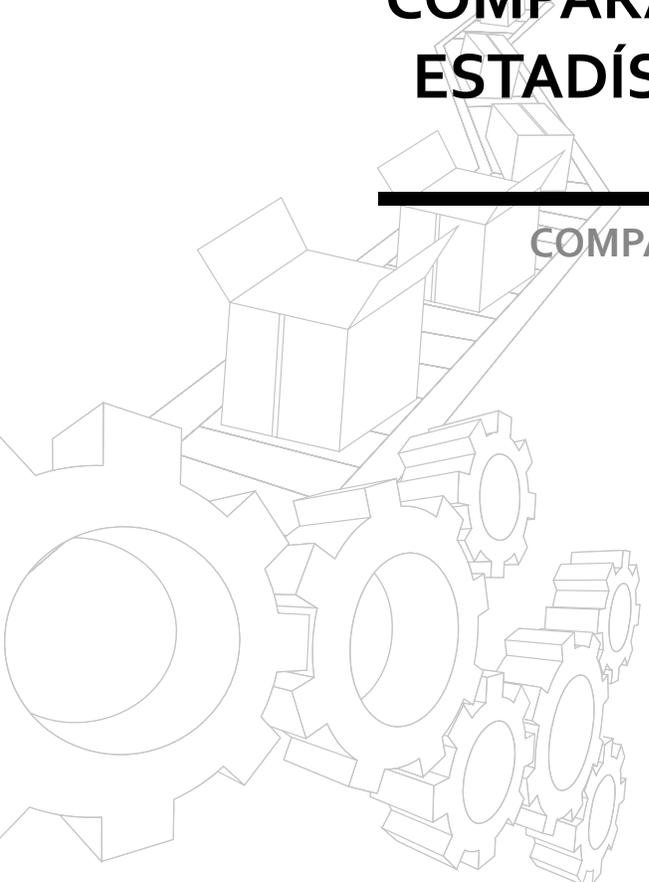


COMPARACIÓN DE METODOLOGÍAS ESTADÍSTICAS EN EL PRONÓSTICO DE LA DEMANDA

COMPARISON OF STATISTICS METHODOLOGIES
ON DEMAND FORECAST



 **Sebastián Ramírez Agudelo**

*Estudiante de la Facultad
de Ingeniería Industrial
Universidad Pontificia Bolivariana
sebastian.ramirezag@alfa.upb.edu.co*

 **Juan Esteban Cardona Ramírez**

*Estudiante de la Facultad de Ingeniería Industrial
Universidad Pontificia Bolivariana
juanesteban.cardona@alfa.upb.edu.co*

 **Daniela González Lotero**

*Estudiante de la Facultad
de Ingeniería Industrial
Universidad Pontificia Bolivariana
daniela.gonzalezlo@alfa.upb.edu.co*

 **Marisol Valencia Cárdenas**

*Ms en Estadística,
Universidad Nacional de Colombia Sede Medellín
solmarival@gmail.com, mvalencia@unal.edu.co*



Empresas de manufactura o de distribución, requieren programar pedidos con anticipación y precisión para una adecuada planeación de operaciones; por ello, un pronóstico de demanda estima las cantidades a ser vendidas, facilitando la planeación de 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, mostrando mejores resultados para el modelo dinámico.

PALABRAS CLAVE

Pronósticos, modelo dinámico, suavización exponencial

RESUMEN

ABSTRACT

Both manufacturing and distribution companies require the scheduling orders in advance and with precision for proper operation planning; therefore demand forecasting estimates quantities to be sold, which facilitates inventory planning and necessary logistic for supply. In this work three statistical models are applied in order to forecast and compare efficiency between them in a recycling company, showing better results for the dynamic model.

KEYWORDS

Forecasts, dynamic model, exponential smoothing.



1. 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 [1].

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, algunos usando distribuciones de probabilidad [2], [3], [4], otros, modelos de pronósticos como los de suavización exponencial, media móvil, ARIMA, entre otros, como se ha visto en la recopilación de [5], reflejando la importancia de este tratamiento de la demanda estimada para mejorar la planeación de inventarios en una empresa manufacturera. Sin embargo, algunos de estos trabajos evidencian una falta de validación necesaria de supuestos en los errores [6], haciendo énfasis principalmente en el proceso matemático, pero generando estimaciones poco confiables.

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"* [7]. 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, toma en cuenta las tendencias lineales locales en evolución dentro de una serie de tiempo y puede usarse para generar pronósticos. Ésta 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 [8]

La pregunta fundamental aquí es si los métodos de suavizamiento exponencial simple, Holt o un modelo dinámico logran capturar adecuadamente toda la influencia que pueden tener los datos históricos, comparando las estimaciones de dichas metodologías en una serie de datos de un producto de una empresa de reciclaje de la ciudad de Medellín.

2. 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 del producto a pronosticar, se hizo un Pareto con respecto a las utilidades; de estos el más representativo fue la Espuma.



Luego sobre la serie temporal del producto elegido se aplicaron las siguientes técnicas de pronósticos:

- Suavización exponencial simple
- Suavización exponencial: Método Holt
- Modelos dinámicos lineales.

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

La estimación parte de la optimización de la suma de los cuadrados de los errores SSE, encontrando el parámetro que la minimiza, lo cual permite ajustar y pronosticar valores de ventas.

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

Donde:

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

α = constante de suavización.

Y_t = valor real de la serie en el periodo t.

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

Este se estima según la minimización de la Suma de los Cuadrados de los Errores.

Modelo suavización exponencial lineal método de Holt

El modelo incorpora dos parámetros: de nivel y de crecimiento, que facilita tomar en cuenta cam-

bios en tendencia de la serie de tiempo. También se basa en la minimización de la suma SSE.

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

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

Donde:

α y γ son constantes de suavizamiento entre 0 y 1 y l_{t-1} y b_{t-1} son estimaciones para el nivel y la tasa de crecimiento en t-1 respectivamente.

$$\text{El pronóstico puntual es: } \hat{y}_{t+\tau}(T) = l_t + \tau b_t$$

Modelo dinámico lineal

El modelo dinámico presentado aquí se basa en la estimación de máxima verosimilitud para una ecuación de regresión múltiple, donde se incorporan variables endógenas (de la misma serie) y exógenas (independientes a esta). Su forma matemática es [g]:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 t + \hat{\beta}_1 Y_{t-1}$$

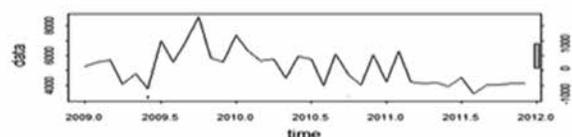
En la ecuación se muestra como ejemplo dos variables explicativas: t que refleja tendencia lineal y un término autorregresivo de orden 1 ().

3. Resultados y Análisis

Se utilizaron los datos de demanda mensual de la espuma en los últimos 3 años de la empresa Recuperar. En esta sección se muestra la aplicación de los 3 modelos explicados a esta serie de datos.

Comportamiento de los datos

FIGURA 1. SERIE DE TIEMPO

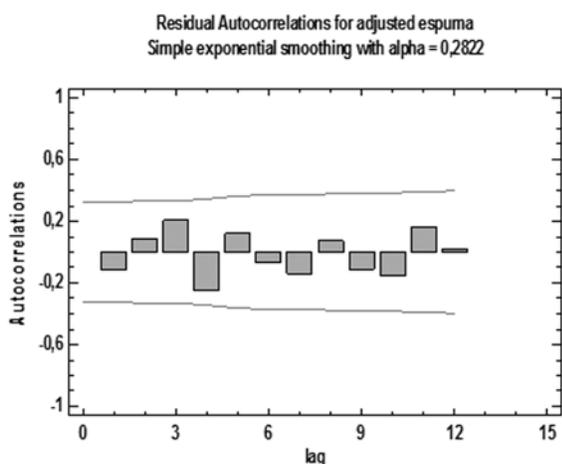


La figura 1 muestra la serie de tiempo, con tendencia levemente descendiente, pero no lineal, y con algunos cambios drásticos en la serie, que podría generar problemas de estimación. Pueden apreciarse alrededor de tres niveles distintos.

Modelo suavización exponencial simple

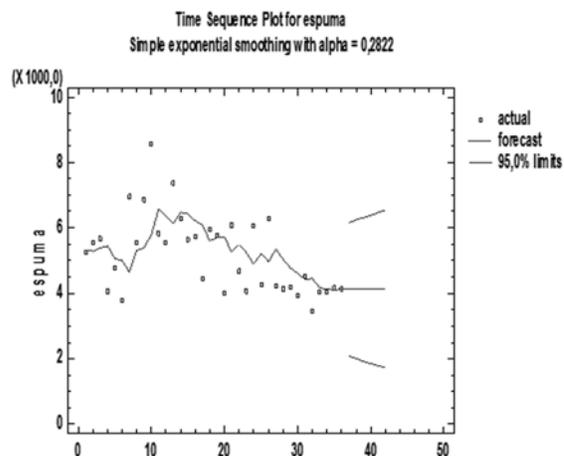
Al ingresar los datos de la serie temporal en el software estadístico Statgraphics, se realizó la optimización del parámetro, obteniendo un α de 0,2822.

FIGURA 2. AUTO CORRELACIÓN EN LOS RESIDUALES



En la figura 2 se puede notar que no hay auto correlación en los residuales, puesto que ninguna barra se sale de las bandas lo que se comprueba con la prueba de Box-Pierce, que con un valor p de 0,683, mayor del 5%, indica independencia en los residuales. Para este modelo se obtuvo un MAPE de 15,92%, mostrando ajuste regular del modelo, especialmente se refleja en los cambios notables de la serie de tiempo.

FIGURA 3. TENDENCIA DE LOS DATOS SEGÚN EL MÉTODO SUAVIZACIÓN SIMPLE



En la figura 3 se puede notar la dispersión de los datos en el tiempo y la tendencia del ajuste (en rojo) con el modelo de suavización simple.

La tabla 1 muestra el pronóstico para los siguientes periodos desde enero hasta junio de 2012. Puede notarse esta tendencia en el último segmento de la figura 2.

TABLA 1. PRONÓSTICO USANDO SUAVIZACIÓN SIMPLE

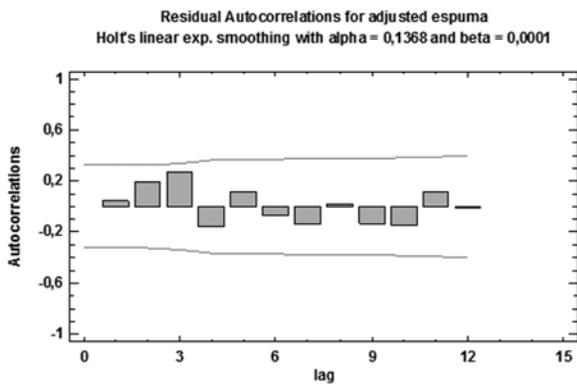
MES	PRONÓSTICO	LIMIT INF	LIMIT SUP
Enero	4121,11	2072,84	6169,37
Febrero	4121,11	1992,84	6249,37
Marzo	4121,11	1915,74	6326,47
Abril	4121,11	1841,25	6400,96
Mayo	4121,11	1769,12	6473,09
Junio	4121,11	1699,13	6543,08

Modelo suavización exponencial lineal método de Holt

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

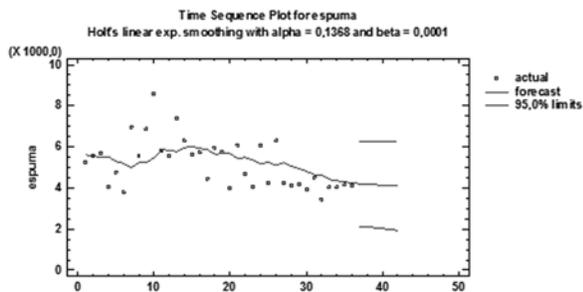


FIGURA 4. AUTO CORRELACIÓN EN LOS RESIDUALES



La figura 4 muestra que no hay auto correlación en los residuales, ya que dichos valores no se salen de las bandas de confianza, lo cual se comprueba con Box-Pierce, que con un valor p de 0,595 mayor de 0.05, indica el cumplimiento de no auto-correlación. El porcentaje de error relativo (MAPE) obtenido es de 15,9714%.

FIGURA 5. TENDENCIA DE LOS DATOS SEGÚN EL MÉTODO DE SUAVIZACIÓN EXPONENCIAL



La figura 5 muestra con rojo la línea de ajuste y los puntos azules los datos observados, indicando muchos valores observados alejados de esta, de forma similar al método de Holt.

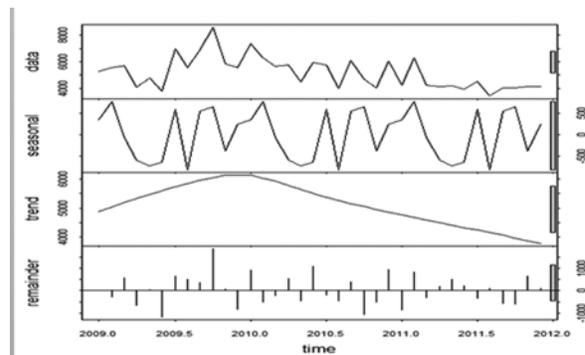
TABLA 2. PRONOSTICO SUAVIZACIÓN MÉTODO DE HOLT

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

Descripción preliminar de la serie

FIGURA 6. DESCOMPOSICIÓN DE LA SERIE EN ESTACIONALIDAD Y TENDENCIA



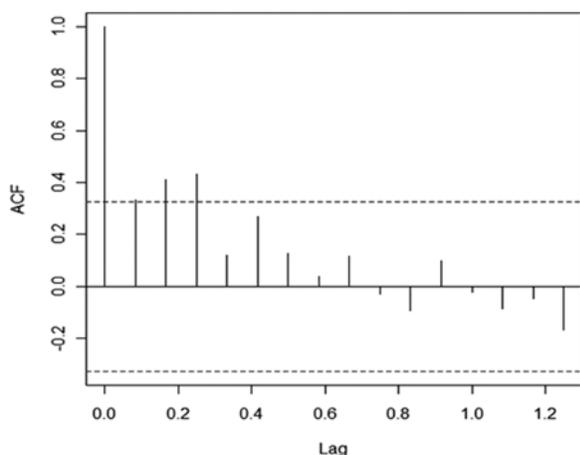
La figura 6 muestra una descomposición con suavización de loess, que sirve para describir el comportamiento de la tendencia y la estacionalidad de manera preliminar a la estimación de modelos dinámicos. La tendencia (trend) se aprecia evidentemente con concavidad hacia abajo lo que muestra un comportamiento de la serie vs el tiempo cuadrático, luego esta es una variable



explicativa a considerar para este modelo. 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 la estimación suavizada; este se muestra pequeño en algunos periodos, pero en otros hay mucha diferencia, por ejemplo, los datos entre 2009.5 y 2010.0, esto significa que en la descomposición no se logró capturar bien la estacionalidad y la tendencia en dichos periodos.

FIGURA 7. GRÁFICA DE AUTOCORRELACIÓN DE LA SERIE



Al observar la figura 7 puede notarse que las autocorrelaciones en los rezagos 1, 2 y 3 son significativamente diferentes de cero y por esto se afirma que hay autocorrelación.

Dichos valores de autocorrelación que se salen de las bandas muestran una dependencia en la serie de orden 1, 2 y 3, por ello se incorporarán en el modelo las variables autorregresivas respectivas.

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 (menor error) 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:

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

Dónde:

t^2 : tiempo cuadrático, indicando la tendencia de grado 2.

Z_{t-2}, Z_{t-3} : variables autorregresivas en la serie de orden 2 y 3 (significativas).

Las demás variables independientes o explicativas son indicadores o dummy.

Pruebas de validación

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 (valores p mayores). Mostrando que existe normalidad y no hay autocorrelación en los residuales, es decir hay independencia en estos.



TABLA 4. PRUEBAS EN LOS RESIDUALES

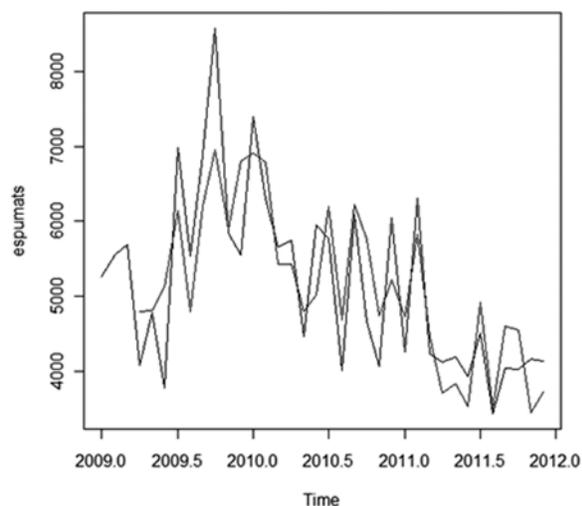
PRUEBAS	VP
JarqueBera Test	0.8465
Shapiro-Wilk	0.7143
Box-Pierce test	0.5882 Lag= 24 0.2437 Lag= 10

Por lo anterior el modelo es confiable estadísticamente. Adicionalmente el MAPE del modelo es 11.5%

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 8. GRÁFICA DE LA SERIE SUPERPUESTA A GRÁFICA DEL MODELO



La figura 8 muestra en azul la fluctuación estimada por el modelo. Puede notarse como en unas partes de los picos hay una diferencia entre el valor estimado y el real, algo que se ve reflejado en el error de pronóstico.

Conclusiones

Entre los dos modelos de suavización exponencial empleados, fue el simple el que obtuvo mejor resultado por tener menor error de pronóstico, aunque la diferencia entre ambos fue de 0.047 % en el indicador MAPE, lo cual muestra resultados muy similares. Sin embargo, al comparar los tres métodos utilizados para realizar los pronósticos el que muestra un mejor ajuste al comportamiento de la serie es el modelo dinámico, que además cumple con la validación de supuestos en los residuales (normalidad, estacionariedad). Esto puede deberse a que dichos modelos permiten incorporar variables endógenas de la serie, así como diferentes tipos de tendencias y estacionalidad del tiempo logrando, muchas veces, capturar adecuadamente la variabilidad de la serie.

Del análisis descriptivo de la serie realizado en las figuras 5 y 6, se extrajo información que aportó de manera significativa en la estimación del modelo dinámico, con mejores resultados que los otros, generando confianza en el método de análisis preliminar empleado para la serie de tiempo. El modelo dinámico estimado facilita la descripción del comportamiento de la venta de la espuma, lo cual ayuda a que la empresa conozca aquellos periodos de mayor venta y genere alternativas de búsqueda de dicho material, generando mayor ingreso en la compañía.

Referencias

- [1] Vidal, V. G.. Modelos de gestión de inventarios en cadenas de abastecimiento: Revisión literaria. Facultad de Ingeniería Universidad de Antioquia (043), 137. 2008.
- [2] Cohen, R. y otros. Forecasting for Inventory Control: An Example of When "Simple" Means "Better". Interfaces, Vol. 16, No. 6 (Nov. - Dec.), pp. 95-99. 1986.
- [3] Gurnani, H. Tang, C. Optimal Ordering Decisions with Uncertain Cost and Demand Forecast Updating. Management Science. Vol. 45, No. 10 (Oct.), pp. 1456-1462. 1999.



- [4] Sethi S. P, et al. Inventory Models with Fixed Costs, Forecast Updates, and Two Delivery Modes. *Operations Research*, Vol. 51, No. 2 (Mar. - Apr.), pp. 321-328. 2003.
- [5] Gutiérrez, V. Vidal, C.J. 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. 2008.
- [6] Hillier, F. S., & Lieberman, G. J. (2002). *Investigación de operaciones*. México: Mc Graw-Hill.
- [7] Krajewski, L., Ritzman, L., & Malhotra, M. (2008). *Administración de operaciones* (8va ed.). México: Pearson educación.
- [8] Hanke, J. E., & Reitsch, A. G. (1996). *Pronósticos en los negocios*. México: Prentice Hall.
- [9] Bowerman, B. L., O'Connell, R. T., & Koehle, A. B. (2007). *Pronóstico, series de tiempo y regresión* (4ta ed.). Monterrey: Cengage Learning.

