

ISSN: 2322-7672

III Encuentro de Investigación Formativa Ingeniería Industrial Medellín

Memorias

Grupo de Investigación en Sistemas
Aplicados en la Industria (GISAI)



**Universidad
Pontificia
Bolivariana**

© xxxxxxxxxxxxxxxxxxxxxxxxxxxx

© Editorial Universidad Pontificia Bolivariana

III Encuentro de Investigación Formativa - Memorias

ISSN: 2322-7672

Primera edición, 2013

Escuela de Ingenierías

Facultad de Ingeniería Industrial

Gran Canciller UPB y Arzobispo de Medellín: Mons. Ricardo Tobón Restrepo

Rector General: xxxxxxxxxxxxxxxxxxxxxxxxxxxx

Vicerrector Académico: Pbro. Jorge Iván Ramírez Aguirre

Editor: Juan José García Posada

Coordinación de producción: Ana Milena Gómez C.

Diagramación: Geovany Snehider Serna Velásquez

Corrector de estilo: Monica Patricia Ospina Toro

Dirección editorial:

Editorial Universidad Pontificia Bolivariana, 2013

Email: editorial@upb.edu.co

www.upb.edu.co

Telefax: (57) (4) 354 4565

A.A. 56006 - Medellín - Colombia

Radicado: 1117-22-03-13

Prohibida la reproducción total o parcial, en cualquier medio o para cualquier propósito sin la autorización escrita de la Editorial Universidad Pontificia Bolivariana.

Comparación de metodologías estadísticas en el pronóstico de la demanda

Sebastián Ramírez

Ingeniería Industrial UPB Colombia
sebastian.ramirezag@alfa.upb.edu.co

Daniela González

Ingeniería Industrial UPB Colombia
daniela.gonzalezlo@alfa.upb.edu.co

Juan Cardona

Ingeniería Industrial UPB Colombia
juanestebean.cardona@alfa.upb.edu.co

Marisol Valencia Cárdenas,

Docente de Ingeniería Industrial de UPB, Medellín
marisol.valencia@upb.edu.co

Resumen

Un pronóstico de demanda estima de forma anticipada las cantidades a ser vendidas, con el fin de facilitar la planeación de sus inventarios y logística necesaria para el abastecimiento. En este trabajo se busca aplicar tres modelos estadísticos para realizar pronósticos comparando la eficiencia entre ellos, en una empresa del sector de reciclaje.

Palabras clave

Pronóstico, modelo dinámico, suavizamiento exponencial.

Introducción

La aleatoriedad de la demanda es frecuente debido a los cambios que se dan en los mercados por cuestiones como crisis económicas, apertura de relaciones comerciales entre otros, es por esto que es importante tener en cuenta modelos estadísticos que logren caracterizarla para hacer un buen pronóstico cada vez más acercado a la realidad futura. En la industria nacional son muy poco frecuentes los métodos para el tratamiento de aleatoriedad de la demanda en los sistemas de diseño y control de inventarios, por su desconocimiento y falta de herramientas propias, necesarias en la optimización de inventarios, mejora de servicio, satisfacción de clientes, aspectos fundamentales en la dinámica productiva del país. (Vidal, marzo 2008).

Algunos de los modelos de pronóstico clásicos basan sus estimaciones en distribuciones probabilísticas e intentan concluir sobre el comportamiento de las variables involucradas y pronosticar respuestas.

Existen trabajos que incorporan el tratamiento de la demanda aleatoria en sistemas de inventarios, usando distribuciones de probabilidad (Cohen et al (1986), Gurnani y Tang (1999), Sethi et al (2003)) , pero también, aquellos que utilizan modelos de pronósticos de demanda para optimizar el manejo de los inventarios, como los de suavización exponencial, media móvil y generados con distribuciones de probabilidad, entre otros, como se ha visto en la recopilación de Gutiérrez y Vidal (2008), reflejando la importancia de este tratamiento de la demanda estimada para mejorar la planeación de inventarios en una empresa manufacturera.

Existen métodos de pronósticos de demanda, tradicionalmente utilizados en la optimización de inventarios, donde se evidencia una falta de validación necesaria de supuestos en los errores (Hillier y Lieberman, 2002), haciendo énfasis principalmente en el proceso matemático y estadístico para realizar los pronósticos y no en dichos supuestos (normalidad y errores no autocorrelacionados).

Dentro de los propósitos de los métodos de series de tiempo se encuentra la captura de la variación aleatoria de la demanda a partir de su comportamiento histórico ya que en estos “se identifican los patrones fundamentales de la demanda que se combinan para producir el patrón histórico observado en la variable dependiente, después de lo cual se elabora un modelo capaz de producir dicho patrón” (Krajewski, Ritzman, & Malhotra, 2008). Es así como los modelos de regresión de series temporales, también llamados dinámicos, incorporan el movimiento histórico de los datos usando variables con el fin de capturar mejor la variabilidad de los datos, entre estas las estacionales, de tendencia, las del pasado endógenas y exógenas, entre otras.

El suavizamiento exponencial simple se basa en promediar o suavizar los valores pasados de una serie en forma exponencialmente decreciente, es decir que la observación más reciente tendrá mayor importancia que la siguiente observación más cercana, y así continuamente hasta encontrar que la observación más antigua será la que menor peso tenga.

Otros métodos como el de suavizamiento de Holt (1957), toma en cuenta las tendencias lineales locales en evolución dentro de una serie de tiempo y puede usarse para generar pronósticos. Esta suaviza el nivel y la pendiente de manera directa al usar diferentes constantes para cada una proporcionando estimados del nivel y de la pendiente que se adaptan a lo largo del tiempo conforme aparecen nuevas observaciones. Una ventaja de la técnica de Holt es su flexibilidad al seleccionar los coeficientes que controlan el nivel y la tendencia (Hanke & Reitsh, 1996).

La pregunta fundamental aquí es, si los métodos de suavizamiento exponencial simple y de Holt logran capturar adecuadamente toda la influencia que pueden tener los datos históricos o si es más eficiente un modelo dinámico, aplicando dichas metodologías para dos productos en una empresa de reciclaje que facilitó la información.

Metodología

Para el desarrollo de este trabajo se utilizaron los datos reales de ventas mensuales de 3 años (2009-2011) de la empresa Recuperar. La empresa en este momento se encarga del reciclaje de más de 20 productos.

Para la elección de los productos a pronosticar, se hizo un Pareto y se analizaron las gráficas de los productos más significativos según las utilidades.

Luego sobre dichos productos se aplicaron las siguientes técnicas de pronósticos:

- Suavización exponencial simple.
- Suavización exponencial: Método Holt.
- Regresión de series temporales o modelos dinámicos.

Cada modelo tiene unas pruebas de validación de residuales diferentes, así como la medición de la media del porcentaje de error absoluto (MAPE), que determina el grado de acierto que puede tener el modelo al pronosticar, en esa medida, el de menor porcentaje de MAPE será el mejor, dado que se validen sus supuestos.

Para la realización de estos métodos se utilizó principalmente el *software* estadístico Statgraphics, el lenguaje estadístico R y herramientas de *Microsoft Excel* 2010.

Modelo suavización exponencial simple

Forma general

$$\hat{Y}_t = \alpha Y_t + (1 - \alpha) \hat{Y}_{t-1}$$

Dónde:

$\hat{Y}_t = \hat{Y}_t$ = pronóstico para el siguiente período.

α = constante de suavización.

$Y_t Y_t$ = valor real de la serie en el período t.

$\hat{Y}_{t-1} \hat{Y}_{t-1}$ = pronóstico para el período t-1.

Este se estima según la minimización de la Suma de los cuadrados de los errores.

Modelo suavización exponencial lineal método de Holt

$$\text{Nivel: } l_t = \alpha Y_t + (1 - \alpha)(l_{t-1} + b_{t-1})$$

Tasa de crecimiento: $b_t = \gamma(l_t l_t - l_{t-1} l_{t-1}) + (1 - \gamma) b_{t-1}$

Donde:

α y γ son constantes de suavizamiento entre 0 y 1 y

$l_{t-1} l_{t-1}$ y $b_{t-1} b_{t-1}$ son estimaciones para el nivel y la tasa de crecimiento en $t-1$ respectivamente.

El pronóstico puntual es

$$\hat{y}_{t+\tau}(T) = l_t \hat{y}_{t+\tau}(T) = l_t + \tau b_t$$

Resultados y Análisis

La tabla 1 muestra la demanda de la espuma en los últimos tres años de la empresa Recuperar. De los cuales se procede a realizar la aplicación de los distintos métodos estadísticos explicados anteriormente.

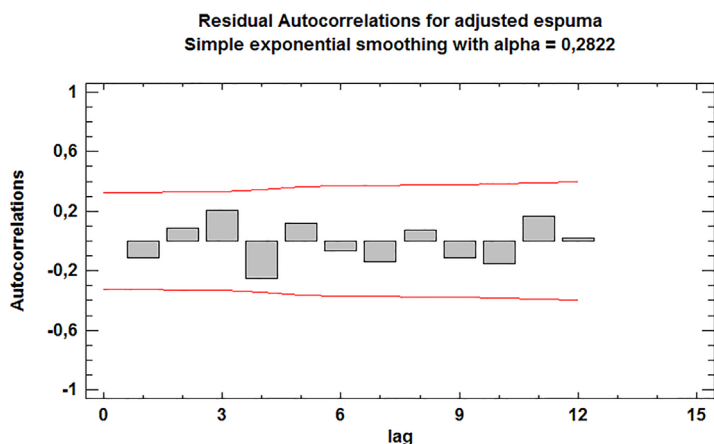
Tabla 1. Demanda de espuma.

	Meses/Año	2009	2010	2011
E S P U M A	1	5256	7392	4256
	2	5550	6299	6306
	3	5687	5660	4233
	4	4072	5740	4113
	5	4785	4461	4187
	6	3781	5955	3923
	7	6983	5777	4508
	8	5541	4007	3436
	9	6882	6089	4041
	10	8578	4662	4026
	11	5843	4058	4161
	12	5556	6052	4138

Modelo suavización exponencial simple

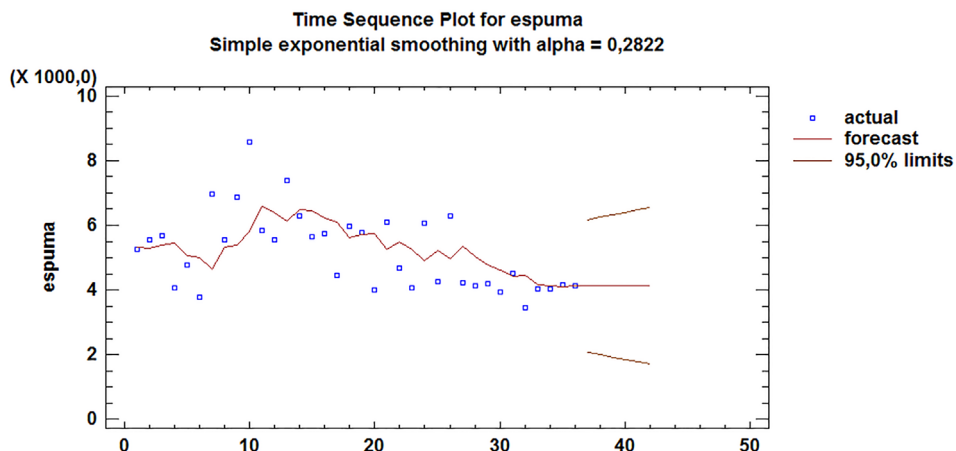
Al ingresar los datos de la tabla 1 en el *software* estadístico Statgraphics y realizar su modelación se pudo obtener que un α para este caso de 0,2822.

Figura1. Auto correlación en los residuales.



En la figura 1 se puede notar que no hay auto correlación en los residuales, puesto que ninguna barra se sale de las bandas. Comprobándolo así con la prueba de *Box-Pierce* en los primeros 12 rezagos se obtuvo un valor p igual a 0,683217 con lo que cumple la prueba de no auto correlación en los residuales. Se obtuvo un MAPE de 15,9241%; lo que muestra que hay una mala captura de la variabilidad de los datos, especialmente lo correspondiente a los picos de la gráfica.

Figura2. Tendencia de los datos según el método suavización simple.



En la figura 2 se puede notar la dispersión de los datos en el tiempo y la tendencia de la modelación con la metodología de suavización simple.

La tabla 2 nos muestra el pronóstico para los siguientes períodos desde enero hasta junio de 2012. Puede notarse esta tendencia en el último segmento de la figura 2.

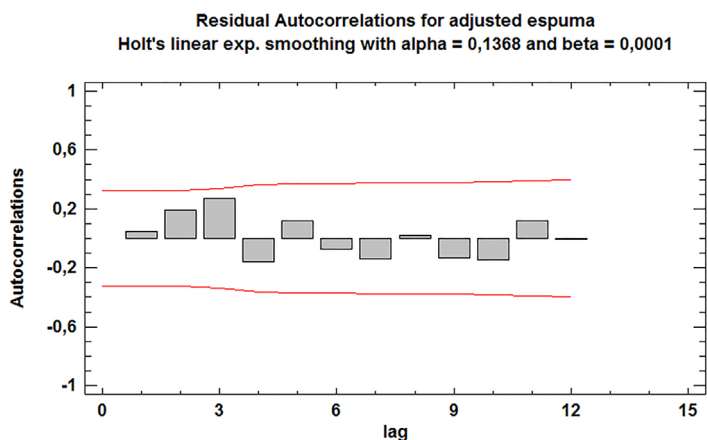
Tabla 2. Pronóstico usando suavización simple.

Mes	Pronóstico	Limit inf	Limit sup
Enero	4121,11	2072,84	6169,37
Febrero	4121,11	1992,84	6249,37
Marzo	4121,11	1915,74	6326,47
Abril	4121,11	1841,25	6400,96
Mayo	4121,11	1769,12	6473,09
Junio	4121,11	1699,13	6543,08

Modelo suavización exponencial lineal método de Holt

Al ingresar los datos se obtuvo los parámetros óptimos: $\alpha=0,1368$ y $\beta=0,0001$

Figura 3. Auto correlación en los residuales.

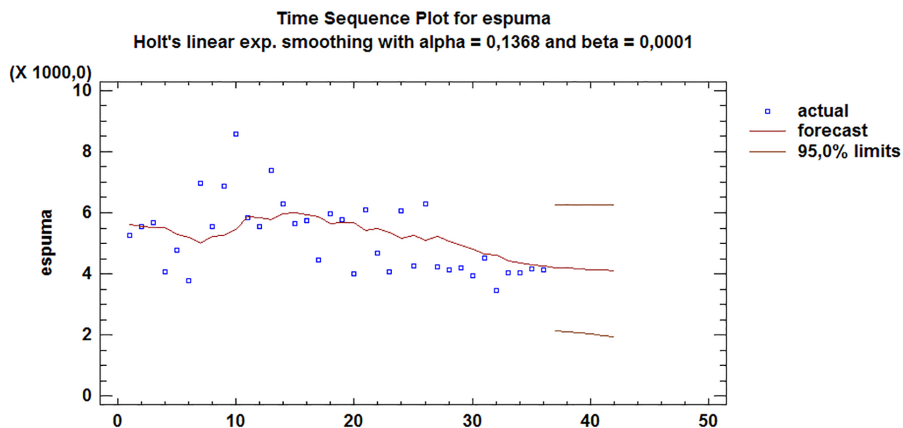


En la figura 3 se puede notar que no hay auto correlación en los residuales, puesto que ninguna barra se sale de las bandas.

En la prueba de *Box-Pierce* en los primeros 12 rezagos se obtuvo un valor p igual a 0,595251 con lo que cumple la prueba de no autocorrelación en los residuales.

El MAPE obtenido es de 15,9714%.

Figura 4. Tendencia de los datos según el método de suavización exponencial.



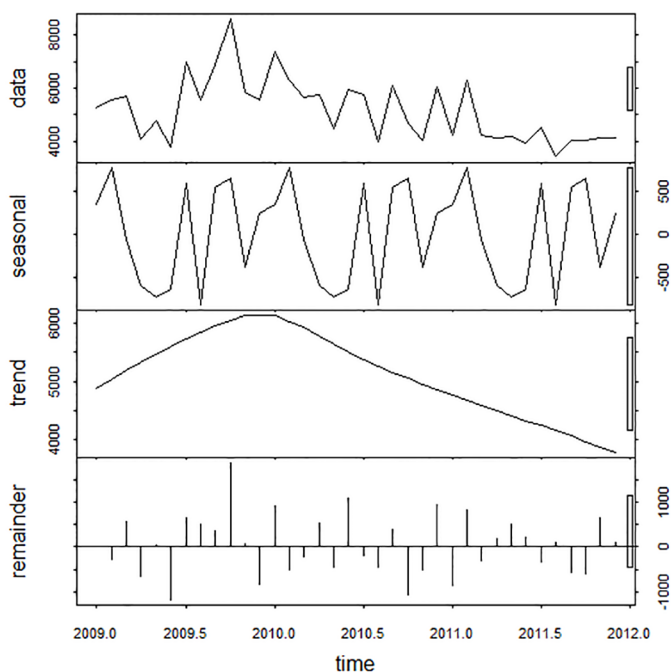
En la figura 4 se puede apreciar lo mismo que en la figura 2, pero, como es lógico en este caso para el método de *Holt*.

Tabla 3. Pronóstico suavización método de *Holt*.

Mes	Pronóstico	Limit inf	Limit sup
Enero	4206,15	2144,91	6267,4
Febrero	4183,03	2102,58	6263,47
Marzo	4159,9	2060,42	6259,38
Abril	4136,77	2018,44	6255,11
Mayo	4113,65	1976,61	6250,69
Junio	4090,52	1934,94	6246,1

Modelo dinámico

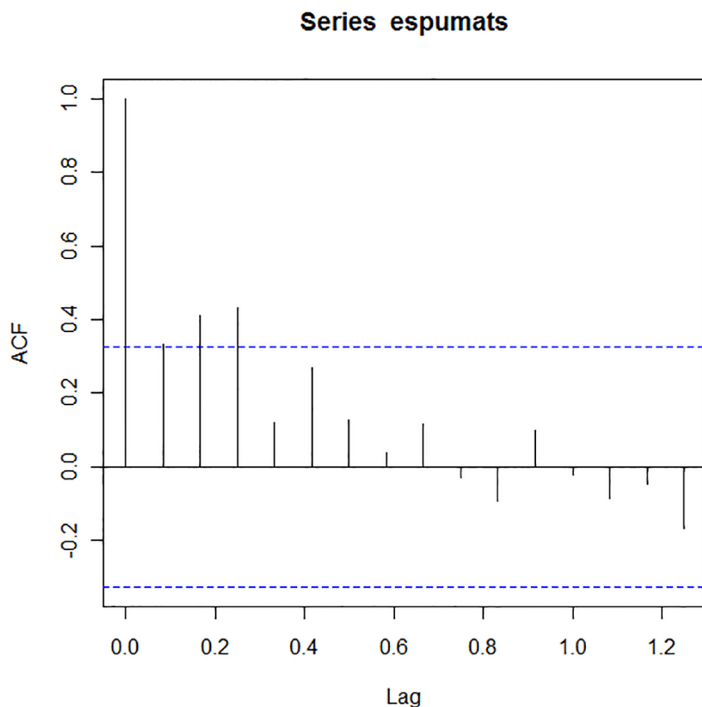
Figura 5. Descomposición de la serie en estacionalidad y tendencia.



En la figura 5, en *trend* se puede ver que la tendencia posee concavidad hacia abajo, lo que muestra que lo más seguro es que el tiempo de la serie sea de tipo cuadrático, por lo que se buscará ingresar al modelo una variable de tiempo de esta dimensión. En la gráfica de la serie (*data*) pueden notarse tres niveles distintos máximos que tiene la gráfica (ver picos). Por esto se buscará ingresar una variable nivel que logre captar ese cambio de altura de la gráfica.

La gráfica *remainder* muestra el residual de lo que el modelo estimó en la gráfica *seasonal* y *trend*. Se observa que hay partes donde la diferencia no es muy grande (donde la línea es corta) pero hay otras partes donde hay mucha diferencia como en el dato que se encuentra entre 2009.5 y 2010.0, esto significa que no se logró capturar bien la estacionalidad y la tendencia en este punto.

Figura 6. Gráfica de autocorrelación de la serie.



Al observar la figura 6 puede notarse que los rezagos 1, 2 y 3 son significativamente diferentes de cero y por esto se afirma que hay autocorrelación. Más adelante en la tabla 4 se muestra la prueba de *Box-Pierce* la cual niega o confirma lo anterior numéricamente.

Los tres valores de autocorrelación que se salen de las bandas se deben tener en cuenta a la hora de estimar un modelo dinámico: Ya que Los valores están mostrando dependencia de orden 1, 2 y 3 en la serie y por ello se incorporarán en el modelo.

Modelo dinámico

El siguiente modelo se logró después de estimar una serie de alternativas posibles basadas en el análisis de las gráficas. Al final, se escogió el modelo que generó el mejor MAPE y cumplió los supuestos en los residuales.

A continuación se escribe el modelo estimado con intercepto, tendencia del tiempo al cuadrado, variable nivel y variable indicadora mes.

Los coeficientes que acompañan a cada variable se muestran en la ecuación a continuación:

$$y = 6,575e+05 - 0,1624 t^2 + 0,216 Z_{t-2} + 0,3441 Z_{t-3} + 196 \text{ nivel} + 1212 \text{ enero} + 2135 \text{ febrero} + 535,1 \text{ marzo} - 67,64 \text{ mayo} + 411,2 \text{ junio} + 1894 \text{ julio} + 560,5 \text{ agosto} + 1672 \text{ septiembre} + 1705 \text{ octubre} + 893,2 \text{ noviembre} + 1029 \text{ diciembre}$$

Pruebas

Tabla 4. Pruebas en los residuales.

Pruebas	Vp
<i>JarqueBera Test</i>	0.8465
<i>Shapiro-Wilk</i>	0.7143
<i>Box-Pierce test</i>	0.5882 Lag= 24
	0.2437 Lag= 10

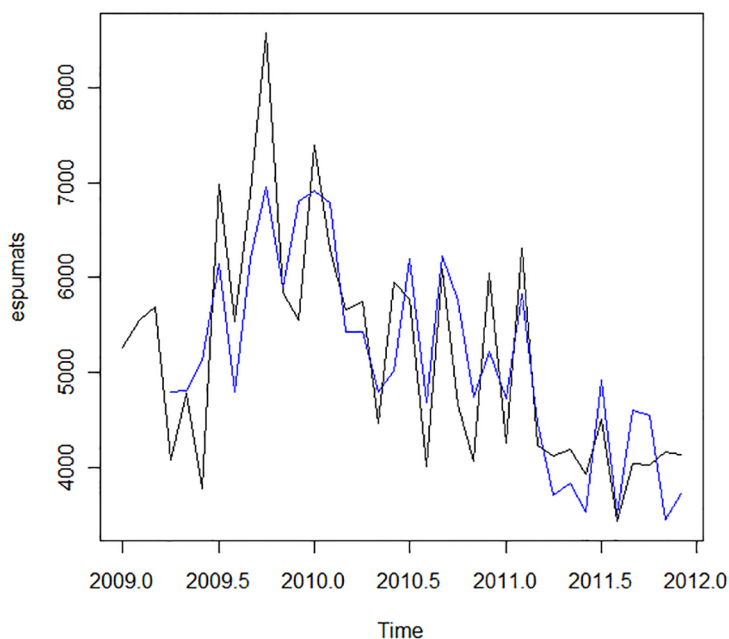
En la tabla 4 se aprecia la validación de los supuestos en los residuales, donde hay cumplimiento de todos a un nivel del 5% de significancia. Mostrando que existe normalidad y autocorrelación en los residuales. La prueba de *Box-Pierce* se hace para 24 y 10 rezagos respectivamente.

Por lo anterior, el modelo es confiable estadísticamente. Adicionalmente, el MAPE del modelo es 11.50487%

Tabla 5. Pronóstico modelo dinámico.

Año	Mes	Pronóstico espuma
2012	Enero	4165,53961
2012	Febrero	5075,56324
2012	Marzo	3419,23598
2012	Abril	3035,71369
2012	Mayo	2868,9801
2012	Junio	2640,5688
2012	Julio	3900,91378

Figura 6. Gráfica de la serie superpuesta a gráfica del modelo.



La figura 6 muestra en azul la fluctuación estimada del modelo. Puede notarse cómo en unas partes de los picos hay una diferencia entre el valor estimado y el real, algo que se verá reflejado en el MAPE.

Conclusiones

Entre los dos modelos de suavización exponencial empleados, el simple obtuvo mejor MAPE, aunque la diferencia entre ambos fue de 0.047 %. Lo cual indica que ambos arrojan resultados muy similares. Sin embargo, al comparar los tres métodos utilizados para realizar los pronósticos y las pruebas de validación requeridas en los residuales, se puede apreciar que el mejor MAPE se obtuvo con el modelo dinámico. Esto puede deberse a que los modelos dinámicos permiten incorporar variables del pasado y diferentes tendencias y estacionalidad del tiempo logrando, muchas veces, capturar adecuadamente la variabilidad de la serie.

Lo que se indagó para el modelo dinámico en la figura 5 y 6 se cumplió y efectivamente, modelando dentro de estas posibilidades se obtuvo el mejor modelo posible.

La prueba de *Bax-Pierce* para este mismo modelo afirma que no hay autocorrelación en los residuales y puesto que, pasa la prueba con una buena cantidad de rezagos, se dice que el modelo es bueno para pronosticar a largo plazo.

Referencias

1. Cohen, R. y otros (1986). *Forecasting for Inventory Control: An Example of When "Simple" Means "Better"*. *Interfaces*, Vol. 16, No. 6 (Nov. - Dec.), pp. 95-99.
2. Gutiérrez, V. Vidal, C.J. (2008). Modelos de Gestión de Inventarios en Cadenas de Abastecimiento: Revisión de la Literatura. *Inventory Management Models in Supply*. Rev. Fac. Ing. Univ. Antioquia N.º 43. Marzo. pp. 134-149.
3. Gurnani, H. Tang, C. (1999). *Optimal Ordering Decisions with Uncertain Cost and Demand Forecast Updating*. *Management Science*. Vol. 45, No. 10 (Oct.), pp. 1456-1462.
4. Hanke, J. E., & Reitsch, A. G. (1996). *Pronósticos en los negocios*. México: Prentice Hall. 552 p.
5. Hillier, F. S., & Lieberman, G. J. (2002). *Investigación de operaciones*. México: Mc Graw-Hill. 950 p.

6. Krajewski, L., Ritzman, L., & Malhotra, M. (2008). *Administración de operaciones* (8va ed.). México: Pearson educación. 1100 p.
7. Sethi S. P, et al. (2003). *Inventory Models with Fixed Costs, Forecast Updates, and Two Delivery Modes*. *Operations Research*, Vol. 51, No. 2 (Mar. - Apr.), pp. 321-328.
8. Vidal, V. G. (marzo 2008). Modelos de gestión de inventarios en cadenas de abastecimiento: Revisión literaria. Facultad de Ingeniería Universidad de Antioquia (043), 137. Pp. 134-149