridakis, Spyros, Wheelwright Steven C. y Rob J. Hyndman, (1998). *Forecasting Methods and applications*. Third edition, John Wiley & Sons, Inc.
- Nerlove, Marc, Grether David, and José L. Carvalho, (1995). *Analysis of Economic Time Series: A Synthesis*. New Milford, Connecticut, USA.: Emerald Group Pub Ltd.
- Pindyck, Robert S & Daniel L. Rubinfeld (1985). *Econometric Models and Economic Forecasts*. Second Edition, International Student Edition.
- Wold Herman, (1938). A study in the analysis of Stationary time series. *Uppsala Almqvist & Wiksell*, VIII, 214 S. Kr. 6.