

**METODOLOGÍA PARA OPTIMIZAR INVENTARIOS DE PRODUCTO
TERMINADO EN UNA EMPRESA, BASADO EN ESTIMACIÓN DE DEMANDA Y
MINIMIZACIÓN DE COSTOS**

**ANDREA BERMÚDEZ ESCOBAR
SEBASTIÁN LONDOÑO OLARTE**

**UNIVERSIDAD PONTIFICIA BOLIVARIANA
ESCUELA DE INGENIERÍA
FACULTAD DE INGENIERÍA INDUSTRIAL
MEDELLÍN**

2013

**METODOLOGÍA PARA OPTIMIZAR INVENTARIOS DE PRODUCTO
TERMINADO EN UNA EMPRESA, BASADO EN ESTIMACIÓN DE DEMANDA Y
MINIMIZACIÓN DE COSTOS**

ANDREA BERMÚDEZ ESCOBAR

SEBASTIÁN LONDOÑO OLARTE

Trabajo de grado para optar al título de Ingeniero Industrial

Directora

MARISOL VALENCIA CÁRDENAS

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA DE INGENIERÍA

FACULTAD DE INGENIERÍA INDUSTRIAL

MEDELLÍN

2013

Nota de aceptación

Firma:

Nombre:

Presidente del jurado:

Firma:

Nombre:

Jurado:

Firma:

Nombre:

Jurado:

AGRADECIMIENTOS

Es el momento de hacer un alto en el camino y agradecerle a aquellos que hicieron parte de esta experiencia, a Dios como principal artífice de cada uno de nuestros triunfos, a nuestros padres por su incondicional apoyo, la educación y el amor que nos han brindado.

Agradecemos el gran apoyo y la colaboración de nuestra directora: Marisol Cárdenas en todo este proyecto, la disposición de siempre compartir su conocimiento.

Agradecimientos a la empresa Auralac por abrirnos sus puertas y permitirnos materializar este proyecto.

TABLA DE CONTENIDO

RESUMEN.....	10
INTRODUCCIÓN.....	11
1. OBJETIVOS	13
1.1 Objetivo General.....	13
1.2 Objetivos específicos.....	13
2. ANTECEDENTES.....	14
3. MARCO TEÓRICO	17
3.1 Estadística descriptiva	17
3.2 Series de tiempo.....	17
3.3 Pronósticos.....	18
3.4 Modelos de pronóstico.....	19
3.4.1 Modelos de suavización	19
3.4.2 Modelos de regresión lineal	20
3.4.3 Modelos Box Jenkins- ARIMA	24
3.4.4 Red Neuronal Artificial (RNA)	26
3.5 Pruebas de validación	28
3.6 Criterios de evaluación	30
3.7 Modelos de inventarios.....	31
4. METODOLOGÍA.....	32
4.1. Estimación de modelos de pronóstico	33
4.2. Modelo de optimización basado en revisión periódica.....	38
5. RESULTADOS	39
5.1 Análisis exploratorio.....	40
5.1.1 Datos extremos por producto	40
5.1.2 Análisis descriptivo de las series	42
5.2 Modelos de pronósticos	51
5.2.1 Modelos de suavizamiento exponencial	52
5.2.2 Modelos dinámicos.....	62
5.2.3. Modelos ARIMA	73

5.3. Modelo de optimización de inventarios	91
CONCLUSIONES	97
RECOMENDACIONES	98

INDICE DE TABLAS

Tabla 1. Identificación tentativa para los Modelos Box Jenkins.....	25
Tabla 2. Demanda real vs pronóstico con suavización exponencial simple, referencia 1200	53
Tabla 3. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1200	54
Tabla 4. Demanda real vs pronóstico con suavización exponencial simple, referencia 1400	56
Tabla 5. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1400	58
Tabla 6. Demanda real vs pronóstico con suavización exponencial simple, referencia 4100	60
Tabla 7. Demanda real vs pronóstico con suavización exponencial Holt, referencia 4100	61
Tabla 8. Asignación de niveles según valor del percentil	63
Tabla 9. ANOVA Tipo III, referencia 1400	64
Tabla 10. Demanda real vs pronóstico con modelo dinámico tentativo, referencia 4100 ..	65
Tabla 11. Niveles estimados para las referencias según el día a pronosticar.....	66
Tabla 12. ANOVA Tipo III final para la referencia 1400	66
Tabla 13. Demanda real vs pronóstico con modelo dinámico final, referencia 4100	67
Tabla 14. ANOVA Tipo III, referencia 1400.....	69
Tabla 15. Demanda real vs pronóstico con modelo dinámico, referencia 1200.....	70
Tabla 16. ANOVA Tipo III, referencia 4100.....	71
Tabla 17. Demanda real vs pronóstico con modelo dinámico, referencia 1400.....	72
Tabla 18. Coeficientes del modelo Box-Jenkins, referencia 1200.....	75
Tabla 19. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1200	75
Tabla 20. Coeficientes del modelo Box-Jenkins para la referencia 1400	78
Tabla 21. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1400	79
Tabla 22. Coeficientes del modelo Box-Jenkins para la referencia 4100	81
Tabla 23. Demanda real vs pronóstico con modelo Box Jenkins, referencia 4100	82
Tabla 24. Demanda real vs pronóstico mediante RNA, referencia 1200	84
Tabla 25. Demanda real vs pronóstico mediante RNA, referencia 1400	87
Tabla 26. Demanda real vs pronóstico mediante RNA, referencia 4100	89
Tabla 27. Resumen de valores de MSE, MAPE de pronóstico y cumplimiento de supuestos de validación	90
Tabla 28. Límite superior e inferior de demanda por producto y día	91
Tabla 29. Costos asociados a los inventarios y utilidad por referencia	91
Tabla 30. Valores de la solución de pedidos e inventarios, referencia 4100.....	92
Tabla 31. Valores de la solución de pedidos e inventarios, referencia 4100.....	93
Tabla 32. Valores de la solución de pedidos e inventarios, referencia 4100.....	93
Tabla 33. Valores de costo reducido	94
Tabla 34. Excedentes y precios sombra.....	94

ÍNDICE DE FIGURAS

Figura 1 Estructura de una RNA Elaboración propia.....	27
Figura 2. Metodología para entrenar y simular una red Neuronal. Elaboración propia	34
Figura 3. Asignación de variables de entrada y de salida a MATLAB	35
Figura 4. Importación de entradas y salidas para la creación de la RNA en MATLAB	35
Figura 5. Creación de la RNA en MATLAB	36
Figura 6. Entrenamiento de la RNA en MATLAB	37
Figura 7. Simulación de la RNA en MATLAB	37
Figura 8. Caja y bigotes para la serie de datos de la referencia 1200.....	40
Figura 9. Caja y bigotes para la serie de datos de la referencia 1400.....	41
Figura 10. Caja y bigotes para la serie de datos de la referencia 4100	42
Figura 11. Salidas por facturación de la referencia 1200 en los últimos dos meses.....	43
Figura 12. Comportamiento de la tendencia en la referencia 1200	43
Figura 13. Estacionalidad en la referencia 1200	43
Figura 14. Facturación promedio según el día de la semana	44
Figura 15. Función de autocorrelación, referencia 1200.....	45
Figura 16. Función de autocorrelación parcial, referencia 1200.....	45
Figura 17. Salidas por facturación de la referencia 1400 en los últimos dos meses.....	46
Figura 18. Facturación promedio según el día de la semana, referencia 1400	47
Figura 19. Comportamiento de la tendencia en la referencia 1200	47
Figura 20. Función de autocorrelación, referencia 1400.....	48
Figura 21. Función de autocorrelación parcial, referencia 1400.....	48
Figura 22. Salidas por facturación de la referencia 4100 en los últimos dos meses.....	49
Figura 23. Facturación promedio según el día de la semana, referencia 4100	50
Figura 24. Comportamiento de la tendencia, referencia 1400	50
Figura 25. Función de autocorrelación, referencia 1400.....	51
Figura 26. Función de autocorrelación parcial, referencia 1400.....	51
Figura 27. Demanda real vs pronóstico con suavización exponencial simple, referencia 1200.....	52
Figura 28. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1200	54
Figura 29. Demanda real vs pronóstico con suavización exponencial simple, referencia 1400.....	56
Figura 30. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1400	58
Figura 31. Demanda real vs pronóstico con suavización exponencial simple, referencia 4100.....	60
Figura 32. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1400	62
Figura 33. Demanda real vs pronóstico con modelo dinámico, referencia 1400	68

Figura 34. Demanda real vs pronóstico con modelo dinámico, referencia 1200	70
Figura 35. Demanda real vs pronóstico con modelo dinámico, referencia 4100	73
Figura 36. Función de autocorrelación diferenciada, referencia 1200.....	74
Figura 37. Función de autocorrelación parcial diferenciada, referencia 1200	74
Figura 38. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1200.....	76
Figura 39. Función de autocorrelación diferenciada, referencia 1400.....	77
Figura 40. Función de autocorrelación parcial diferenciada, referencia 1400	77
Figura 41. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1400.....	79
Figura 42. Función de autocorrelación diferenciada, referencia 4100.....	80
Figura 43. Función de autocorrelación parcial diferenciada, referencia 4100.....	80
Figura 44. Demanda real vs pronóstico con modelo Box Jenkins, referencia 4100.....	82
Figura 45. Estructura de la RNA, referencia 1200	83
Figura 46. Rendimiento del modelo de RNA, referencia 1200.....	84
Figura 47. Demanda real vs pronóstico mediante RNA, referencia 1200	85
Figura 48. Estructura de la RNA, referencia 1400	86
Figura 49. Rendimiento del modelo de RNA, referencia 1400.....	86
Figura 50. Demanda real vs pronóstico mediante RNA, referencia 1400	87
Figura 51. Estructura de la RNA, referencia 4100	88
Figura 52. Rendimiento del modelo de RNA, referencia 4100.....	88
Figura 53. Demanda real vs pronóstico mediante RNA, referencia 4100	89
Figura 54. . Solución del sistema de inventarios en LINGO.....	92

RESUMEN

El manejo adecuado de los inventarios es uno de los aspectos determinantes para la rentabilidad de una empresa manufacturera, en especial, del sector lácteo, con tantas variaciones en su demanda y el corto ciclo de vida de sus productos. Este trabajo parte de estimar modelos de pronóstico de demanda para tres productos elaborados por una empresa del sector lácteo de Antioquia, comparando entre algunos métodos clásicos y redes neuronales, con el fin de encontrar el método más eficiente y acertado para efectuar predicciones de ventas. Éstas son entradas para proponer un modelo de optimización de inventarios, que permita determinar una política adecuada para su manejo.

Se describen y se utilizan los modelos de pronóstico: Suavización exponencial simple, Suavización exponencial Holt, modelos dinámicos y de Box-Jenkins, por medio de los programas estadísticos *Statgraphics* y *R*. Se muestra el proceso como se desarrollan cada uno de ellos, su respectivo análisis y se determina cual es el modelo clásico más eficiente.

Seguido a esto, se formula un modelo de red neuronal artificial (RNAs), para cada una de las series, usando el software *MATLAB* y se compara con el modelo clásico elegido. Se detalla el proceso de estimación de las RNA y se selecciona el modelo de pronóstico final.

Por último se realiza el modelo de optimización de inventarios, en el software *LINGO*, usando los pronósticos obtenidos de los modelos de RNAs, que resultaron los más eficientes. Se propone finalmente un esquema de pedidos y almacenamiento de inventarios a costo óptimo.

INTRODUCCIÓN

Los inventarios son considerados el amortiguador de los procesos de producción y distribución; gracias a ellos se pueden disponer de insumos o de productos terminados en las cantidades necesarias en el momento oportuno (Vidal & Gutiérrez, 2008); una inadecuada administración de ellos conllevaría a la inmovilización de capital en mercancías, reflejado en la liquidez de las compañías, además de sobrecostos de operación.

Una inadecuada planeación puede llevar a excesos o agotamientos de producto para el abastecimiento. Tener inventarios al alza puede generar sobrecostos, pérdida de productos, baja en precios, y por lo tanto, una reducción en el margen de utilidad; por otro lado, tener inventarios a la baja, por limitaciones de almacén o presupuesto, causaría pérdidas en ventas, y en el mercado a largo plazo.

Ambos aspectos son conflictivos, reduciendo el inventario se minimiza la inversión, pero se corre el riesgo de no poder satisfacer la demanda de las operaciones de la empresa. Si se tienen grandes cantidades de inventario, se disminuye la posibilidad de generar agotados e interrumpir líneas de producción continuas, pero también se aumenta la inversión. Por ello es imprescindible explorar modelos más adecuados para administrar adecuadamente los inventarios, integrando con estos, técnicas estadísticas para el pronóstico de la demanda (Correa & Gómez, 2009; Simchi-Levi, Kaminsky & Simchi-Levi, 2008).

Los pronósticos aunque no aseguran una predicción exacta, si permiten una disminución de la incertidumbre, ya que se anticipan a hechos futuros basados en la historia del comportamiento de la demanda, o en análisis cualitativo, permitiendo así tomar decisiones para la planeación de la producción.

En la industria de alimentos, los modelos de inventarios que contemplan una adecuada predicción de la demanda son de gran ayuda para la gestión óptima del abastecimiento, especialmente en los productos perecederos como los lácteos, debido a su corta vida útil y la importancia de su calidad relacionada con la salud humana (Doganis, Alexandridis, Patrinos, & Sarimveis, 2006).

En este trabajo, se parte de la comparación del pronóstico de las ventas de tres productos lácteos elaborados por la empresa ubicada en el Oriente Antioqueño "Auralac", generado por los modelos clásicos de Suavización exponencial simple, Suavización exponencial Holt, modelos dinámicos y de Box-Jenkins, frente a

pronósticos estimados mediante un modelo de redes neuronales artificiales. Su desempeño se probará con la debida verificación de supuestos respectivos y los indicadores del error como: el Error cuadrático medio (MSE sus siglas en inglés) y el error absoluto medio relativo (MAPE sus siglas en inglés).

El pronóstico de ventas derivado de estos modelos será utilizado como insumo para un modelo de optimización, usando, revisión periódica que permita encontrar la forma de emisión de pedidos con mínimo costo acorde con la logística interna antes del abastecimiento.

1. OBJETIVOS

1.1 Objetivo General

Optimizar el sistema de inventarios de la empresa Auralac, basado en la estimación de demanda eficiente.

1.2 Objetivos específicos

- Realizar revisión y descripción bibliográfica sobre los modelos de pronóstico de Suavización exponencial simple, Suavización exponencial Holt, modelos dinámicos, de Box-Jenkins y de RNAs.
- Realizar exploración del modelo de revisión periódica, para minimización de costos de inventarios.
- Realizar análisis y selección de los datos proporcionados por la empresa Auralac.
- Estimar el pronóstico de demanda mediante los modelos clásicos y determinar el más eficiente.
- Estimar el pronóstico de demanda mediante un modelo de RNAs y compararlo con el más eficiente de los modelos clásicos.
- Aplicar un modelo de optimización de inventarios donde se emita un pedido a mínimo costo, basado en la demanda estimada y en una técnica de revisión periódica.

2. ANTECEDENTES

Las series de tiempo son métodos estadísticos que permiten predecir la demanda futura a partir de datos históricos de ventas (Alonso & Villegas, 2005). Dichos modelos dependen de factores como el tiempo, patrones de tendencia, estacionalidad y el comportamiento de la demanda.

Los métodos pronósticos se han vuelto necesarios para enfrentar la aleatoriedad inherente a demandas variables como la de los productos lácteos, enfrentadas a la apertura de nuevos mercados muy competitivos. Son de gran ayuda para planear cantidades a producir diariamente, facilitando la integración de la información compartida con proveedores y cliente y con ello, el abastecimiento en la cadena láctea. A continuación se revisan elementos asociados a técnicas para el modelamiento de inventarios.

Las características de las RNAs permiten la modelación de procesos complejos, como es el caso de pronósticos de demanda con alta variabilidad, que se presenta en este trabajo. Cuando la complejidad del problema lo amerita, o no es posible obtener una solución exacta, los modelos heurísticos permiten a través de un conjunto de estrategias de búsqueda, dar soluciones particulares a estos casos donde se tienen altos niveles de incertidumbre. Entre estas técnicas se encuentran las redes neuronales artificiales (Tabares & Hernández, 2006).

Doganis et al., (2006) unificaron dos métodos heurísticos para la previsión de ventas en el sector lácteo: una arquitectura de red neuronal radial y un algoritmo genético. Estas nuevas metodologías de series de tiempo para la previsión de volúmenes de ventas diarias en productos de corta vida útil como es el caso de la leche, resultan mucho más eficientes en los pronósticos de demanda respecto a modelos sencillos como los lineales y exponenciales con una diferencia alrededor del 6%.

Watson (1987), afirmó que los pronósticos de demanda y los subsistemas de reordenamiento no son independientes entre sí. Patrones de demanda desiguales pueden causar fluctuaciones en los parámetros de la demanda pronosticada, lo que implica un aumento en el costo promedio anual de ordenar y almacenar, y una discrepancia entre el nivel de servicio esperado y el logrado.

Este algoritmo computacional puede facilitar la solución de problemas de inventarios ya que permite la optimización de cualquier proceso que transforme variables de entrada en variables de salida. Las RNAs han mostrado aplicaciones

exitosas, como: la planificación de lotes de producción y requerimiento de materiales (MRP) plantado por Lotfi & Choueiki (2000); la fiabilidad de una red de suministros Srivaree-ratana, Konak & Smith (2002), el mantenimiento y la clasificación de inventarios para una empresa que realiza costeos (Partovi & Anandarajan, 2002); en este mismo año, Koh, Hwang, Sohn & Ko (2002), desarrollaron un modelo para obtener la cantidad de pedidos y el nivel de inventario óptimos para una empresa que abastece su demanda estacionaria por medio de productos nuevos y reciclados, al igual que la debida política de recuperación de esta última.

Los modelos RNAs se han utilizado para la predicción en casos en que la demanda presenta variabilidad y estacionalidad. Au, Choi & Yu (2008) realizaron un modelo de redes neuronales basado en los datos históricos de una empresa de confecciones; el cual su capacidad de predicción para la producción a corto plazo, fue muy acertado frente a otros modelos tradicionales como el SARIMA, especialmente cuando la serie de datos presenta incertidumbre y una demanda con estacionalidad y tendencia. Gutierrez, Solis & Mukhopadhyay (2008) aplicaron pronósticos con RNAs para demanda, comparando su desempeño frente a métodos tradicionales de series de tiempo tales como: suavizado exponencial simple, *Croston*, y aproximación *Syntetos-Boylan*; en esta comparación se concluyó que los modelos de redes neuronales funcionan mejor que los métodos tradicionales. También han sido aplicados para pronosticar demanda de energía diaria y horaria (Medina & García, 2005; Rueda, Velásquez & Franco, 2011).

Vaklieva-Bancheva, Shopova, Espuña, & Puigjaner (2006) proponen un modelo que responde a las fluctuaciones de las demandas rápidamente, desde una mirada de la cadena de suministro del sistema de producción y la importancia de integrar la capacidad de producción y los tiempos cortos de respuesta, característicos en los productos perecederos; para ello definieron variables como costos de transporte, producción, demandas, capacidades de distribución, distancias entre las plantas, mercados y centros de distribución. El tiempo de respuesta para éste tipo de alimentos debe estar acompañado de una adecuada administración del inventario.

En lo relacionado a la formulación de inventarios se muestran a continuación los temas logísticos considerados en el trabajo presentado.

Chiu (2003) determinó el tamaño del lote óptimo de inventario para minimizar los costos de producción, considerando variables como: una tasa aleatoria de productos defectuosos, su proceso de reparación y el retrasos en los pedidos; siguiendo la misma línea en el 2009 Wee & Chung desarrollaron un modelo integrado que optimizaba los inventarios deteriorados y las debidas políticas para su reposición. Los autores se apoyaron en un modelo de revisión periódica que relacionaba la producción de los proveedores y los tiempos de almacenamiento, así consideraron a partir de inventarios imperfectos y con escases de pedido, un modelo óptimo de inventario.

Vidal & Londoño (2004) revisaron técnicas de modelos de inventarios, de pronósticos de demanda y su integración para la optimización de estos. Dentro de las estrategias investigadas por los autores para la optimización de inventarios en una cadena de suministros, se encuentra la obtención y transmisión de información en tiempo real de los puntos de consumo y las respectivas líneas de productos; la consolidación de centros de distribución y bodegas con estandarización de productos, lo cual evita el mantenimiento de una diversidad de inventarios que sólo difieren en aspectos menores de estructura física como: el peso, el tamaño y la forma. Según los autores el principal problema de los modelos es la estimación de inventarios de seguridad y los puntos de re-orden, sólo como un simple promedio de la demanda, ignorando su variabilidad y especialmente en los tiempos de reposición, los cuales en la mayoría de los casos, no coincide con el valor adecuado obtenido a través de un análisis estadístico formal.

Benseman (1986), desarrolló un modelo de programación lineal para los planes de producción a mediano plazo de la compañía láctea más grande de nueva Zelanda, donde se muestran las fluctuaciones en la disponibilidad y calidad de la leche entera, los costos, la capacidad instalada, la productividad de los productos y sus precios. Gracias a la implementación del modelo la compañía ha ahorrado más de un millón de dólares por año. El autor concluye que los modelos matemáticos permiten considerar más información y tomar mejores decisiones.

Rob & Broekmeulen (2009) proponen y evalúan una nueva política de reaprovisionamiento utilizando un modelo de simulación del proceso de venta al por menor de productos perecederos, en el cual integran un factor que no se tenía en cuenta con anterioridad en ambientes estocásticos, la edad del inventario. Los autores concluyen que tener este factor en cuenta conduce a una reducción significativa en los costos, pero que para ello se requieren registros información detallada por medio de tecnologías modernas.

En el recorrido de la utilización de pronósticos para la optimización de inventarios a través del tiempo, es apreciable que algunos modelos resultan ser más eficientes que otros dependiendo de las variables consideradas. Integrar el modelo más eficiente de pronósticos con la optimización de inventarios de producto terminado y lograr una reducción de costos asociados a él, en un caso aplicado, resulta de gran interés para la industria láctea, y es allí donde este trabajo desea apuntar.

3. MARCO TEÓRICO

3.1 Estadística descriptiva

Cuartiles

Los cuartiles son tres valores que dividen el conjunto de datos en cuatro partes iguales, conteniendo cada una de ellas el 25% de las observaciones. El tercer cuartil constituye el cuarto superior del conjunto de datos, el segundo cuartil es idéntico a la mediana, que deja el 50% por debajo de las observaciones y el primer cuartil separa el cuarto inferior de los tres cuartos superiores.

Rango intercuartil

Una medida de dispersión resistente a los valores apartados es la dispersión de los cuartiles f , dada por:

$$f = \text{CuartilSuperior} - \text{CuartilInferior}$$

Diagrama de caja y bigotes

Es una representación gráfica de un conjunto de datos o valores, que permite visualizar la dispersión, el grado y la dirección del sesgo. El gráfico comprende: la media, la mediana, el primer y tercer cuartiles, y por ende el rango intercuartil, el máximo y el mínimo de los valores y los datos que están muy alejados de la mayoría, llamados **datos atípicos**, que quedan localizados fuera del tercer cuartil más 1.5 veces el rango intercuartil y fuera del primer cuartil menos 1.5 veces el rango intercuartil (Ruíz & Morillo, 2004).

3.2 Series de tiempo

Una serie de tiempo o serie temporal, se conoce como el conjunto de datos que son obtenidos y registrados de forma sucesiva con intervalos fijos en un periodo determinado, de esta forma se realizan observaciones a través del tiempo a una variable particular.

Tal como lo afirma Pepió (2001), el objetivo del análisis de una serie temporal es la identificación de patrones de comportamiento para prever la evolución futura, asociadas a fenómenos aleatorios; de esta forma el estudio del comportamiento pasado permite crear un modelo probabilístico para la predicción del futuro, los cuales son denominados procesos estocásticos. Para identificar y posteriormente evaluar ese patrón, es conveniente analizar varios componentes dentro de la

serie, lo cual se conoce como “descomposición”, entre estos componentes se encuentran: Tendencia, ciclo, estacionalidad y fluctuaciones irregulares.

Tendencia

La tendencia es el componente que se encarga de representar el crecimiento (o la declinación) subyacente en una serie de tiempo evaluado generalmente a largo plazo.

Ciclo

El comportamiento cíclico es una serie de fluctuaciones de más de un año de duración, son generalmente asociadas a las condiciones económicas cambiantes. *“En la práctica resulta difícil identificar ciclos y con frecuencia se le considera parte de la tendencia (...) es conocido como ciclo-tendencia”* (Hanke & Wichern, 2006).

Estacionalidad

Las variaciones estacionales se conocen a las oscilaciones que se producen y repiten en periodos cortos de tiempo, las cuales son asociadas a factores dinámicos en el tiempo, por ejemplo: clima, costumbres, sucesos relativos al calendario, como periodos de vacaciones, días feriados, entre otros (Bowerman & O’Connell, 2007).

Fluctuaciones irregulares

Consiste en las fluctuaciones impredecibles o aleatorias, dichas fluctuaciones representa lo que “resta” en una serie temporal después de ser explicados la tendencia, los ciclos y las variaciones estacionales (Bowerman & O’Connell, 2007).

3.3 Pronósticos

Los pronósticos son las predicciones de los hechos y condiciones futuras, su aplicación en los procesos de toma de decisiones son de gran importancia para algunas empresas, en especial si son comerciales, tal como lo afirma Bowerman & O’Connell (2007). Simchi-Levi et al., (2008) afirman que pronósticos permiten reaccionar con rapidez en los procesos administrativos de planeación, programación, producción y control, logrando así un incremento en la eficiencia operaciones.

Los pronósticos se pueden hacer a corto, mediano y largo plazo, a través de dos enfoques, tal como existen en todos los modelos decisiones, uno es el análisis

cuantitativo y otro es el enfoque cualitativo. Render (2004) define los pronósticos cualitativos como aquellos que incorporaran factores intuitivos, basados en la experiencia y el sistema de valores de los encargados de la toma de decisiones; por otro lado, los pronósticos cuantitativos se apoyan de modelos matemáticos basados en datos históricos o variables causales para pronosticar

3.4 Modelos de pronóstico

3.4.1 Modelos de suavización

Suavización exponencial simple

El método de suavización exponencial usa un promedio ponderado de los valores pasados de la serie de tiempo, así lo afirma Anderson & Sweeney (2008), en este modelo sólo se elige un peso para la observación más reciente, los demás datos se calculan de forma automática.

La ecuación para el modelo de suavización exponencial es:

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

Donde,

F_{t+1} : Pronóstico de la serie de tiempo para el periodo $t+1$

Y_t : Valor real en el periodo t de la serie de tiempo

F_t : Pronóstico para el periodo t de la serie de tiempo

α : Constante de suavizamiento, varía de 0 a 1

Según Hanke & Wichern (2006), la constante de suavizamiento determina la medida en que un dato actual influye en la siguiente observación de la serie temporal; por ello entre más próximo sea a la unidad, el nuevo pronóstico será más cercano la observación actual, y de forma contraria, entre más cercano a cero, será similar a la predicción del periodo inmediatamente anterior.

La elección de la constante α , se realiza con base a un proceso iterativo para minimizar el error cuadrático medio (MSE).

Suavización exponencial Holt

El método de suavización exponencial simple no tiene en cuenta la tendencia, es por ello que el método de Holt (en honor a su creador) la incluye, lo cual permite integrar las tendencias lineales locales en evolución dentro de una serie de tiempo.

Para Hanke & Wichern (2006), el modelo de Holt suaviza el nivel y la pendiente de manera directa, a través de dos constantes de ponderación: α (alfa) y β (beta) respectivamente. Según los autores una de las ventajas del modelo es la flexibilidad al seleccionar los coeficientes anteriormente mencionados.

Para la estimación del pronóstico con este modelo es necesario resolver las siguientes ecuaciones:

- **Estimación nivel actual:**

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + T_{t-1})$$

- **Estimación de la tendencia**

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$

- **Pronostico en el periodo “p”**

$$\hat{Y}_{t+p} = L_t + pT_t$$

Donde,

L_t : Valor nuevo suavizado

Y_t : Valor real de la serie en el periodo t (actual)

T_t : Tendencia estimada

p: Periodo a pronosticar

$\hat{Y}_{t+p}$: Pronóstico para el periodo p

α : Constante de suavizamiento para el nivel, varía de 0 a 1

β : Constante de suavizamiento para la tendencia, varía de 0 a 1

3.4.2 Modelos de regresión lineal

La relación entre dos o más variables relacionadas de un modo no determinístico, se conoce como análisis de regresión (Devore, 2004), en la cual se estudian las variaciones de una variable respuesta en función de una o más variables independientes o predictivas, dependiendo del número de estas variables pueden ser regresión simple (una variable) o regresión múltiple (más de una variable).

Modelo de regresión lineal simple

Es la relación matemática más simple entre dos variables (una dependiente y otra independiente), en un modelo probabilístico el valor esperado de Y (variable dependiente) es una función línea de la variable independiente X , (Montgomery, Peck, & Vining, 2006).

La ecuación que define este modelo es:

$$Y = \beta_0 + \beta_1 X + \varepsilon$$

Donde,

Y : Variable respuesta o variable dependiente

X : Valor de la variable predictora

β_0 : Valor de la variable dependiente cuando la variable predictora (X) es cero

β_1 : Coeficiente de la regresión lineal, indica el incremento en la media de la variable respuesta por cada unidad de aumento en la variable independiente

ε : Desviación aleatoria o término de error aleatorio

Modelos de regresión múltiple

Como se explicó anteriormente este modelo incluye más de una variable independiente, estas variables pueden ser de muchos tipos, para nuestro caso se evaluarán modelos tentativos de pronósticos como: regresión múltiple con componentes: lineales, polinómicas, trigonométricas (senoidales o cosenoidales), auto-regresivos y variables categóricas (Bowerman & O'Connell, 2007).

Regresión múltiple con componente lineal

Al igual que la regresión lineal simple, las variables son una función lineal, pero en este caso se incluyen más de una variable regresora (Montgomery et al., 2006)

En general se puede relacionar la respuesta con k regresores o variables predictoras con el siguiente modelo:

$$Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \varepsilon$$

Donde Y es la variable respuesta, X_j con $j = 0, 1, \dots, k$ son las variables independientes, los coeficientes β_j con $j = 0, 1, \dots, k$ se conocen como coeficientes de regresión o coeficientes de regresión parcial, los cuales representan el cambio esperado en la respuesta si se mantienen las demás variables independientes constantes, y al igual que el modelo de regresión simple se tiene un error aleatorio ε .

Regresión con componente polinómico

Esta regresión se analiza de forma análoga a los dos modelos anteriores. El modelo puede ser nombrado según el grado del polinomio, por ejemplo si el polinomio es de grado dos, se habla de un modelo de segundo orden.

La ecuación para un modelo de regresión polinomial de k -ésimo grado es:

$$Y = \beta_0 + \beta_1 X + \beta_2 X^2 + \dots + \beta_k X^k + \varepsilon$$

En este caso el valor esperado de Y es una función polinomial de X de k -ésimo grado, con coeficientes de regresión y error aleatorio similar a la regresión múltiple.

Para Montgomery et al (2006) los modelos polinomiales son útiles como funciones de aproximación a relaciones no lineales, desconocidas o en algunos casos complejas.

Regresión múltiple con componente senoidal

Combinar términos polinomiales y trigonométricos son convenientes para los modelos, según Montgomery et al (2006) en especial cuando la serie temporal presenta estacionalidad.

Al modelo de regresión se le incluye las variables:

$$\text{Sen}\left(\frac{2\pi n}{L} t\right)$$

Donde,

n : Número entero que determina la cantidad de armónicos en la función trigonométrica

L : Periodo en el cual la serie temporal es estacional

t : Variable de tiempo

La función seno, puede ser cambiada por la función coseno. Con este tipo de funciones es posible describir movimientos ondulatorios más complejos (Aguirre, 1994), donde el tiempo es la variable independiente para el componente senoidal; como todo modelo de regresión tiene asociado un coeficiente que en el modelo determina la amplitud de la función trigonométrica.

Regresión múltiple con componente autorregresivo

La autocorrelación dentro de las series temporales hace que los valores de la variable dependiente en un periodo estén relacionados linealmente con los valores de la variable dependiente en otro periodo, así lo afirma Hanke & Wichern (2006). De esta forma es posible resolver el problema de autocorrelación que se presentan en las series de tiempo.

Los modelos de regresión que incluyan éste componente se conocen como *modelos autorregresivos*. Los modelos de primer orden se comportan de forma genérica con la siguiente ecuación:

$$Y_t = \beta_0 + \beta_1 Y_{t-1} + \varepsilon$$

En la ecuación anterior se presenta el rezago para el periodo inmediatamente anterior (Y_{t-1}), el cual puede ser remplazado por Y_{t-n} , para realizar un rezago en el n-ésimo periodo.

Regresión múltiple con variable indicadora

Las variables independientes cuantitativas pueden ser agregadas a los modelos dinámicos a través de niveles, los cuales son descritos con variables ficticias o también llamadas indicadoras, el efecto que estas variables tienen sobre el modelo dinámico puede ser analizado gracias a las pruebas de significancia (Bowerman & O'Connell, 2007).

De forma análoga, los modelos econométricos dinámicos al igual que los modelos de regresión lineal, también están formados por una o varias ecuaciones en las que la variable explicada (variable endógena), la cual es influida por alguna otra variable del modelo (variables explicativas). Los modelos dinámicos se diferencian de los modelos de regresión cuando además de incluir variables explicativas o llamadas exógenas, también se incluyen variables endógenas retardadas las cuales incluyen al modelo valores del pasado de la serie temporal, tal como lo afirma Caridad y Ocerín (1998).

3.4.3 Modelos Box Jenkins- ARIMA

Un modelo Box Jenkins es una declaración algebraica que muestra como una variable (Z_t) de la serie de tiempo está relacionada con sus propios valores pasados ($Z_{t-1}, Z_{t-2}, \dots, Z_{t-n}$).

Es especialmente adecuado para realizar pronósticos a corto plazo de series con variación estacional, incluyendo cambios en los patrones estacionales y cuando la serie de tiempo es estacionaria. (Pankratz, 1983)

Entre los modelos Box Jenkins se encuentra:

- **Modelo autorregresivo de orden p, AR (p):**

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

Donde,

θ_0 : Constante.

ϕ_1 : Parámetro que relaciona a z_t con z_{t-1} .

a_t : Choque aleatorio estadísticamente independiente, y con distribución normal, también llamado ruido blanco.

- **Modelo de media móvil de orden q, MA (q):**

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Donde,

μ : Constante.

θ_1 : Parámetro que relaciona a a_t con a_{t-1} .

a_t : proceso estocástico llamado ruido blanco, el cual tiene media cero, varianza constante e incorrelación.

- **Modelo autorregresivo y de medias móviles, ARMA(p, q):**

$$Z_t = \theta_0 + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Esta clase de modelo es una alternativa natural cuando el número de parámetros necesitado puede ser excesivamente grande, frente al número disponible de datos recolectados para el análisis (Bowerman & O'Connell, 2007).

Para identificar cuál de estos modelos se ajusta mejor a los datos, se seleccionan tentativamente uno o varios modelos Box Jenkins por medio de las gráficas de autocorrelación (**ACF**) y de autocorrelación parcial (**PACF**), ambas derivadas de los datos de la serie. Éstas representan un correlograma de rezagos de la serie de tiempo, que explican el grado de autocorrelación en la serie rezagada K periodos.

La selección del tipo de modelo depende de funciones de autocorrelación específicas, como se muestra a continuación:

Tabla 1. Identificación tentativa para los Modelos Box Jenkins.

Modelo	Función de autocorrelación (ACF)	Función de autocorrelación parcial (PACF)
MA(q)	Decrece con rapidez a cero después del orden q.	Decrece exponencialmente, senoidalmente o lentamente.
AR(p)	Decrece exponencialmente, senoidalmente o lentamente.	Decrece con rapidez a cero después del orden p.
ARMA(p, q)	Decrece exponencialmente, senoidalmente o lentamente.	Decrece exponencialmente, senoidalmente o lentamente.

Fuente: Bowerman & O'Connell, 2007

Cuando la serie de tiempo no es estacionaria se puede recurrir a una transformación llamada diferenciación, frecuentemente aplicada a los datos para inducir una media estacionaria. Así la denominación de los modelos se modifica a ARI, MAI o ARIMA según el caso.

La estacionalidad puede ser modelada con los modelos ARIMA estacionales multiplicativos, quedando un modelo que describe una parte no estacional combinada con una estacional (Bowerman & O'Connell, 2007), de tipo:

SARIMA (p, d, q) X (P, D, Q) [L]

Donde,

L: Periodo estacional

p: Orden AR no estacional

d: Orden de diferenciación en componente no estacional

q: Orden MA no estacional

P: Orden de SAR o componente autorregresivo estacional

D: Diferenciación del componente estacional

Q: Orden SMA o media móvil del componente estacional

3.4.4 Redes Neuronales Artificiales (RNAs)

Una red neuronal es un centro de procesamiento de información de cuyas interacciones locales depende el comportamiento del conjunto del sistema.

Las RNAs son una de las ramas más destacadas del campo científico de la Inteligencia Artificial. Éstas tratan de simular el proceso de razonamiento humano, caracterizado por el aprendizaje a través de la experiencia y la extracción de conocimiento a partir de un conjunto de datos, e imitan esquemáticamente la estructura neuronal del cerebro.

El elemento fundamental de los sistemas neuronales biológicos es la neurona, ya que es la clave para el procesamiento de la información. De su cuerpo celular parte un denso árbol de ramificaciones compuesto por dendritas, donde llega la información, se procesa y se emite una señal por medio del axón o rama principal.

La estructura de la red está conformada por una capa de entrada, de salida y capas ocultas. Estas últimas reciben estímulos y emiten salidas dentro del sistema. En ellas se lleva a cabo el procesamiento básico de la información. (Flórez & Fernández, 2008).

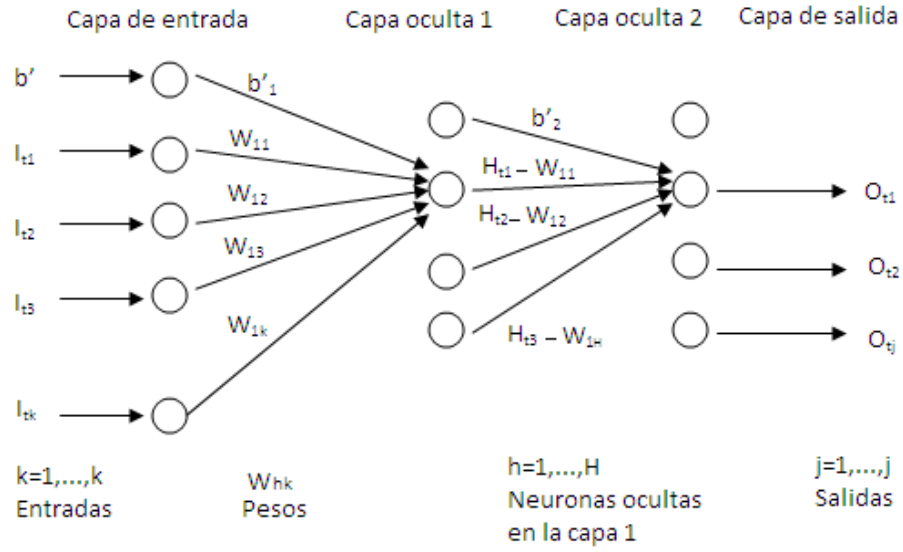


Figura 1. Estructura de una RNA Elaboración propia

En la primera etapa se agrega la información con una combinación lineal de los pesos asignados y neuronas de entrada y en la segunda etapa se aplica una función de transferencia hacia las neuronas de salida.

En cada neurona ocurren dos procesos, hay entradas y hay un umbral o señal mínima que se transmite a las otras neuronas, la cual está dada por:

$$b_h o b_j = f(\overline{W_{1k}} \cdot \overline{I_{tk}} + \overline{b'_1})$$

Los pesos W_{hk} pueden representarse matricialmente como:

$$IW^{10} = \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1k} \\ \vdots & \vdots & & \vdots \\ W_{h1} & W_{h2} & \dots & W_{hk} \end{bmatrix} IW^{21} = \begin{bmatrix} W_{11} & W_{12} & \dots & W_{1h} \\ \vdots & \vdots & & \vdots \\ W_{j1} & W_{j2} & \dots & W_{jh} \end{bmatrix}$$

La capa oculta 1 está definida por la función:

$$Net_{th} = b'_h + W_{h1}I_{t1} + W_{h2}I_{t2} + \dots + W_{hk}I_{tk}$$

Un ejemplo de una función de transferencia para las neuronas que salen de la capa oculta 1, es la función logística:

$$H_{t_h} = \frac{1}{1 + e^{-Net_{t_h}}}$$

La capa oculta 2 está definida por la función:

$$Net_{t_j} = b_j^2 + W_{j_1}H_{t_1} + W_{j_2}H_{t_2} + \dots + W_{j_k}H_{t_H}$$

La función logística de transferencia para las neuronas que salen de la capa oculta 2 será entonces:

$$\theta_{t_j} = \frac{1}{1 + e^{-Net_{t_j}}}$$

Generalmente en la capa final de salida se usa una función lineal:

$$W_i(w, x) = \sum_{j=1}^n w_{ij} x_j$$

La función de salida de la red es:

$$O_{t_j} = f^2 \left[b^2 + IW^{21} \left(f' (b' + IW^{10} I_{t_k}) \right) \right]$$

3.5 Pruebas de validación

Las pruebas de validación o la validación de supuestos; se conocen como aquellas pruebas que se realizan sobre los residuales del modelo, donde se formulan algunos supuestos, los cuales son precisos verificarlos; esto permite que a la hora de hacer inferencias sobre los datos estimados, se tenga una validez estadística que lo sustenta, pero aún así con los riesgos inherentes que se tiene con estas decisiones estadísticas.

Con estas pruebas también es posible verificar cualquier anomalía que se tenga sobre el ajuste de los modelos de pronóstico.

Los supuestos generales que se contemplan para los modelos son: la independencia, la normalidad, la media cero y la varianza constante.

Para nuestro caso, se desarrollan las siguientes pruebas:

Prueba de normalidad Jarque-Bera

Con esta se determina la bondad de ajuste a la distribución normal de una variable aleatoria. La prueba realizada por C.M. Jarque y A.K. Bera se basa en el estudio de la forma de la distribución, examinando las discrepancias respecto a la campana uniforme que siguen la distribución normal; estas discrepancias se evalúan con respecto a la simetría y la kurtosis.

Esta prueba se realiza sobre los residuos de un modelo estimado a través de un proceso de máxima verosimilitud, el cual se obtiene con la ecuación:

$$JB = N * \left[\frac{S^2}{6} + \frac{(K - 3)^2}{24} \right] \approx \chi^2$$

Donde N es el tamaño de la muestra, S es el coeficiente de asimetría y K es la kurtosis o apuntamiento. La prueba se realiza bajo la hipótesis nula de normalidad, el cual para ser aceptada, el estadístico de Jarque-Bera debe ser mayor al grado de significancia.

Prueba de incorrelación de Box-Pierce

Permite determinar si todos los coeficientes de autocorrelación del conjunto son cero, es decir si son independientes. Para los residuales de un modelo de pronóstico la estadística Q de Box-Pierce tiene una distribución chi cuadrada con los grados de libertad iguales a m , menos el número de parámetros estimados en el modelo. Está definida por:

$$Q = n(n + 2) \sum_{k=1}^m \frac{r_k^2}{n - k}$$

En esta ecuación, n es el número de observaciones en la serie de tiempo, k el retraso del tiempo, m el número de retrasos de tiempo a ser probados y r_k la

función de autocorrelación de la muestra de los residuales retrasados k periodos. (Hanke & Wichern, 2006)

De manera alterna se puede calcular e interpretar el valor p generado mediante el estadístico Q de prueba. Si éste es superior al grado de significancia se confirma que hay incorrelación en los residuales.

3.6 Criterios de evaluación

Gracias a los criterios de evaluación, es posible comparar los valores pronosticados frente a los valores reales de una serie de datos; de esta forma se evalúa comparativamente la efectividad entre diferentes modelos de pronósticos, al igual que permite realizar un seguimiento y verificación de su desempeño.

Siempre se busca el valor menor en los criterios de evaluación, entre los cuales se encuentran:

Error cuadrático medio (MSE)

Es una forma de medir el error global del pronóstico. Consiste en el promedio de los cuadrados de las diferencias entre los valores pronosticados y observados. (Render, 2004). La función que lo define es:

$$MSE = \frac{\sum (\text{Errores de pronóstico})^2}{n}$$

Error porcentual absoluto medio (MAPE)

Es otra forma de medir el error global del pronóstico. El MAPE es el promedio de las diferencias absolutas entre los valores pronosticados y los reales, expresado como porcentaje de los valores reales (Ramírez, 2013). Su fórmula es:

$$MAPE = \frac{\sum_{i=1}^n |real_i - pronóstico_i| / real_i}{n} * 100$$

3.7 Modelos de inventarios

Los inventarios son uno de los mayores rubros en los estados financieros de las compañías, razón por la cual una acertada administración permite un adecuado nivel de servicio a un costo razonable. Los inventarios permiten amortiguar las diferencias que se presentan entre la oferta y la demanda, según Izar (1996), entre otras ventajas de tener un adecuado nivel de inventarios se encuentran:

- Satisfacer la demanda aún durante los tiempos de entrega - *lead time* -
- Oportunidad de descuentos por la compra de insumos en lotes de mercancías mayores
- Regular la carga de producción ante una demanda estacionaria

Para responder las preguntas de cuándo y cuánto siempre buscando el mínimo costo para la compañía y el mínimo capital ocioso, es posible usar modelos de inventarios que pueden ser determinísticos en los cuales se conoce con certidumbre la demanda de los productos o un modelo estocásticos en los cuales se pueden describir con una distribución de probabilidad.

Para la estimación del inventario, se tienen asociados algunos componentes que son generales e independientes del modelo seleccionado para su determinación, los cuales son:

- Demanda
- Costos de ordenar
- Costos de mantener o almacenar
- Costos de penalización por faltantes
- Tiempo de entrega

Otro tipo de clasificación para los inventarios se relaciona con la forma en que se revisa el inventario (Hillier & Liberman, 2002) puede ser de revisión continua, en el cual se hace un pedido en el momento en que el inventario baja de un punto de reorden establecido o un modelo de revisión periódica en el cual se verifica el nivel de inventarios con intervalos discretos. Este último modelo será el empleado para realizar el modelo de inventarios.

Modelo determinístico de revisión periódica

En un modelo determinístico con revisión periódica la demanda varía entre dos límites: un límite superior y otro inferior, en el cual se deben planear para los

siguientes n periodos respecto a cuanto producir u ordenar para reabastecer el inventario al principio de cada periodo.

La demanda se denota con la variable r_i la cual varía para el periodo $i = 1, 2, \dots, n$

Las variables de costos son:

K = Costo de preparación para producir u ordenar artículos para reabastecer el inventario al inicio del periodo

c = Costo unitario de producir u ordenar cada artículo

h = Costo de mantener en inventario cada artículo que queda al final del periodo

El objetivo del modelo es minimizar el costo total durante los n periodos, el cual se logra sólo si se ignoran los costos fijos que afectan la producción.

La solución del modelo se logra mediante una fórmula basada en una política óptima de un programa de producción en el cual sólo se produce cuando el nivel de inventario es cero. Además, en la restricción de los pedidos no necesariamente debe producirse el valor de demanda diaria, sino que puede aumentarse el lote de producción hasta máximo la suma de demandas de los n periodos, dependiendo del tipo de proceso (Taha, 2004)

4. METODOLOGÍA

El trabajo parte de un análisis descriptivo, seguido de una aplicación de modelos estadísticos clásicos y Redes Neuronales Artificiales para realizar pronósticos de ventas de 3 productos de la empresa Auralac, seleccionando el mejor de estos con base en los criterios MAPE y el MSE; seguido a esto, se usan los mejores pronósticos para un modelo de optimización de inventarios, basado en revisión periódica, que definirá las decisiones para programar pedidos y almacenamiento de inventarios al mínimo costo.

Para comenzar, se explicará el proceso de análisis descriptivo.

Se utilizaron los datos históricos de salidas por facturación, consideradas ventas, desde el primero de enero de 2010 hasta el 24 de diciembre de 2012, de tres productos elaborados por la empresa Auralac, proporcionados por ellos, de manera confidencial, sin dar sus nombres.

Para identificar los datos atípicos se realizó un diagrama de caja y bigotes para la serie de datos de salidas por facturación de cada una de las referencias, en el software *Statgraphics*. Luego se hizo una ponderación de cada uno de ellos con sus dos datos anteriores, con el fin de evitar sesgos en los pronósticos.

Para la identificación de patrones en la series se generaron, mediante el software *R*, las gráficas ACF y PACF; se graficó la serie de tiempo sólo para los dos últimos meses de cada una de las referencias, con el objetivo de tener una mejor visualización de la correlación y estacionalidad; así como gráficos de tendencia y variabilidad. Además se realizó la prueba de autocorrelación Box-Pierce para completar el análisis de estacionariedad.

4.1. Estimación de modelos de pronóstico

Para los siguientes siete días, se estimaron modelos dinámicos, de Suavización exponencial y de Box-Jenkins para pronosticar las ventas de los siguientes siete días, por medio de los programas estadísticos *Statgraphics* y *R*. Además se elaboró una RNA para cada una de las series, mediante la herramienta de software matemático *MATLAB*.

Inicialmente se realizó un modelo de Suavización exponencial simple y Suavización exponencial Holt para cada una de las referencias, mediante la función automática del software *Statgraphics*.

Para la definición de variables del modelo dinámico (las variables son diferentes, según el comportamiento de cada serie), se utilizó el análisis descriptivo y las gráficas ACF y PACF. Debido a los cambios estructurales que presentan de forma general en las tres series y las evidencias de estacionalidad cada 7 días, se incluyeron variables ficticias al modelo como: día de la semana y nivel. Luego de la estimación de los modelos, se seleccionaron aquellas que presentaron significancia inferior al 5% o 10% (encima del 95% o 90% de confianza) y se estimó el pronóstico.

Se analizaron las gráficas ACF y PACF para cada una de las series y a partir de ello se definieron los modelos Box Jenkins que más se ajustaran a los datos. Además se hizo uso de la función automática del software *Statgraphics* para generar más modelos en cada uno de los productos.

Verificación de supuestos

Se realizó la prueba de validación de normalidad Jarque-Bera, de autocorrelación Box-Pierce y el MSE, para cada uno de los modelos. Pero el criterio final de selección del modelo más indicado para la estimación del pronóstico fue el MAPE.

Según esto se determinó que de los modelos clásicos, los más eficientes son los modelos dinámicos y por lo tanto se prosiguió con la comparación de éstos con las RNAs.

Proceso de entrenamiento y simulación de la Red neuronal

El proceso para la estimación de las RNAs se explica en la figura 2, partiendo de las variables más significativas derivadas del proceso anterior.

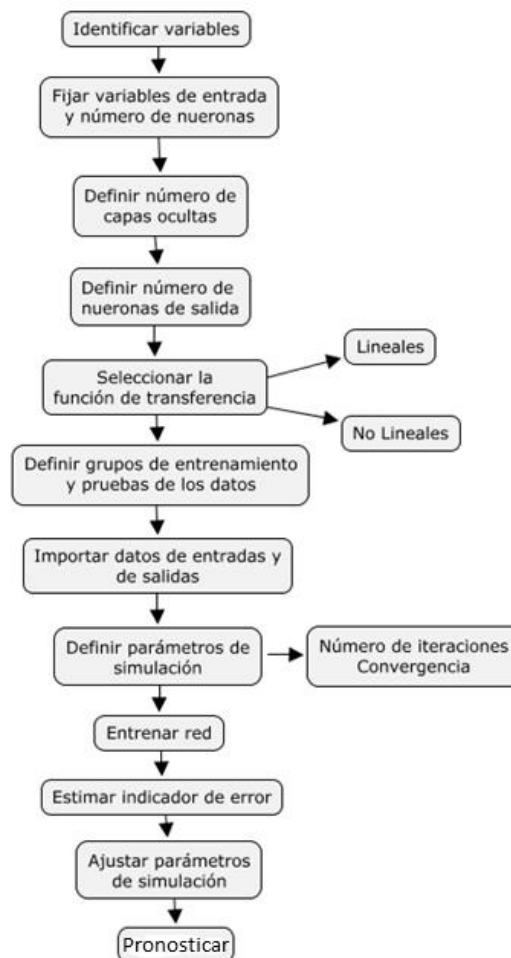


Figura 2. Metodología para entrenar y simular una red Neuronal. Elaboración propia

Las variables significativas tomadas de los modelos dinámicos, constituyen las entradas de los modelos, los datos históricos de ventas las salidas. Se crearon bases de datos en *Microsoft Excel* para cada una de las referencias y se ingresaron a *MATLAB*.

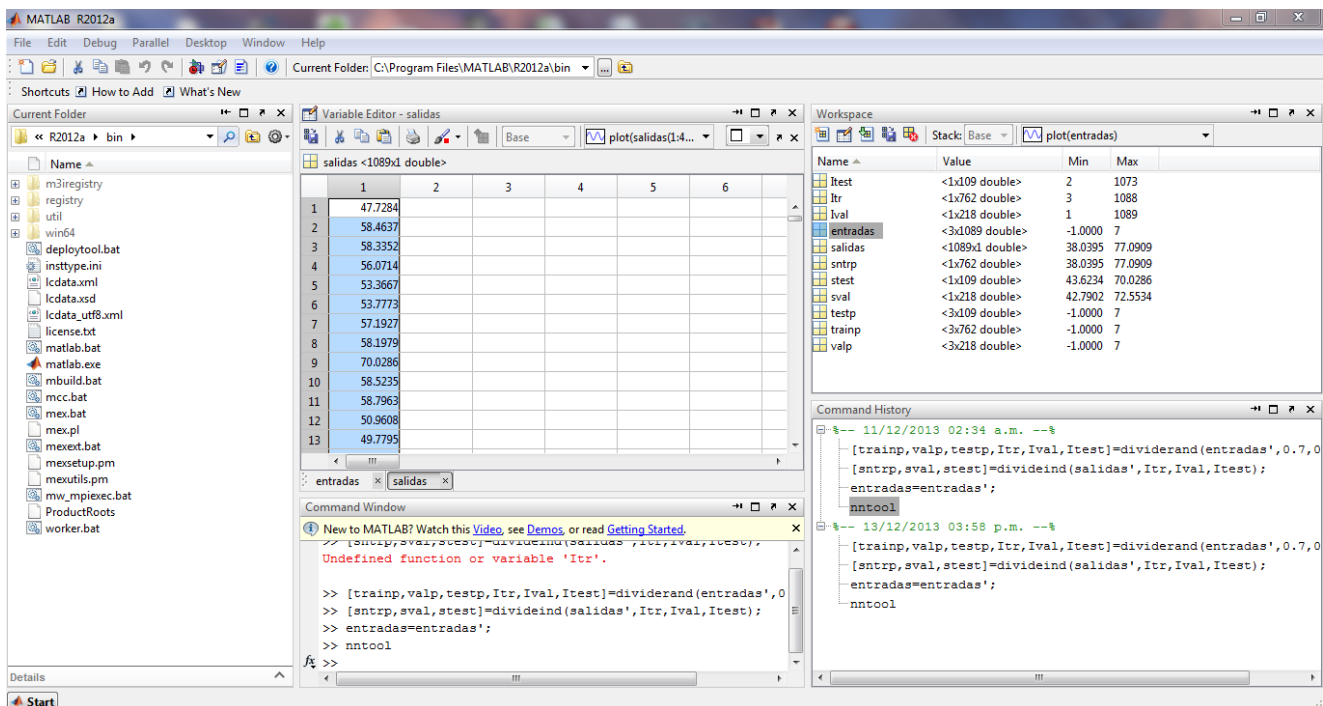


Figura 3. Asignación de variables de entrada y de salida a MATLAB

Se dividen las entradas y las salidas en tres subgrupos y se importan para la creación de la red.

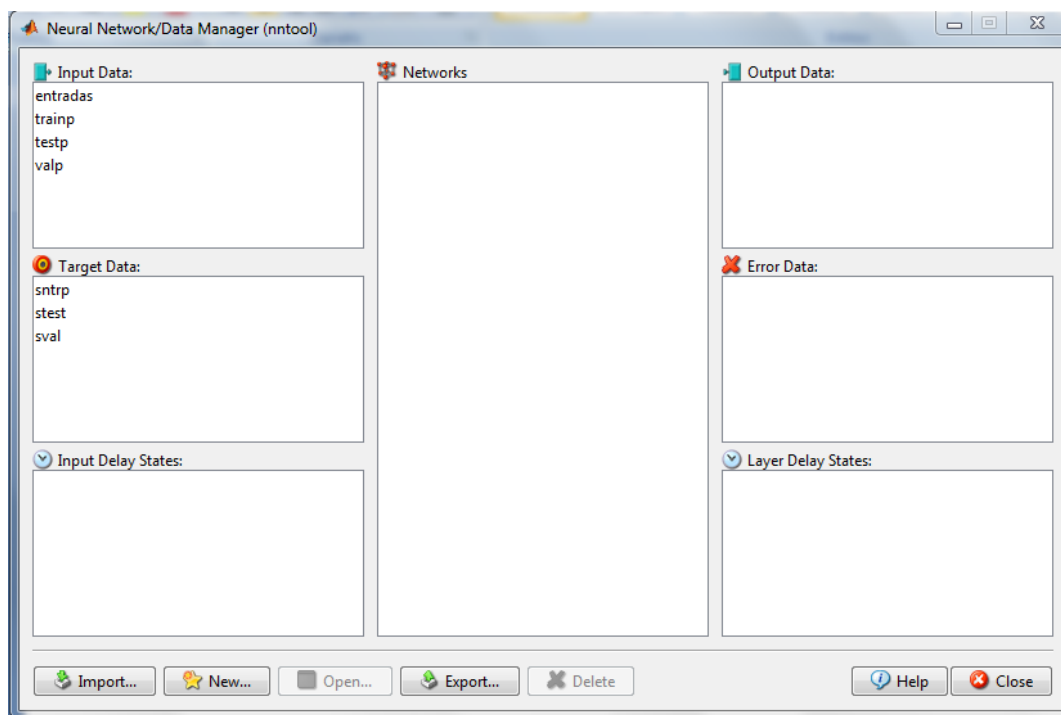


Figura 4. Importación de entradas y salidas para la creación de la RNA en MATLAB

Para la creación de la red se ingresaron el número de neuronas y de capas ocultas, la función de rendimiento, de entrenamiento, de aprendizaje y de transferencia.

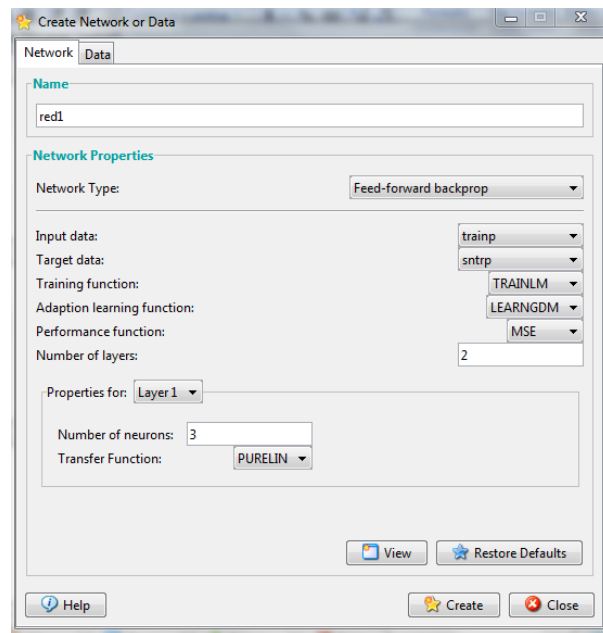


Figura 5. Creación de la RNA en MATLAB

Luego se realizó el entrenamiento de la red. Se definieron el número de iteraciones y se evaluó el rendimiento de la red por medio del MSE, para compararlos en una primera instancia con los modelos dinámicos.

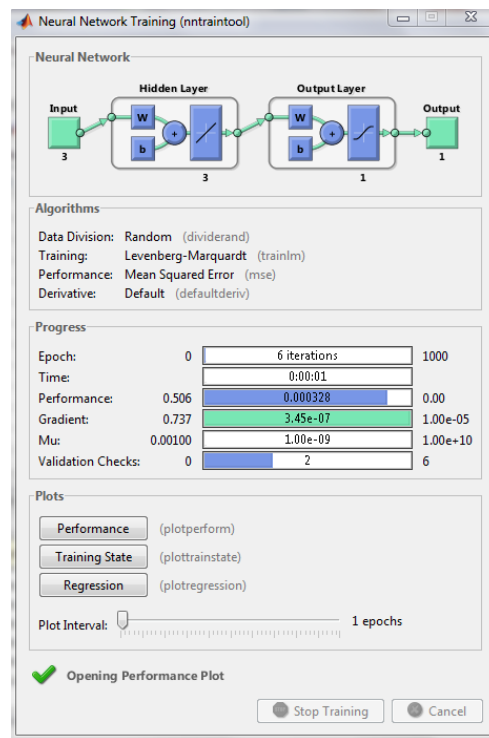


Figura 6. Entrenamiento de la RNA en MATLAB

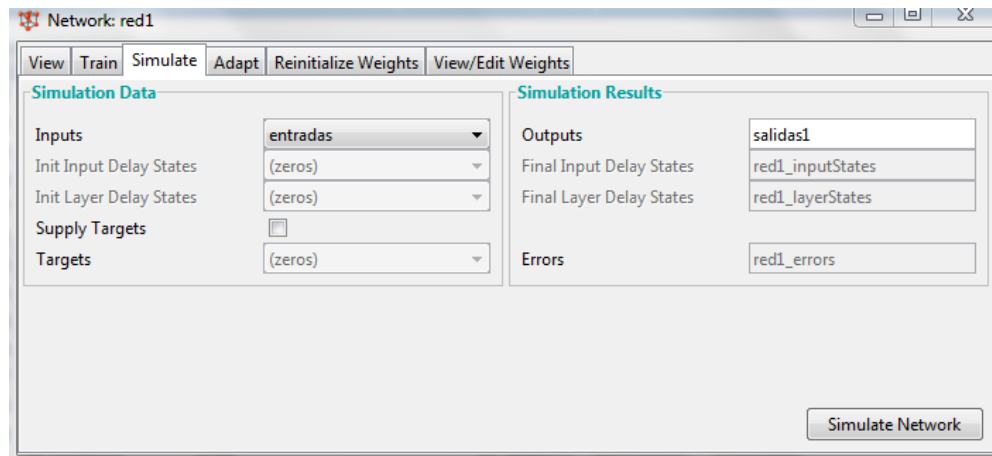


Figura 7. Simulación de la RNA en MATLAB

Por último se ingresaron las entradas de los próximos siete días, se simuló la red y se llevaron los pronósticos al programa Microsoft Excel para estimar el MAPE de ajuste y de pronóstico y a partir de éstos elegir el mejor modelo.

Por último se integró el pronóstico del mejor modelo a un modelo de revisión periódica, realizado en el programa *LINGO*, con el fin de optimizar el sistema de inventarios de la empresa.

4.2. Modelo de optimización basado en revisión periódica

El objetivo del modelo de revisión periódica fue determinar las cantidades óptimas a producir en la semana del 25 al 31 de diciembre de 2012, de cada una de las referencias, de manera que se minimicen los costos de los inventarios, asegurando que se cumpla con la utilidad operacional mínima esperada por la empresa.

Para la estimación del modelo se tuvo en cuenta los siguientes componentes asociados a los inventarios, para cada una de las referencias:

- Demanda.
- Costo de producción.
- Costo de mantener.
- Utilidad.
- Niveles máximos a almacenar semanalmente por referencia.

Las cantidades óptimas a producir deben cumplir con la demanda mínima de los siete días o periodos siguientes (pronóstico) y no deben superar el tope designado por política de la empresa para cada una de las referencias. Además se debe asegurar que éstas no superen la capacidad de producción diaria.

Definición de parámetros y variables

Sean

x_{ij} : Cantidades de venta del i -ésimo producto ($i= 1, 2, 3$) para el j -ésimo periodo ($j=1, 2, \dots, 7$). Estas son cantidades pronosticadas con el modelo más eficiente.

P_{ij} : Unidades a pedir para ser producidas del producto i en el periodo j .

I_{ij} : Unidades a guardar en inventario del producto i en el periodo j .

Cp_j : Capacidad de producción de la planta en el periodo j .

N_{ij} : Nivel máximo de inventario para el producto i en el periodo j .

CM_i : Costo unitario de mantener inventario para el i -ésimo producto, igual en todos los períodos.

CP_i : Costo unitario de producción para el i -ésimo producto, igual en todos los períodos.

Función objetivo:

$$\text{Minimizar } \sum_{i=1}^3 \sum_{j=1}^7 CM_i I_{ij} + CP_i P_{ij}$$

Sujeto a:

Restricción de inventarios para cada periodo:

$$I_{ij} = P_{ij} + I_0 - x_{ij}$$

Restricción de almacenamiento por producto:

$$\sum_{j=1}^7 I_{ij} \leq \sum_{j=1}^7 N_{ij} \quad \text{Para } i = 1, 2 \text{ y } 3.$$

Restricción de producción por período:

$$\sum_{i=1}^3 P_{ij} \leq Cp_j \quad \text{Para } j = 1, 2, \dots, 7.$$

La programación debe cumplir con la mínima demanda:

$$\sum_{j=1}^7 P_{ij} \leq \sum_{j=1}^7 LI_{ij}$$

$$x_{ij}, P_{ij}, I_{ij} > 0$$

Se ingresaron la función objetivo y restricciones escritas en extenso en el software LINGO y se estimó el modelo de revisión periódica.

5. RESULTADOS

En este capítulo se presentarán los análisis exploratorios descriptivos, realizados a las series de ventas de las tres referencias y posteriormente se mostrarán las estimaciones de la demanda mediante modelos estadísticos, las debidas pruebas de validación y criterios de comparación, los cuales permiten seleccionar el modelo más eficiente de pronóstico, usando los criterios de menor MSE y MAPE.

5.1 Análisis exploratorio

5.1.1 Datos extremos por producto

Referencia 1200

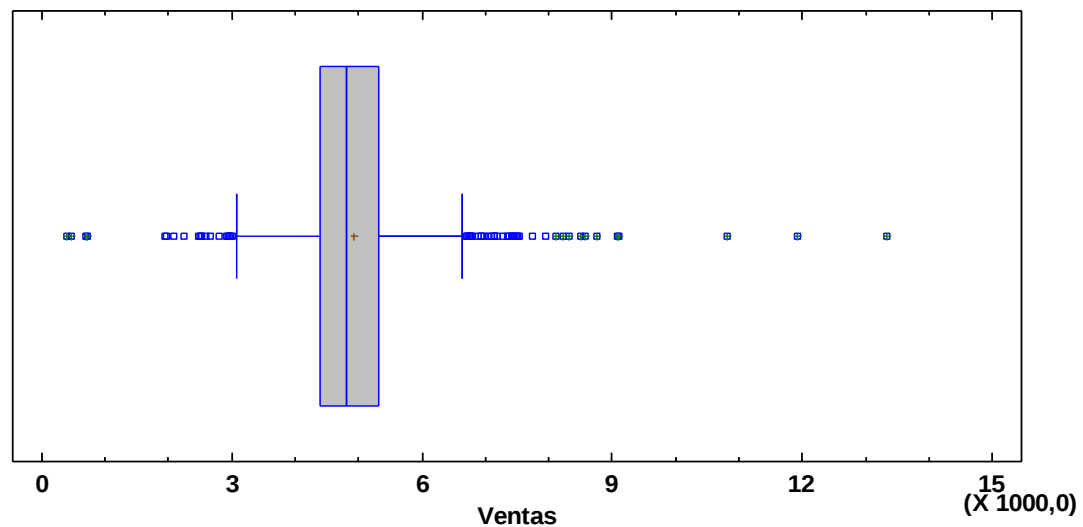


Figura 8. Caja y bigotes para la serie de datos de la referencia 1200

En la figura 8 es posible notar una pequeña asimetría a derecha, ya que el rectángulo de la izquierda es menor al rectángulo del lado derecho, y la media (cruz roja entro de los recuadros) es mayor que la mediana (línea que separa en dos los rectángulos). Lo que indica una mayor dispersión de la mediana al cuartil 3, que del cuartil 1 a la mediana. Además la presencia de un mayor número de valores atípicos por encima del límite superior, indica una alta dispersión de los datos por arriba del tercer cuartil. Esto muestra una mayor acumulación de los datos por debajo de la mediana, en un menor rango de ventas y una larga cola de datos dispersa hacia los valores altos de ventas.

Referencia 1400

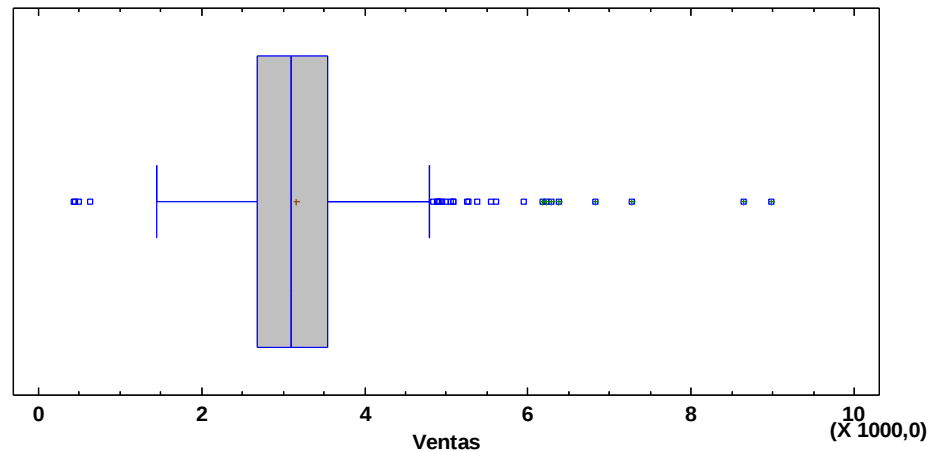


Figura 9. Caja y bigotes para la serie de datos de la referencia 1400

De acuerdo con la anterior figura se puede afirmar que la distribución de las ventas está sesgada levemente a la derecha, ya que la media es un poco mayor que la mediana. Todos los valores atípicos se encuentran por encima del límite superior, lo cual evidencia una alta dispersión de los datos por encima del tercer cuartil. Esto muestra una mayor acumulación de los datos por debajo de la mediana, en un menor rango de ventas y una larga cola de datos dispersa hacia los valores altos de ventas.

Referencia 4100

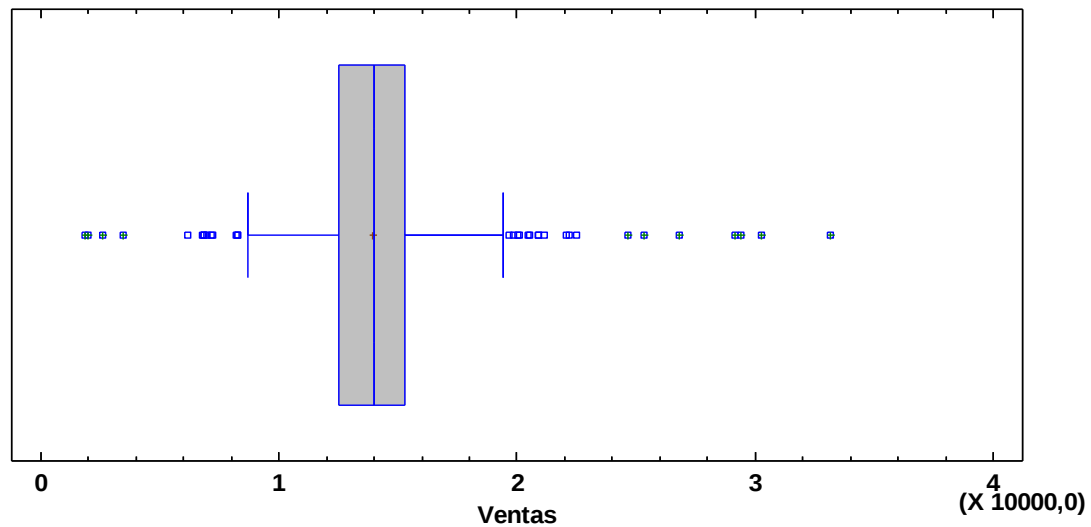


Figura 10. Caja y bigotes para la serie de datos de la referencia 4100

El diagrama es simétrico debido a que la media y la mediana son iguales, pero la mayoría de los datos atípicos se encuentran por encima del límite superior, por lo que se puede afirmar que hay mayor dispersión de los datos por arriba del tercer cuartil. Esto evidencia una acumulación equitativa de los datos por debajo y por arriba de la mediana y una larga cola de datos dispersa hacia los valores altos de ventas.

Los altos picos de facturaciones por ventas, para las tres referencias, generan ruido en la estimación de los modelos de pronóstico; por lo tanto se realiza una ponderación mediante un promedio ponderado de estos datos extremos para asegurar que la serie tenga un comportamiento más normal.

5.1.2 Análisis descriptivo de las series

Referencia 1200

Se utilizó la serie de la demanda diaria durante tres años desde el 2010 hasta 2012, que equivale a 1089 datos, por lo tanto no se muestra la gráfica de la serie completa ya que no permitiría hacer un análisis descriptivo de la referencia; en la figura 11 se observa el comportamiento para los últimos dos meses en esta referencia.

Se puede apreciar alta variabilidad de los datos y presencia de puntos extremos, representados en valores muy altos o muy bajos de ventas.

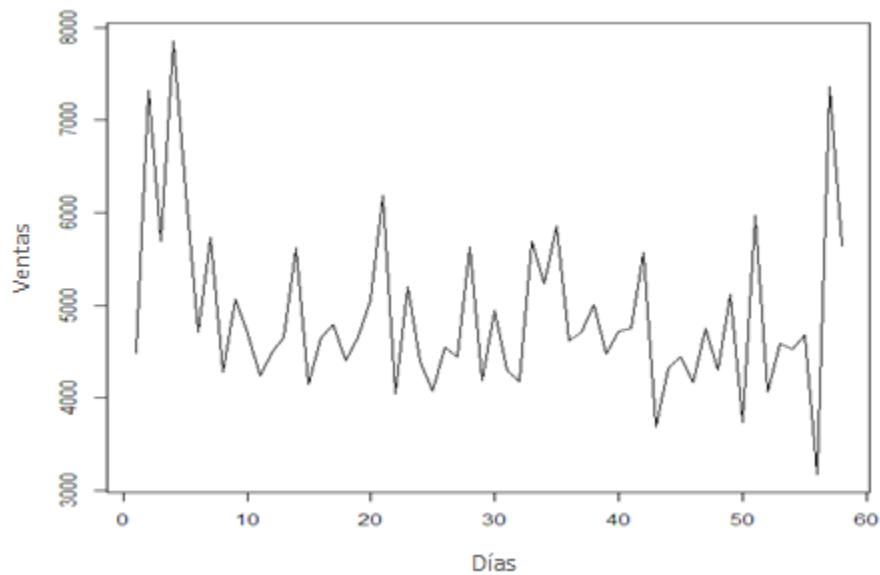


Figura 11. Salidas por facturación de la referencia 1200 en los últimos dos meses

La siguiente figura muestra una tendencia no lineal. Se explorará si una variable polinómica o componente senoidal, pueden ser incluidas en la estimación del modelo dinámico.

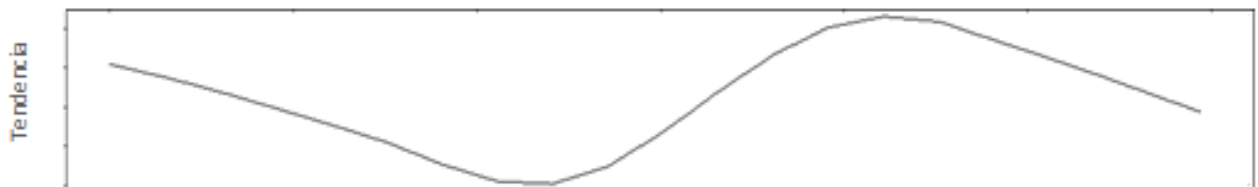


Figura 12. Comportamiento de la tendencia en la referencia 1200

La siguiente figura muestra la descomposición estacional al eliminar componente de tendencia en la serie.

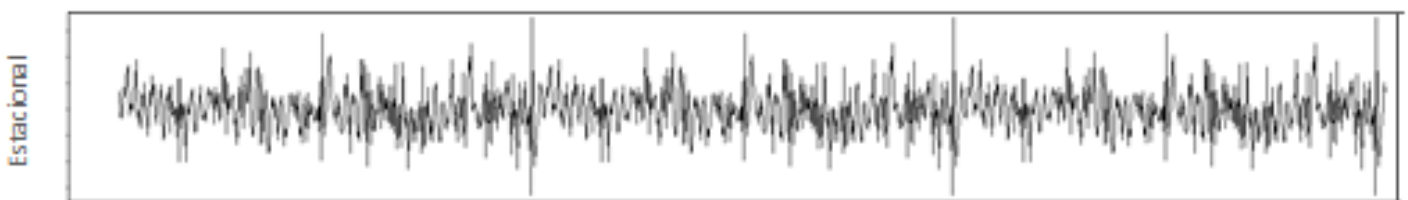


Figura 13. Estacionalidad en la referencia 1200

Como se visualiza en la figura 13, debido a la cantidad de datos de la serie, no es fácil de visualizar un patrón estacional en la serie en esta gráfica, sólo se aprecia una variabilidad alta. Por lo tanto, se recurre a la ayuda de tablas dinámicas para encontrar el valor de ventas promedio que se presenta para cada día de la semana, las cuales se aprecian en la siguiente figura:

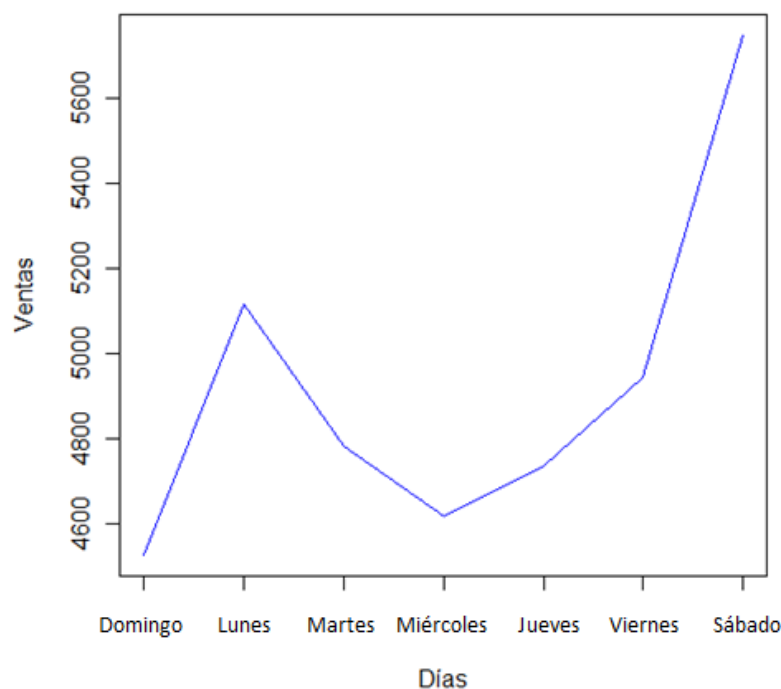


Figura 14. Facturación promedio según el día de la semana

En la figura se evidencia el aumento de la facturación por ventas a finales de la semana específicamente después del día miércoles hasta el sábado y valores muy bajos en los días domingo y martes, entre estos dos días se observa un alto pico de facturación para el lunes.

También se analizan las figuras de la función de autocorrelación (ACF) y la función de autocorrelación parcial (PACF) para determinar comportamiento correlacionado en la serie y detección de estacionalidad posible en la serie.

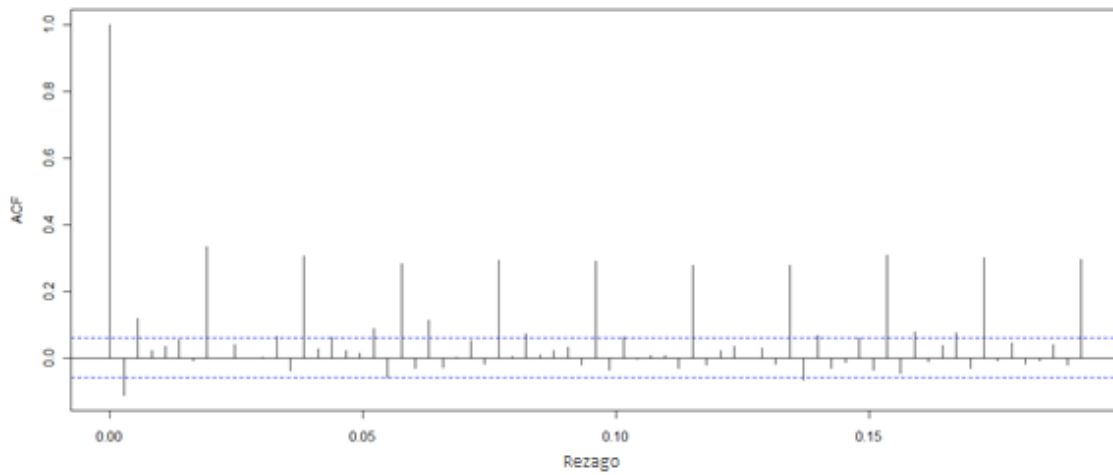


Figura 15. Función de autocorrelación, referencia 1200

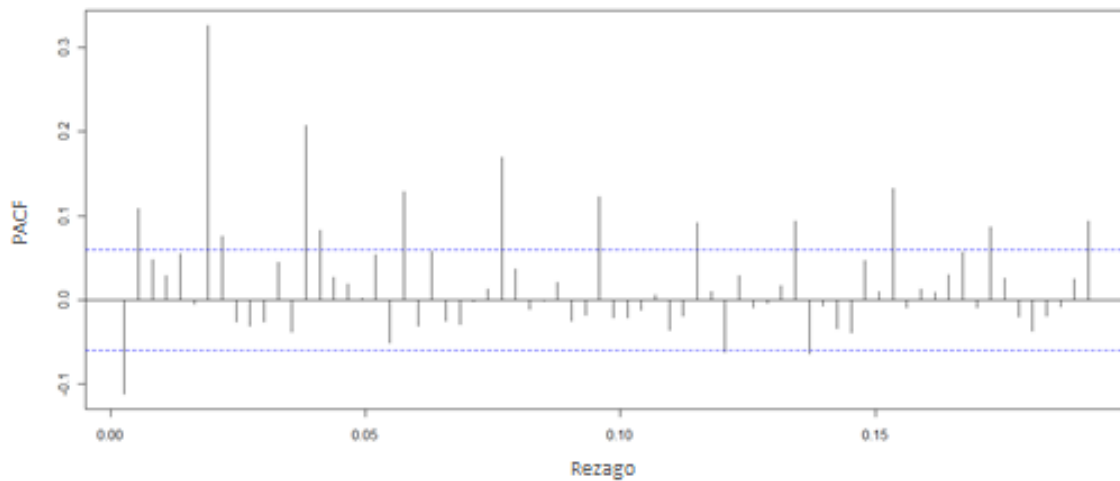


Figura 16. Función de autocorrelación parcial, referencia 1200

En las figuras 15 y 16 para la Función de autocorrelación AFC y PACF respectivamente, obtenidas del software *R*, se observa valores de autocorrelación muy altos los cuales se encuentran por fuera de las bandas, se aprecia como el patrón se repite cada 7 rezagos, lo cual nos indica la estacionalidad en la serie para esta referencia; los rezagos de orden 1 y 2 también son significativamente diferentes de cero. La prueba de Box-Pierce con un valor p de $2,2 \times 10^{-16}$ indicando una alta autocorrelación en la serie confirmando el análisis anterior.

Referencia 1400

Similar a la anterior referencia, se muestra en la figura 17 las ventas diarias de los últimos dos meses del año 2012, debido al gran volumen de datos que presenta la serie; se puede observar una variabilidad alta en la serie, con presencia de altos picos de facturación, seguido de bajos niveles de facturación.

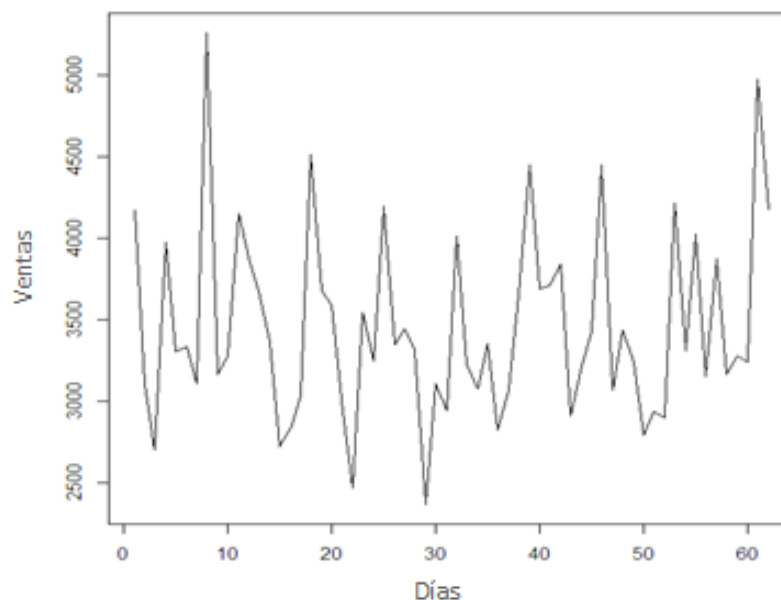


Figura 17. Salidas por facturación de la referencia 1400 en los últimos dos meses

Para determinar un posible comportamiento estacional se recurre a un gráfico de los promedios de ventas por día de la semana, que permite realizar un mejor análisis. Las facturaciones muestran un aumento acelerado hacia el final de la semana después del día miércoles hasta el día sábado, este último día y los días lunes tienen los picos más altos de ventas.

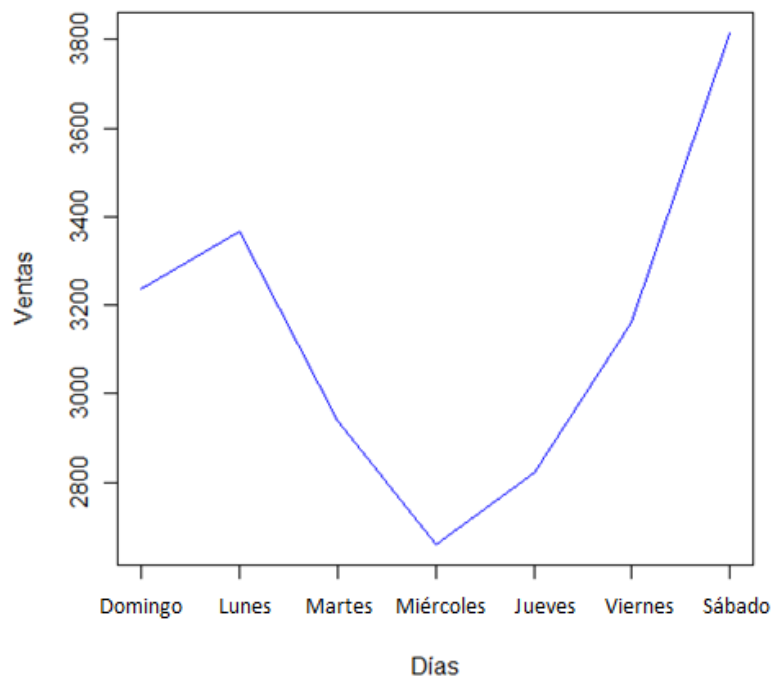


Figura 18. Facturación promedio según el día de la semana, referencia 1400

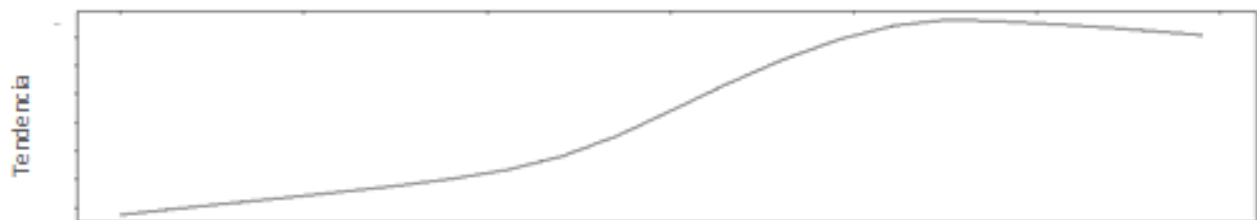


Figura 19. Comportamiento de la tendencia en la referencia 1200

En la figura anterior es posible evidenciar una cierta tendencia no lineal para la serie temporal la cual crece paulatinamente, es posible incluir una tendencia polinómica tentativamente para la estimación del modelo dinámico.

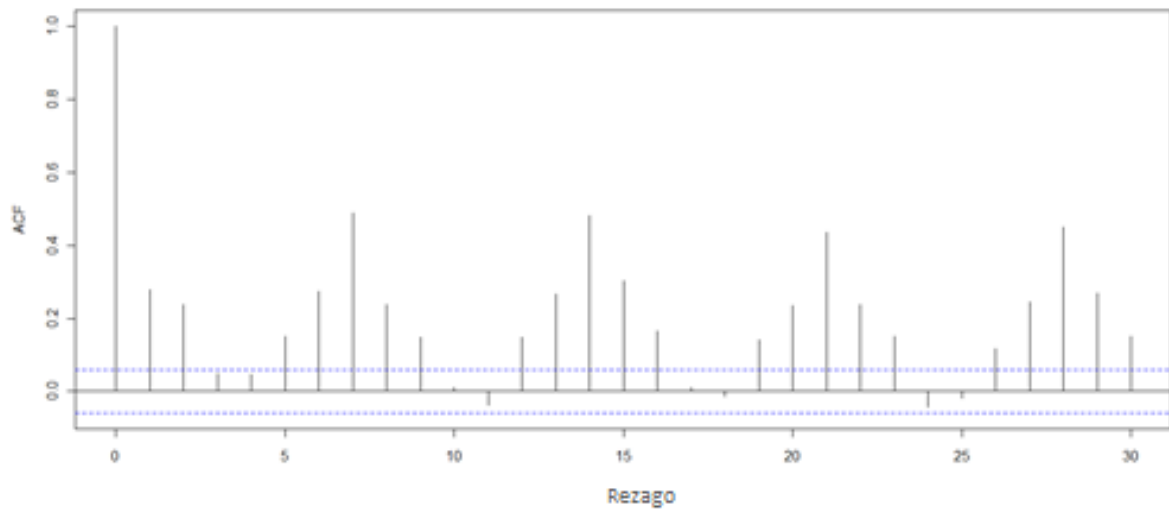


Figura 20. Función de autocorrelación, referencia 1400

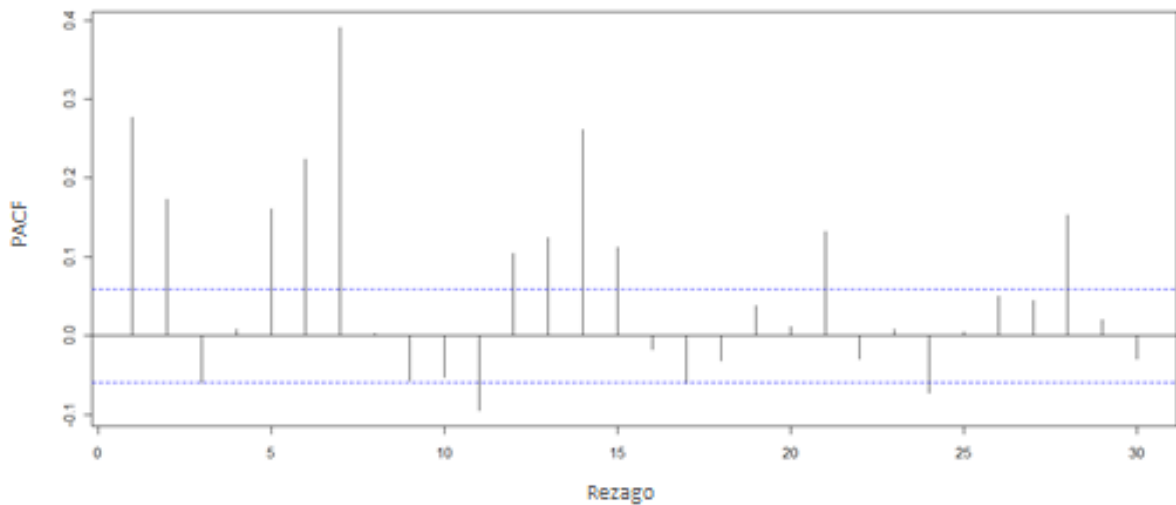


Figura 21. Función de autocorrelación parcial, referencia 1400

La prueba de Box-Pierce con un valor p de $2,2 \times 10^{-16}$ indica una alta autocorrelación en la serie, por lo que se puede afirmar que la serie no presenta estacionariedad. Las gráficas ACF y PACF evidencian lo mismo. Muestran además altos valores de autocorrelación que salen de las bandas de confianza, en especial, altos valores cada 7 rezagos, lo que indica la presencia de estacionalidad de orden 7; de estos, los rezagos 3 y 4 se encuentran por dentro de las bandas de confianza. Estas dependencias tan marcadas con el pasado de la serie pueden ser incluidas como variables endógenas rezagadas en el modelo dinámico.

Referencia 4100

Similar a las anteriores referencias, se grafican los valores de ventas diarias de los últimos 2 meses del año 2012 de la referencia 4100. Se observa que la serie presenta un comportamiento aleatorio con picos altos y bajos no tan grandes como las referencias anteriores.

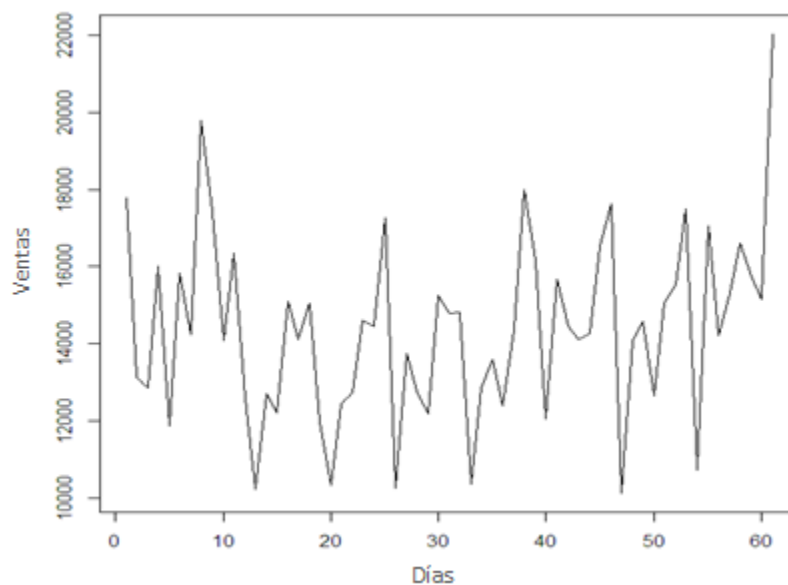


Figura 22. Salidas por facturación de la referencia 4100 en los últimos dos meses

La prueba de Box-Pierce con un valor p de $2,2 \times 10^{-16}$ indica una alta autocorrelación en la serie, por lo que se puede afirmar que la serie no presenta estacionariedad.

En la figura 23 se muestran los promedios de ventas diarias según el día de la semana, evidenciando un patrón similar a las referencias anteriores.

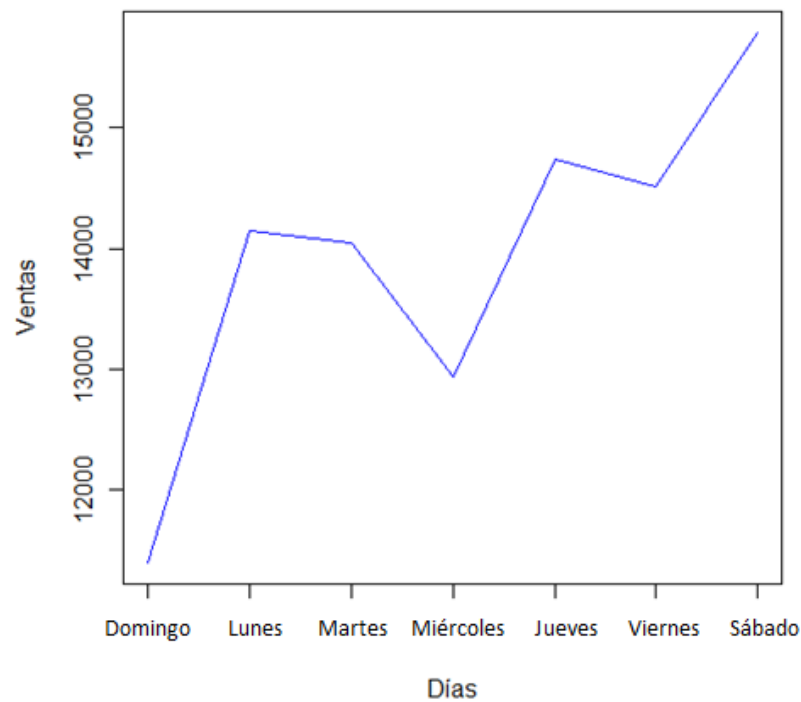


Figura 23. Facturación promedio según el día de la semana, referencia 4100

El comportamiento de las salidas por facturación promedio crecen al final de la semana, desde el día miércoles hasta el día sábado. El domingo presenta un nivel bajo de facturación al igual que el día miércoles presenta una pequeña disminución en el nivel de facturación.

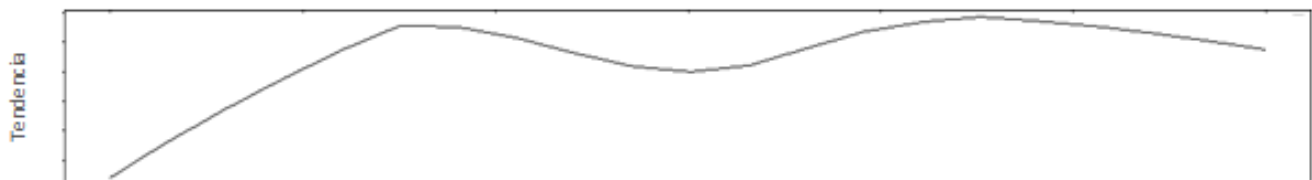


Figura 24. Comportamiento de la tendencia, referencia 1400

La figura de tendencia, muestra un crecimiento lineal sólo al principio de la serie, posteriormente es no lineal. En los primeros registros de la facturación por ventas, con el tiempo, se evidencia un cambio de comportamiento de la serie posiblemente senoidal.

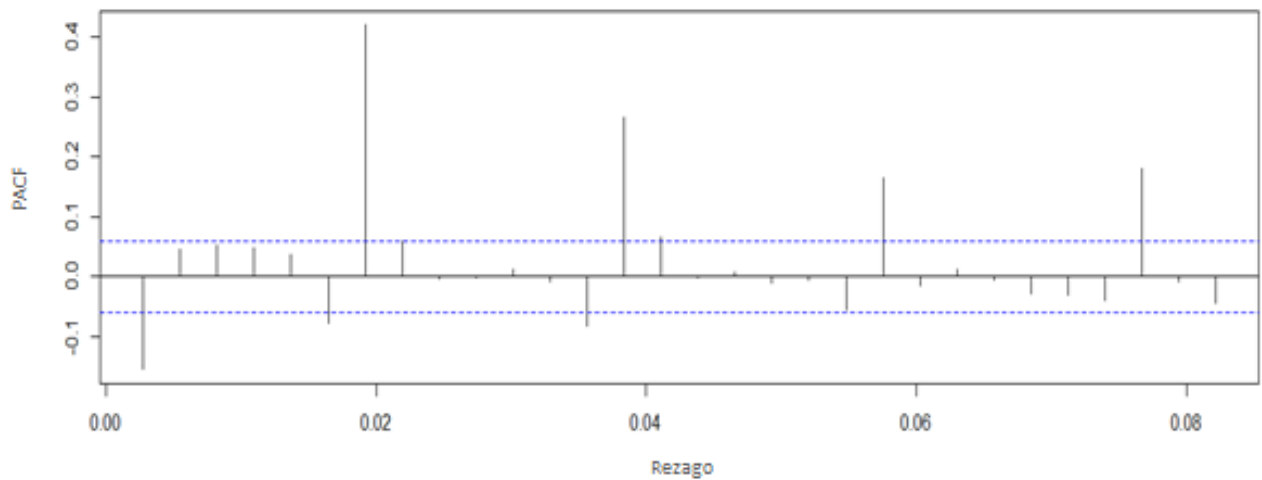


Figura 25. Función de autocorrelación, referencia 1400

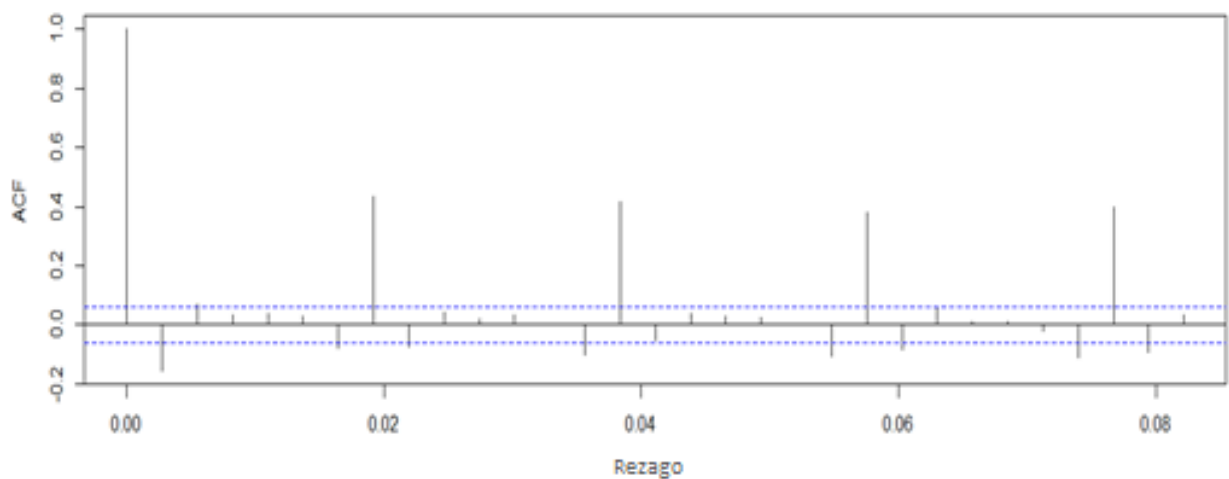


Figura 26. Función de autocorrelación parcial, referencia 1400

La fuerte autocorrelación de la prueba de Box Pierce y las anteriores gráficas de la ACF y la PACF, muestra los fuertes ciclos que se repite cada 7 periodos. Las dependencias tan marcadas con el pasado de la serie, especialmente en los rezagos 1, 6 y 7 pueden ser incluidas como variables endógenas en el modelo dinámico al encontrarse por fuera de las bandas de seguridad.

5.2 Modelos de pronósticos

En esta sección se muestra la estimación de modelos clásicos para las tres series de tiempo de ventas: Suavización exponencial simple y Holt, Box-Jenkins, Dinámicos y de RNAs, con el análisis respectivo, validación de supuestos e indicadores de error. Se calculó MAPE de ajuste con los datos usados para

estimar el modelo, y MAPE de pronóstico, para 7 días siguientes de la serie, indicador más adecuado para la elección del mejor modelo.

5.2.1 Modelos de suavizamiento exponencial

Referencia 1200

- Modelo de suavizamiento exponencial simple

El modelo es estimado a través de una minimización del error SSE, para encontrar la constante de suavizamiento, con la ayuda del software *StatGrapics*. Dicha constante es $\alpha (\alpha) = 0.0267$, con base en ella y los demás términos de la ecuación descrita en el marco teórico, se obtienen los valores ajustados (ver figura 27)

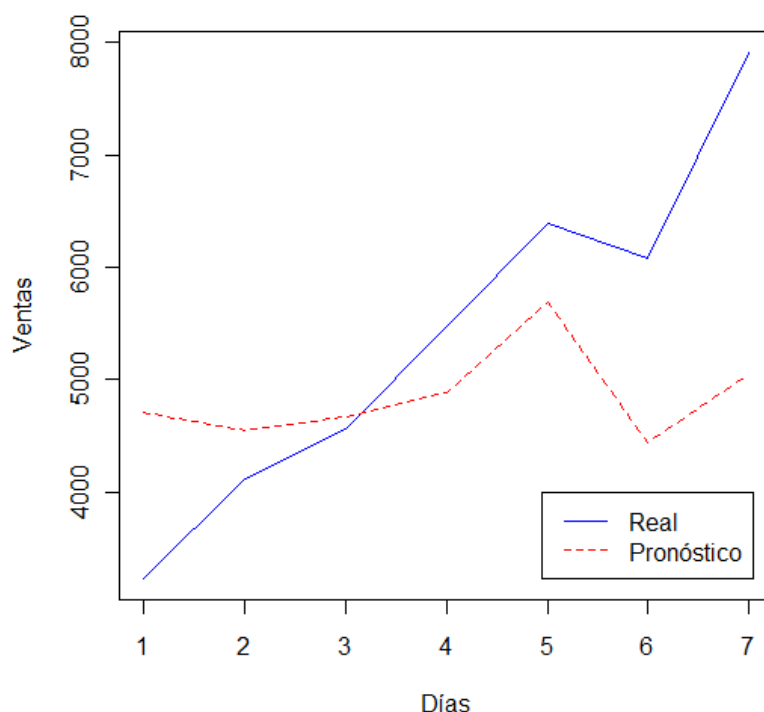


Figura 27. Demanda real vs pronóstico con suavización exponencial simple, referencia 1200

Como se observa en la figura 27 el ajuste (línea punteada) no se ajusta adecuadamente al comportamiento demanda real (línea continua) para la semana siguiente pronosticada.

El valor de MAPE de ajuste aproximado de 9.65% y un MSE de 433014.26, pero en la tabla 2 se muestra el MAPE de pronóstico, con 7 días de pronóstico, el cual es el más importante a la hora de evaluar la efectividad del modelo:

Tabla 2. Demanda real vs pronóstico con suavización exponencial simple, referencia 1200

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	3233	4712.84	45.8%	20.47%
26/12/2012	4120	4547.83	10.4%	
27/12/2012	4561.35	4662.68	2.2%	
28/12/2012	5473.62	4883.99	10.8%	
29/12/2012	6385.89	5685.28	11.0%	
30/12/2012	6081.8	4442.34	27.0%	
31/12/2012	7906.34	5042.32	36.2%	

Fuente: Elaboración propia

Validación de los residuales

La prueba de incorrelación como parte de la validación de los residuales del modelo, se obtiene con la ayuda de *StatGraphics*, en los primeros 24 rezagos un valor p equivalente a 1.34×10^{-5} , el cual es menor al nivel de significancia del 5%, permite rechazar la hipótesis sobre incorrelación, por lo tanto los residuales se encuentran correlacionados.

La prueba de normalidad de Jarque-Bera arroja un valor $p < 2.2 \times 10^{-16}$, por lo tanto con una significancia del 5% se concluye que los residuales del modelo no provienen de una distribución normal

Debido a las pruebas de validación de los residuales y al alto MAPE de pronóstico, este modelo no es confiable para realizar pronósticos.

- Suavización exponencial Holt

Este modelo se caracteriza por tener dos parámetros (de nivel y de tendencia) que posibilitan adecuarse con mayor facilidad a los pronósticos de una serie, en especial cuando poseen tendencia, con el software *StatGraphics* es posible hallar la constante de suavizamiento para el nivel (alfa) y la constante de suavizamiento para la tendencia (beta), con las cuales se minimice el error MSE.

Para esta referencia se tiene un valor alfa= 0.052 y un beta = 0.0056.

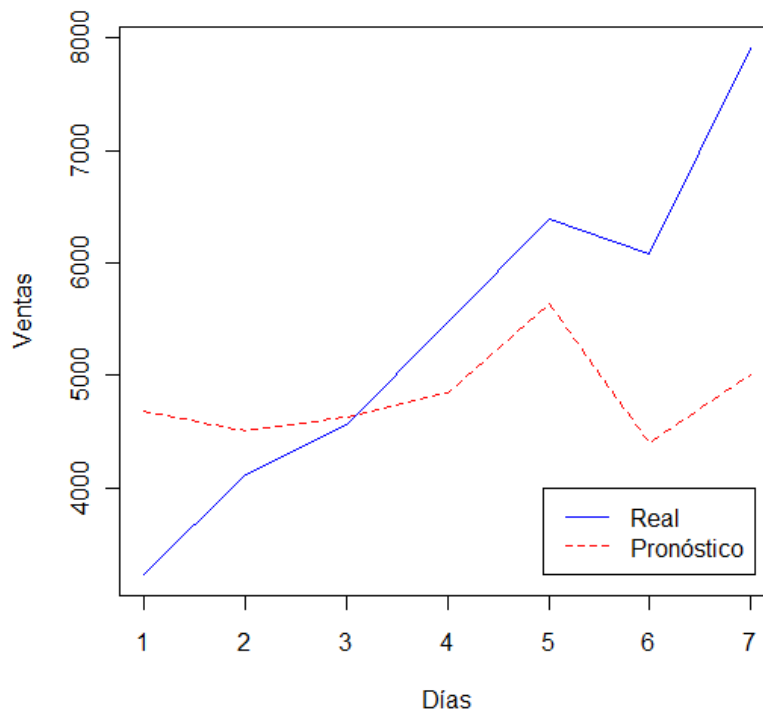


Figura 28. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1200

La figura 28 se muestra un ajuste del pronóstico (línea punteada) muy similar al modelo de suavización exponencial simple, de igual forma, esta serie no captura de forma correcta el comportamiento que presenta la serie real (línea continua).

El porcentaje de error porcentual absoluto medio para el ajuste es de 9.66% y un MSE de 435022.12, pero en base a los pronósticos hallados para la semana del 25 al 31 de diciembre, se tiene un MAPE de pronóstico del 20.45% tal como se muestran en la siguiente tabla:

Tabla 3. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1200

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	3233	4675.58	44.62%	20.45%
26/12/2012	4120	4511.07	9.49%	
27/12/2012	4561.35	4624.16	1.38%	
28/12/2012	5473.62	4842.76	11.53%	
29/12/2012	6385.89	5636.28	11.74%	
30/12/2012	6081.8	4403.26	27.60%	
31/12/2012	7906.34	4997.06	36.80%	

Fuente: Elaboración propia

Validación de los residuales

Para los 24 primeros rezagos, se tiene un valor p de 1.077×10^{-6} , el cual no es mayor al nivel de significancia, por lo tanto no cumple con la prueba de Box-Pierce con una significancia del 5%, corroborando que los residuales se encuentran correlacionados.

La prueba de Jarque-Bera estima un valor $p < 2.2 \times 10^{-16}$, por lo tanto no hay criterios para aceptar la hipótesis nula, los residuales del modelo no siguen una distribución normal.

Con un MAPE de pronóstico tan alto y las pruebas de validación inaceptables, el modelo de suavización exponencial de Holt no es un modelo confiable para la estimación de datos en el futuro, al generarse posibles sesgos en las predicciones.

Se concluye de lo anterior que los modelos de suavización exponencial no son buenos para el pronóstico de la demanda en la referencia 1200, tal como se muestra en las tablas 2 y 3, especialmente en los días 25 y 31 de diciembre donde más se encuentran alejados el pronóstico de los datos reales.

Referencia 1400

- Modelo de suavizamiento exponencial simple

Con una constante de suavizamiento exponencial $\alpha = 0.0677$ se realiza el pronóstico (línea punteada) para la referencia de producto terminado 1400, en la cual es posible observar la carencia del modelo para seguir el comportamiento real de la serie temporal (línea continua).

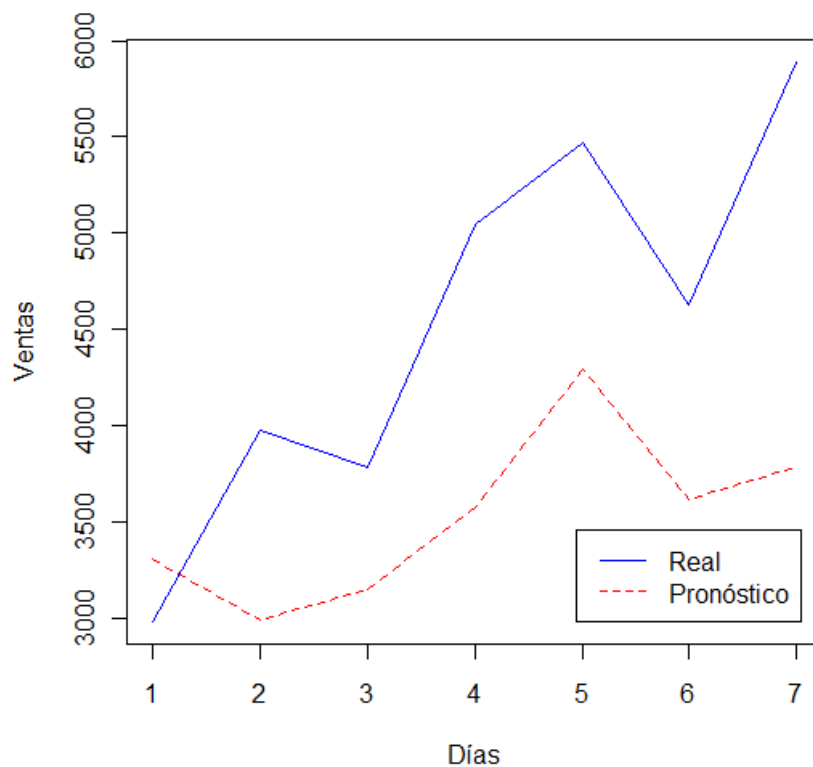


Figura 29. Demanda real vs pronóstico con suavización exponencial simple, referencia 1400

La estimación del modelo genera un MAPE de ajuste aproximado del 11.65% y un MSE de 242877.19. Los pronósticos del modelo se encuentran en la siguiente tabla:

Tabla 4. Demanda real vs pronóstico con suavización exponencial simple, referencia 1400

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	2984.25	3303.67	10.7%	22.92%
26/12/2012	3979	2985.8	25.0%	
27/12/2012	3780.05	3151.52	16.6%	
28/12/2012	5045.52	3575.57	29.1%	
29/12/2012	5465.98	4291.85	21.5%	
30/12/2012	4625.06	3617.99	21.8%	
31/12/2012	5886.44	3780.81	35.8%	

Fuente: Elaboración propia

Analizando cuantitativamente el error porcentual entre los valores reales de la serie y las observaciones calculadas con el modelo, se obtiene para la semana del 25 al 31 de diciembre un MAPE de pronóstico del 22.92%, el cual excede alrededor de 10 puntos respecto al MAPE de ajuste, corroborando la poca eficiencia del modelo para capturar el comportamiento de la serie de tiempo.

Validación de los residuales

El estadístico de prueba que se realiza sobre los residuales, arroja para los primeros 24 rezagos un valor p equivalente a 1.32×10^{-8} , el cual al ser menor del nivel de significancia (5%) corrobora que hay correlación en los residuales del modelo.

A su vez la prueba de normalidad de Jarque-Bera arroja un valor $p < 2.2 \times 10^{-16}$, lo que permite concluir que los residuales no provienen de una distribución normal.

Basados en el MAPE y las pruebas de validación sobre los residuales, se concluye que el modelo no es confiable para determinar los niveles de demanda para la referencia 1400.

- Suavización exponencial Holt

La figura 30 presenta un comportamiento similar al modelo de suavizamiento exponencial simple, en este caso se realiza una doble suavización de $\alpha = 0.1047$ y $\beta = 0.0051$ como constantes para el suavizamiento del nivel y la tendencia respectivamente.

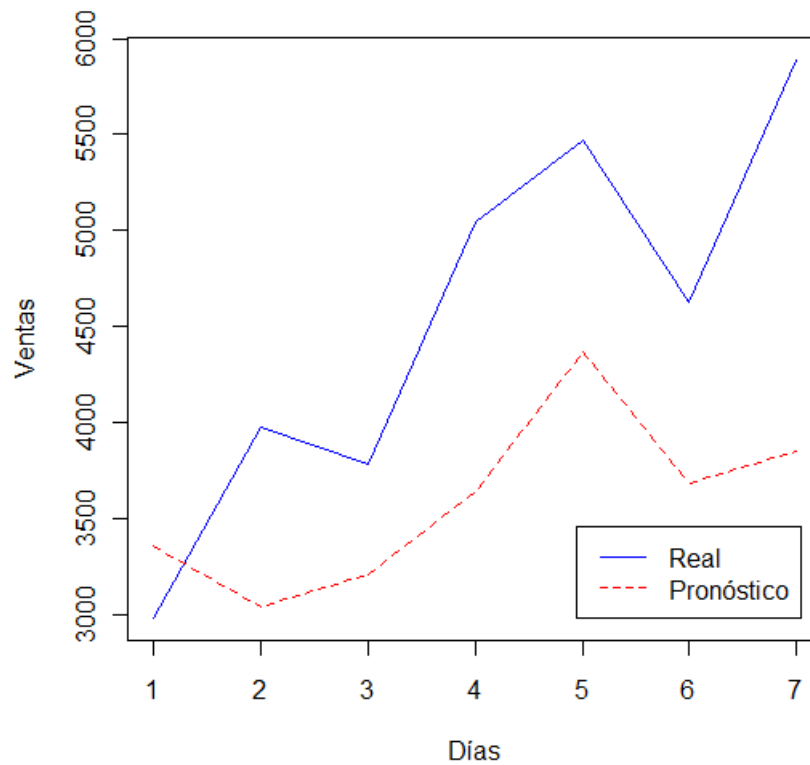


Figura 30. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1400

Con este modelo se logra un MAPE de ajuste aproximadamente de 11.66% y un MSE de 245292.02; con la ecuación resultante se realiza el pronóstico para los siete datos siguientes, los cuales se resumen en la Tabla 5

Tabla 5. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1400

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	2984.25	3359.48	12.57%	22.04%
26/12/2012	3979	3037.07	23.67%	
27/12/2012	3780.05	3206.52	15.17%	
28/12/2012	5045.52	3638.98	27.88%	
29/12/2012	5465.98	4369.16	20.07%	
30/12/2012	4625.06	3684.18	20.34%	
31/12/2012	5886.44	3851.04	34.58%	

Fuente: Elaboración propia

El MAPE de pronóstico es del 22.04%, mostrando un ajuste inadecuado, si se desea realizar un pronóstico para la referencia 1400.

Validación de los residuales

En los 24 primeros rezagos se obtuvo un valor p de 1.115×10^{-9} , el cual nos asegura que hay correlación en los residuales y la prueba de normalidad Jarque-Bera calcula un valor $p < 2,2 \times 10^{-16}$; por lo tanto el modelo no es confiable.

Con base al MAPE de pronóstico y las pruebas de incorrelación de Box-Pierce, se concluye que los modelos de suavización exponencial simple y Holt no son adecuados para el pronóstico de la demanda en la referencia 1400, tal como se muestra en las tablas 4 y 5, especialmente en los días 25 y 31 de diciembre donde más se encuentran alejados el pronóstico de los datos reales.

Referencia 4100

- Modelo de suavizamiento exponencial simple

La constante de suavizamiento en la cual la serie temporal tiene el menor error cuadrático medio MSE, después de realizar varias iteraciones, corresponde a un *alfa* de 0.0432, un MAPE de ajuste del 9.24% y un MSE de 3067585.64.

En la siguiente figura se muestran los pronósticos (línea punteada) versus los datos reales (línea continua) para el periodo comprendido entre el 25 y el 31 de diciembre del 2012; se evidencia como el modelo no captura de forma eficiente el comportamiento de la serie estacional.

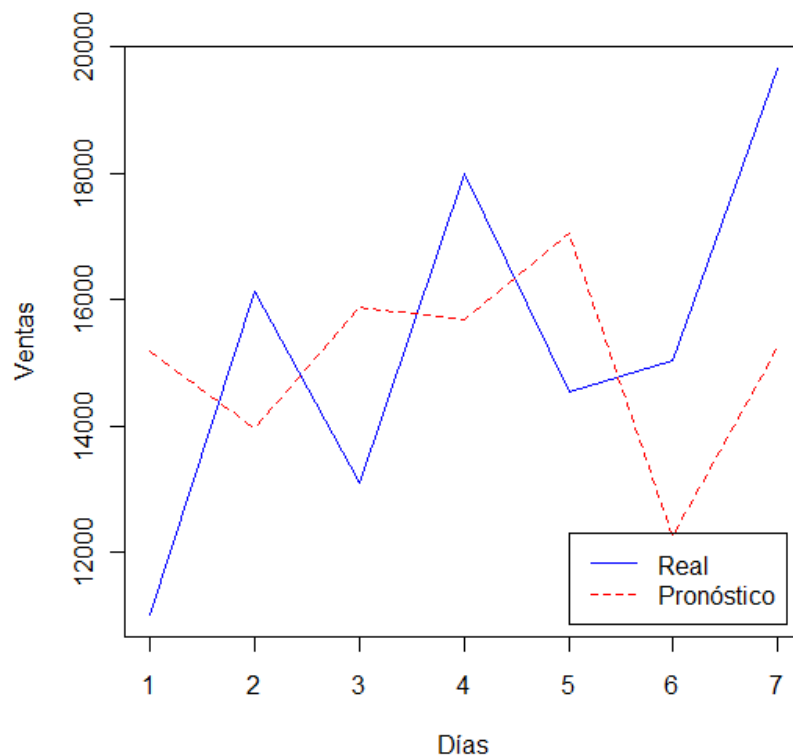


Figura 31. Demanda real vs pronóstico con suavización exponencial simple, referencia 4100

En la siguiente tabla se muestran los valores puntuales del pronóstico de la demanda para cada día, el cual arroja un MAPE de pronóstico correspondiente al 20.53%.

Tabla 6. Demanda real vs pronóstico con suavización exponencial simple, referencia 4100

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	11017.4175	15177.4	37.8%	20.53%
26/12/2012	16123.05	13961.3	13.4%	
27/12/2012	13104.815	15881.8	21.2%	
28/12/2012	18004.0725	15684.2	12.9%	
29/12/2012	14532.2424	17061	17.4%	
30/12/2012	15048.18	12243.2	18.6%	
31/12/2012	19657	15246.4	22.4%	

Fuente: Elaboración propia

Validación de los residuales

Para los 24 primeros rezagos, el estadístico de prueba que se realiza sobre los residuales calcula un valor p equivalente a 0.000431, el cual al ser menor que el grado de significancia (5%) corrobora que hay correlación en los residuales del modelo. En la prueba Jarque-Bera, se estima un valor $p < 2,2 \times 10^{-16}$

- Suavización exponencial Holt

Las constantes de suavizamiento que posee este método corresponden a un $\alpha=0.058$ y un β equivalente a 0.0054, como constante para el nivel y la tendencia respectivamente.

Con ayuda del programa *StatGraphics* se calcula un MAPE de ajuste aproximadamente del 9.30% y un MSE de 3086120.82; en la siguiente tabla se muestran los pronósticos para la semana siguiente de la serie original con el respectivo error relativo, y el MAPE de pronóstico.

Tabla 7. Demanda real vs pronóstico con suavización exponencial Holt, referencia 4100

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	11017.4175	15508.2	40.76%	20.63%
26/12/2012	16123.05	14272.1	11.48%	
27/12/2012	13104.815	16242.9	23.95%	
28/12/2012	18004.0725	16048.2	10.86%	
29/12/2012	14532.2424	17465	20.18%	
30/12/2012	15048.18	12538.8	16.68%	
31/12/2012	19657	15621.7	20.53%	

Fuente: Elaboración propia

En la figura 32 se muestra el valor pronosticado (línea) punteada contra la serie real (línea continua), se observa como el modelo no tiene la capacidad de capturar el comportamiento de la serie a través del tiempo.

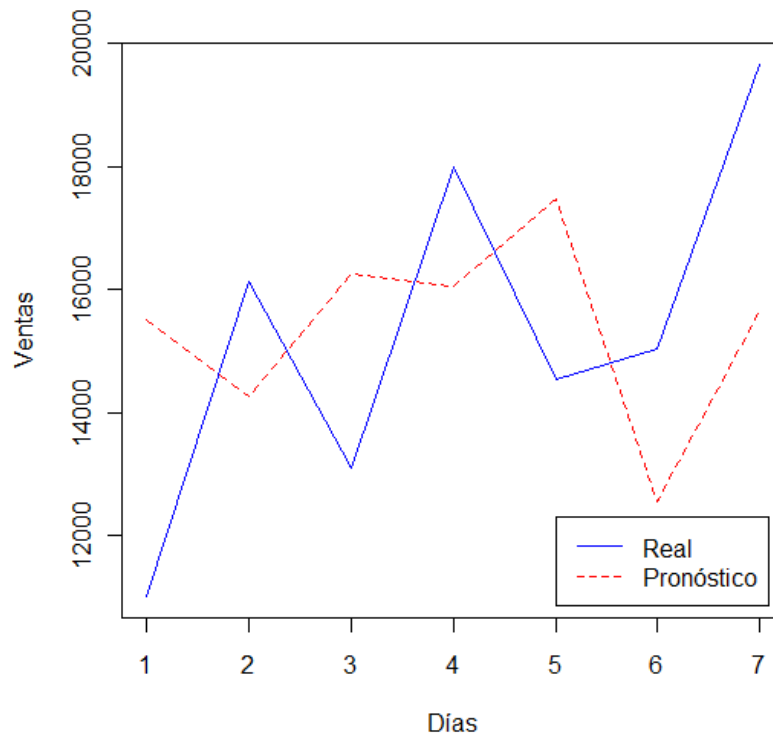


Figura 32. Demanda real vs pronóstico con suavización exponencial Holt, referencia 1400

Prueba de Incorrelación de Box-Pierce

Para los primeros 24 rezagos se obtuvo un valor p equivalente a 0.000159 para la prueba de Box-Pierce, este valor al ser menor el grado de significancia, se afirma que los residuales están auto-correlacionados, lo cual podría generar sesgos en las predicciones. La prueba de Jarque-Bera arroja un valor $p < 2,2 \times 10^{-16}$, por lo tanto los residuales no provienen de una distribución normal.

El MAPE de pronóstico tan alto (alrededor del 20%) muestra que los modelos de suavización exponencial simple y Holt no son eficientes al momento de realizar predicciones de demanda para la referencia 4100, tal como se muestra en las tablas 6 y 7, especialmente en los días 25 y 31 de diciembre donde más se encuentran alejados el pronóstico de los datos reales.

5.2.2 Modelos dinámicos

Para cada referencia se hicieron varias estimaciones con base en el proceso descrito en el capítulo 5.1, Las pruebas de significancia ayudaron a la

incorporación o no de las variables que afectan con mayor proporción el pronóstico de la demanda.

Para corregir la correlación en los residuales, validación importante cuando se trata de una serie temporal, se ensayaron correctivos como:

- Transformaciones sobre la variable respuesta
- Incorporación de variables del pasado (variables autorregresivas - endógenas)
- Incorporar variables explicativas o de nivel (exógenas)

En especial esta última técnica, tuvo el mayor impacto en la mejora de la estimación del modelo dinámico final.

Debido a la detección de patrones estacionales encontrados, así como el juicio de quienes trabajan en la planeación de la producción de la empresa Auralac, afirman que los niveles en la demanda son fuertemente marcados por el día de la semana y las fechas especiales del año, se incluyeron estas dos variables exógenas: días de la semana y nivel de demanda para cada día, las cuales pueden justificar los posibles cambios estructurales en la serie.

La alta dispersión que presenta la serie de datos entre los años 2010 y 2012, permitió establecer 6 niveles de demanda en las tres referencias, tomando como punto de división los percentiles, de esta forma se logra la agrupación de los datos de forma estratificada: Los niveles 1 y 6 para los datos más extremos, los niveles 2 y 3 cuando los datos se encuentran por debajo de la media y los niveles 4 y 5 cuando se encuentran por encima de la media; tal como se muestra en la tabla número 8.

Tabla 8. Asignación de niveles según valor del percentil

NIVEL	PERCENTIL
Nivel 1	0 - 5
Nivel 2	5 - 25
Nivel 3	25 - 50
Nivel 4	50- 75
Nivel 5	75 - 95
Nivel 6	95 -100

Fuente: Elaboración propia

La inclusión de estas variables (especialmente el nivel) fue de gran ayuda para la disminución del MAPE y corregir la correlación en los residuales.

Para mostrar el acierto de incluir el nivel en un modelo dinámico se tomará la referencia 1400 como base, y el procedimiento realizado fue el mismo para las referencias restantes.

En el análisis descriptivo de la serie 1400 se tiene variables tentativas como:

- Tendencia polinómica
- Componente senoidal
- Variables autorregresivas de orden: 1, 2, 6, 7

Usando un método de selección BACKWARD¹ y haciendo transformaciones sobre la variable respuesta (raíz cuadrada, raíz cúbica, logaritmo), de todas las variables tentativas, sólo se incluyen en el modelo las variables: tiempo, sinoidal con periodo 3, rezago 6, rezago 7, sin intercepto y con transformación logarítmica, bajo un grado de significancia del 0.05, tal como lo muestra la siguiente tabla ANOVA tipo III:

Tabla 9. ANOVA Tipo III, referencia 1400

	Suma de cuadrados	Grados de Libertad	Valor F	Pr(>F)
Tiempo	5.806	1	184.292	< 2.2e-16
Sin3	1.507	1	47.837	7.92E-12
L(log(serie),6)	0.477	1	15.14	0.0001059
L(log(serie),7)	6.132	1	194.656	< 2.2e-16
Residuales	33.959	1078		

Fuente: Programa estadístico R

En esta tabla, todas las variables finales son significativas.

El programa R calcula un MAPE de ajuste alrededor del 13.68%.

¹ Metodología de selección en la cual se incluyen todas las variables de la ecuación y después se elimina la variable que tenga una significación mayor. Si una variable cumple con los requisitos de eliminación, se recalcula el modelo sin dicha variable. El proceso se detiene cuando todas las variables tengan un nivel de significancia menor al estipulado. (Álvarez, 1994)

La ecuación para el modelo es:

$$Z_t = 0.002 * T + 0.0566 * \text{Sen}\left(\frac{3\pi * t}{7}\right) + 0.1086 * Z_{t-6} + 0.389 * Z_{t-7} + \varepsilon$$

Donde $Z_t = \ln(y_t)$

Haciendo pronósticos puntuales para la semana del 25 al 31 de diciembre del 2012 se obtiene:

Tabla 10. Demanda real vs pronóstico con modelo dinámico tentativo, referencia 4100

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	2984.25	3322.11157	11.32%	22.85%
26/12/2012	3979	3522.97054	11.46%	
27/12/2012	3780.05	3267.76873	13.55%	
28/12/2012	5045.52	3309.25562	34.41%	
29/12/2012	5465.98	3449.32304	36.89%	
30/12/2012	4625.06	3999.29503	13.53%	
31/12/2012	5886.44	3602.22288	38.80%	

Fuente: Elaboración propia

El modelo arroja un valor p equivalente a 7.837e-12, por lo tanto los residuales no provienen de una distribución normal y para la prueba de Box Pierce se obtiene un valor p de < 2.2e-16 entonces los residuales están auto-correlacionados.

Teniendo un MAPE de pronóstico tan alto y ninguna prueba de validación sobre los residuales es aceptable, se toma la decisión de incorporar variables exógenas al modelo: día de la semana y los niveles.

Al realizar el pronóstico para la semana siguiente no se tiene certeza qué nivel de demanda pueda tener, por ello su asignación será a través de la opinión de expertos y con base en los datos obtenidos en los años anteriores.

Usando tablas dinámicas se filtra la serie para los últimos tres meses del año 2012, se agrupan por la cantidad de niveles de demanda existente en cada día de la semana, los cuales se muestran en la siguiente tabla:

Tabla 11. Niveles estimados para las referencias según el día a pronosticar

	REFERENCIA 1200						REFERENCIA 1400						REFERENCIA 4100					
DÍA/NIVEL	1	2	3	4	5	6	1	2	3	4	5	6	1	2	3	4	5	6
Domingo	2		1	3	5	1			1	1	4	6	3		1	7		1
Lunes		1	3	2	3	4				1	5	7	1	2	1	1	3	5
Martes	4		3	2	3		2	4	5		1			5	4	2	1	
Miércoles		5	4		1	2		2	4		5	1		3	1	1	5	2
Jueves			4	3	3	1				5	7				7	2	3	
Viernes		3	2	3	4					1	5	6			1	1	3	7
Sábado	1		1	2	2	4				1	4	7				6	4	2

Fuente: Elaboración propia

Los datos de la tabla 11 que se encuentran en negrilla, son aquellos que tienen mayor frecuencia absoluta para cada una de las referencias, según el día de la semana y el nivel correspondiente.

Estimando un nuevo modelo dinámico para la referencia 1400, se tiene la siguiente tabla ANOVA tipo III.

Tabla 12. ANOVA Tipo III final para la referencia 1400

	Suma de cuadrados	Grados de libertad	Valor F	Pr(>F)
Intercepto	75715	1	36634.6371	< 2.2e-16
Cos4	96	1	46.6522	1.42E-11
Día	87	7	5.9848	7.43E-07
Nivel	19105	5	1848.7771	< 2.2e-16
Residuales	2222	1075		

Fuente: Programa estadístico R

El modelo dinámico para la referencia 1400 es el siguiente:

$$\begin{aligned}\sqrt{Y_t} = & 44.67 + 0.44 * \cos((4\pi * t)/7) + 0.064 * \text{Lunes} - 0.041 * \text{Martes} - 0.508 \\ & * \text{Miércoles} - 0.259 * \text{Jueves} - 0.144 * \text{Viernes} + 0.665 * \text{Sábado} \\ & + 1.114 * \text{Domingo} + 5.126 * \text{Nivel2} + 9.375 * \text{Nivel3} + 12.646 \\ & * \text{Nivel4} + 17.086 * \text{Nivel5} + 23.488 * \text{Nivel6} + \varepsilon\end{aligned}$$

Empleando la ecuación anterior y asignando el nivel para cada día de la semana con base en la tabla 11 obtenemos los pronósticos para la semana del 25 al 31 de diciembre del 2012, los cuales se detallan en la siguiente tabla:

Tabla 13. Demanda real vs pronóstico con modelo dinámico final, referencia 4100

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	2984.25	2948.09816	1.2%	7.44%
26/12/2012	3979	3787.38427	4.8%	
27/12/2012	3780.05	3817.83574	1.0%	
28/12/2012	5045.52	4665.46928	7.5%	
29/12/2012	5465.98	4776.47361	12.6%	
30/12/2012	4625.06	4838.48027	4.6%	
31/12/2012	5886.44	4693.3362	20.3%	

Fuente: Elaboración propia

El modelo arroja un MAPE de pronóstico aproximadamente del 7,4%, lo cual corrobora el acierto de inclusión las dos variables exógenas logrando una disminución en el MAPE alrededor de 15 puntos.

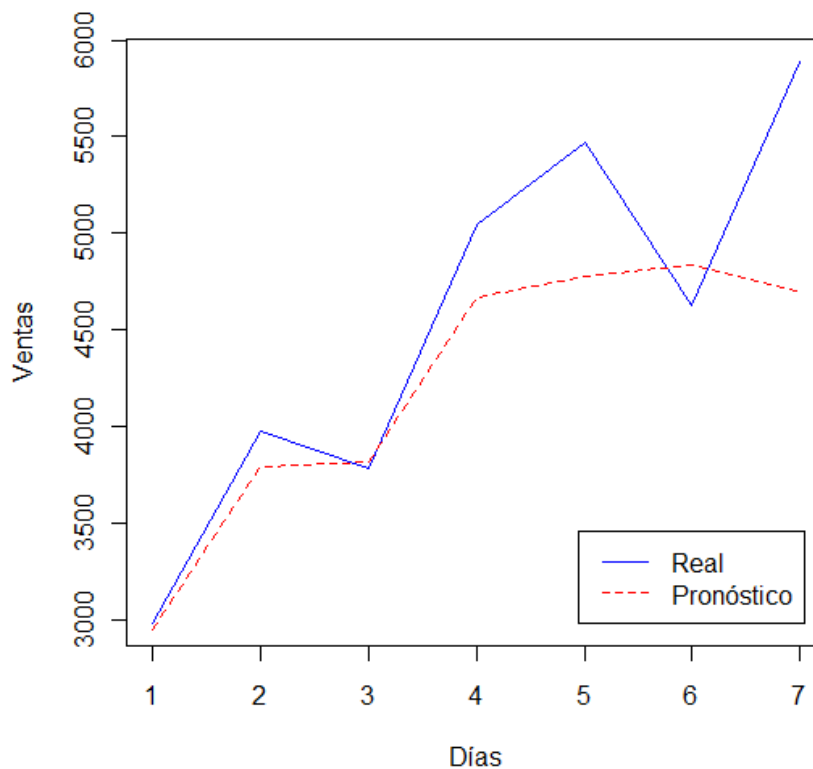


Figura 33. Demanda real vs pronóstico con modelo dinámico, referencia 1400

En la figura 33 se muestra como el modelo trata de imitar el comportamiento de la serie original (línea continua) contra los valores pronosticados (línea punteada). Se aprecia que los datos del 31 y el 29 de diciembre, son los más discrepan con la realidad, los cuales generan más peso en la ponderación del error o MAPE.

Validación de los residuales

La prueba de normalidad de Jarque-Bera determina un valor p de $< 2.2 \times 10^{-16}$, lo cual afirma que los residuales no provienen de una distribución normal, pero la prueba de Box- Pierce arroja un valor p para los primeros 48 rezagos de 0.2957, al ser mayor al nivel de significancia del 5%, se afirma que los residuales de los errores no están correlacionados.

Se concluye que el modelo es adecuado para pronosticar la demanda de la referencia 1400. Por la ausencia de normalidad en los residuales, se recomienda que el pronóstico sea realizado sólo a corto plazo.

A continuación se muestran los modelos estimados para las dos referencias restantes, por practicidad, de todos los modelos ensayados para estos productos terminados, sólo se mostrarán los modelos finales.

Referencia 1200

El modelo dinámico para esta referencia, tiene una transformación logarítmica sobre la variable respuesta, las variables que fueron incluidas se encuentran en la siguiente tabla ANOVA tipo III, las cuales se afirma que son significativas para el modelo con una significancia del 10%.

Tabla 14. ANOVA Tipo III, referencia 1400

	Suma de cuadrados	Grados de libertad	Valor F	Pr(>F)
Tiempo	0.0211	1	8.3011	0.004041
L(log(serie), 6)	0.0094	1	3.6881	0.055067
Día	0.0335	7	1.8796	0.06954
Nivel	19.4273	5	1527.6286	< 2.2e-16
Residuales	2.719	1069		

Fuente: Programa estadístico R

La tabla de coeficientes obtenida con el software R, permite generar la siguiente ecuación:

$$Z_t = 0,0052 * T + 0,021 * Z_{t-6} - 2,535 * Lunes - 2,532 * Martes - 2,535 * Miércoles - 2,5335 * Jueves - 2,539 * Viernes - 2,52 * Sábado - 2,543 * Domingo + 0,2386 * Nivel2 + 0,3282 Nivel3 + 0,4214 * Nivel4 + 0,5423 * Nivel5 + 0,716 * Nivel6$$

Donde $Z_t = \ln(y_t)$

Se estima un MAPE de ajuste para el modelo alrededor del 3.34%, y de pronóstico de 5.97%, como se observa en la Tabla 15, donde además, se ven los datos pronosticados para la referencia 1200 para la última semana del año 2012, usando el modelo dinámico anteriormente descrito se tiene los siguientes niveles de demanda:

Tabla 15. Demanda real vs pronóstico con modelo dinámico, referencia 1200

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	3233	3348.86584	0.03583849	5.97%
26/12/2012	4120	4237.04848	0.02840983	
27/12/2012	4561.35	4648.28658	0.0190594	
28/12/2012	5473.62	5675.77327	0.03693228	
29/12/2012	6385.89	7004.85004	0.0969262	
30/12/2012	6081.8	5723.12639	0.05897491	
31/12/2012	7906.34	6787.13764	0.14155758	

Fuente: Elaboración propia

La siguiente figura muestra como el modelo es capaz de capturar el comportamiento que presenta la serie temporal, entre los datos reales (línea continua) contra la serie pronosticada (línea punteada), es apreciable como en los días 29 y 31 de diciembre las líneas no se traslapan, los cuales posiblemente son quienes tienen mayor peso sobre la ponderación del error.

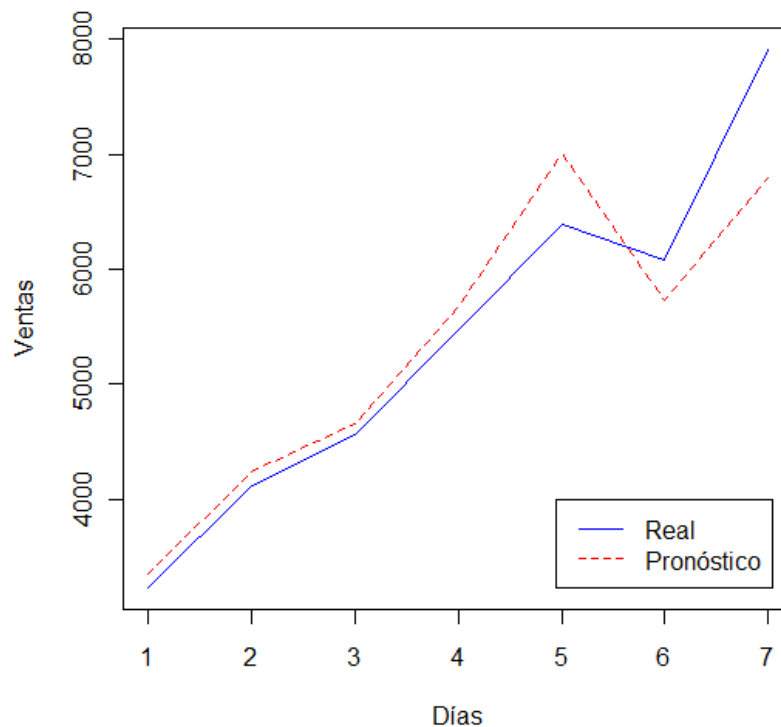


Figura 34. Demanda real vs pronóstico con modelo dinámico, referencia 1200

Validación de los residuales

La prueba de normalidad sobre los residuales, Jarque-Bera, arroja un valor p inferior a 2.2×10^{-16} , por lo tanto, se afirma que los residuales no provienen de una distribución normal.

La prueba de Box- Pierce para los 24 primeros rezagos de la serie tiene un valor p de 0.7712, al ser mayor al grado de significancia del 5%, no hay criterios para rechazar la hipótesis sobre la incorrelación de los residuales de los errores.

La prueba nos permite analizar la dependencia de orden del pasado de la serie, lo cual asegura que el modelo es eficiente para el pronóstico de la demanda para la referencia 1200, pero se recomienda realizarlo a corto plazo.

Referencia 4100

El modelo dinámico final para esta referencia incluye solamente las variables exógenas de nivel y día de la semana, las cuales cumplen con prueba de significancia, con una confianza del 95%; como se ven en la siguiente tabla ANOVA tipo III (valor p inferior al 5%).

Tabla 16. ANOVA Tipo III, referencia 4100

	Suma de cuadrados	Grados de libertad	Valor F	Pr(>F)
Intercepto	4254.4	1	1.93E+06	< 2.2e-16
Día	0	7	2.78E+00	0.007195
Nivel	16.6	5	1.51E+03	< 2.2e-16
Residuales	2.4	1076		

Fuente: Programa estadístico R

Comprobada la significancia de las variables y con ayuda de la tabla de coeficientes, se construye la siguiente ecuación para la referencia 4100:

$$\begin{aligned} \ln(y) = & \text{[} 9,13 + 0,0154 * \textit{Lunes} + 0,1529 * \textit{Martes} + 0,0171 * \textit{Miércoles} + 0,0206} \\ & * \textit{Jueves} + 0,0184 * \textit{Viernes} + 0,0237 * \textit{Sábado} + 0,097 * \textit{Domingo} \\ & + 0,2188 * \textit{Nivel2} + 0,3452 * \textit{Nivel3} + 0,4408 * \textit{Nivel4} + 0,5382} \\ & * \textit{Nivel5} + 0,697 * \textit{Nivel6} \text{]} \end{aligned}$$

Se estima que el modelo tiene un MAPE de ajuste aproximadamente del 3.45%.

Empleando la ecuación descrita anteriormente se hallan los valores puntuales pronosticados para la semana del 25 al 31 de diciembre del 2012, los cuales se encuentran en la tabla 17.

Tabla 17. Demanda real vs pronóstico con modelo dinámico, referencia 1400

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	11018.00	11666.3044	0.05884048	3.27%
26/12/2012	16124.00	16084.9929	0.00241919	
27/12/2012	13105.00	13308.8433	0.01555462	
28/12/2012	18005.00	18883.9758	0.04881843	
29/12/2012	14533.00	14689.1294	0.01074309	
30/12/2012	15049.00	15809.6217	0.050543	
31/12/2012	19657.00	18826.4674	0.04225124	

Fuente: Elaboración propia

En la figura 35 se corrobora el acierto en las variables incluidas en el modelo, las cuales permiten realizar un pronóstico (línea puntuada) que se ajusta de forma adecuada al comportamiento real de la serie original (línea continua), las pequeñas variaciones están ponderadas en el MAPE de pronóstico del 3.27%.

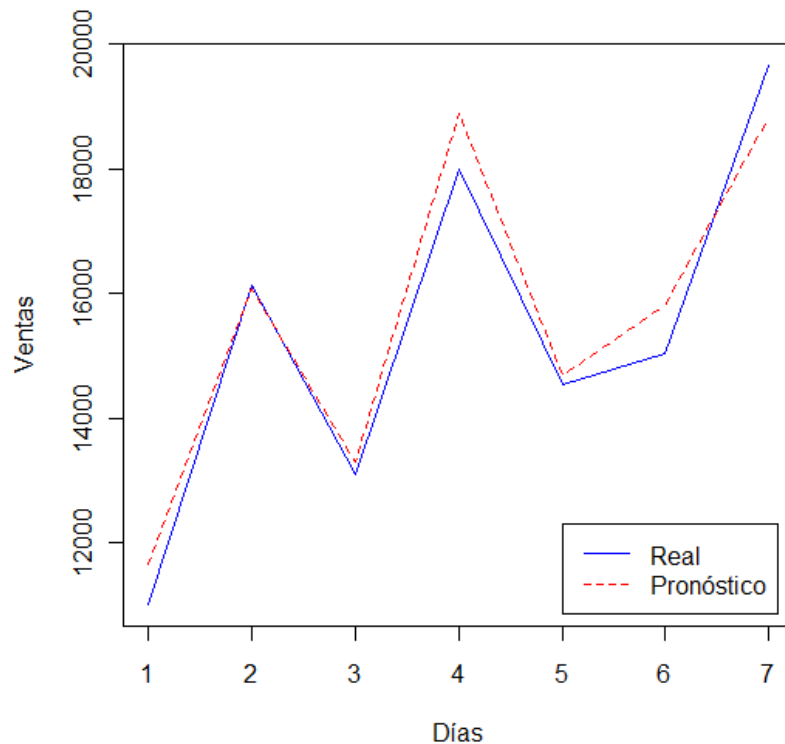


Figura 35. Demanda real vs pronóstico con modelo dinámico, referencia 4100

Validación de los residuales

La prueba de normalidad Jarque-Bera estima un valor p inferior a $2.2e-16$, al ser menor al grado de significancia (5%), se concluye que los residuales no provienen de una distribución normal, pero la prueba de Box-Pierce estima un Valor P de 0.8316, por lo tanto se afirma que los residuales no están correlacionados.

Se concluye que el modelo dinámico es acertado al momento de realizar pronósticos de demanda para la referencia 4100, sin embargo, al tener problemas de normalidad se recomienda realizar pronósticos a corto plazo solamente.

5.2.3. Modelos ARIMA

Para identificar el modelo Box Jenkins que mejor se ajuste a los datos, se utilizan las gráficas ACF y PACF, pero dada la falta de estacionariedad encontrada en el análisis descriptivo, se debe realizar primero una diferenciación en las series. La diferencia de orden 1 sería: $Z_t = Y_t - Y_{t-1}$.

Referencia 1200

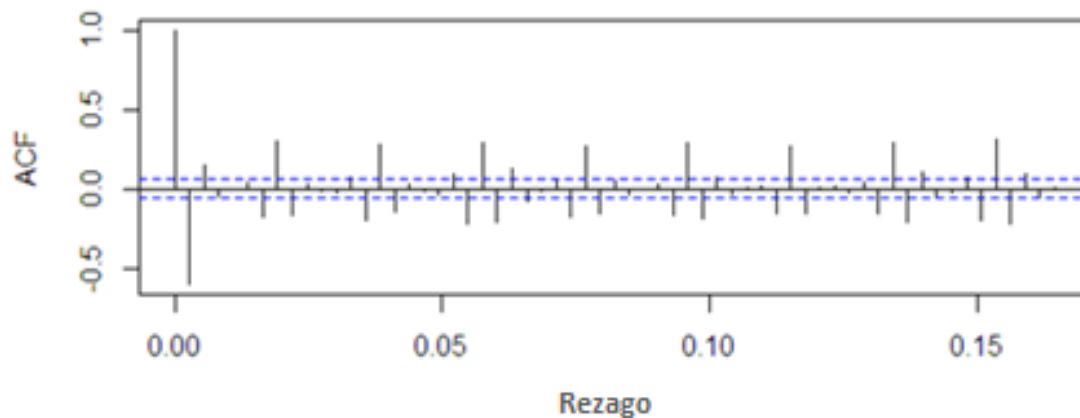


Figura 36. Función de autocorrelación diferenciada, referencia 1200

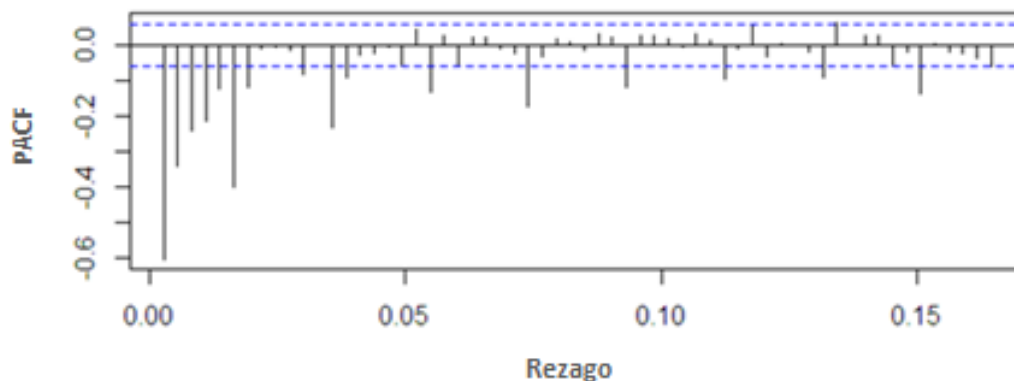


Figura 37. Función de autocorrelación parcial diferenciada, referencia 1200

En las figuras 36 y 37 los valores de auto correlación que se encuentran por fuera de las bandas de seguridad siguen un patrón de orden 7. Lo que evidencia la estacionalidad de la serie.

En la ACF los siete primeros valores decrecen rápidamente después del segundo rezago. En la PACF son significativamente diferentes de cero, especialmente el primer rezago, ya que se encuentran por fuera de las bandas. Se puede apreciar en ambas gráficas un patrón senoidal. Esto sugiere un modelo tipo ARMA (1,2) para la parte no estacional.

En la ACF y la PACF los valores de autocorrelación de orden 7 decrecen lentamente después del orden 2L (14). El comportamiento indica un modelo tipo SARMA (2,2) para la parte estacional.

El modelo tentativo producto del análisis de las gráficas ACF y PACF es un SARIMA (1, 1,2) x (2, 0, 2) [7].

La prueba de Box-Pierce arroja un valor p de 0.0190907, mayor al grado de significancia del 1%, por lo tanto no hay autocorrelación en los residuales, pero la prueba normalidad de Jarque-Bera no se cumple ya que el valor p es de 2.2×10^{-16} . El modelo determina un MAPE de ajuste del 9.89%, de pronóstico del 22.1% y un MSE de 431501.24.

La función automática de Statgraphics estima un modelo SARIMA (2, 0, 1) x (0, 1, 1) [7], con un MAPE de ajuste del 9.69% y un MSE de 417304.11.

El modelo cumple con la prueba de Box-Pierce con un grado de significancia del 5%, ya que el valor p es de 0.0510907, pero no con la prueba de normalidad debido a que el valor p es de 2.2×10^{-16} . El MAPE de pronóstico es del 20.92%, inferior al del modelo estimado anteriormente, por lo cual se elige éste último como modelo Box-Jenkins para esta referencia.

Tabla 18. Coeficientes del modelo Box-Jenkins, referencia 1200

Parámetro	Coeficientes	Error estándar	T
AR(1)	0.824657	0.0335407	24.5867
AR(2)	0.164099	0.030763	5.33431
MA(1)	0.948324	0.0171708	55.229
SMA(1)	0.975582	0.00324065	301.045

Fuente: Programa estadístico R

La siguiente es la ecuación del modelo estimado.

$$Z_t = -0.824657Z_{t-1} - 0.164099Z_{t-2} - 0.948324a_{t-1} - 0.975582a_{t-7}$$

Los pronósticos para la semana del 25 al 31 de diciembre del 2012 se muestran en la siguiente tabla:

Tabla 19. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1200

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	3233	4544.17	41%	20.92%

26/12/2012	4120	4589.92	11%
27/12/2012	4561.35	4775.32	5%
28/12/2012	5473.62	4688.16	14%
29/12/2012	6385.89	5601.37	12%
30/12/2012	6081.8	4392.24	28%
31/12/2012	7906.34	5112.21	35%

Fuente: Elaboración propia

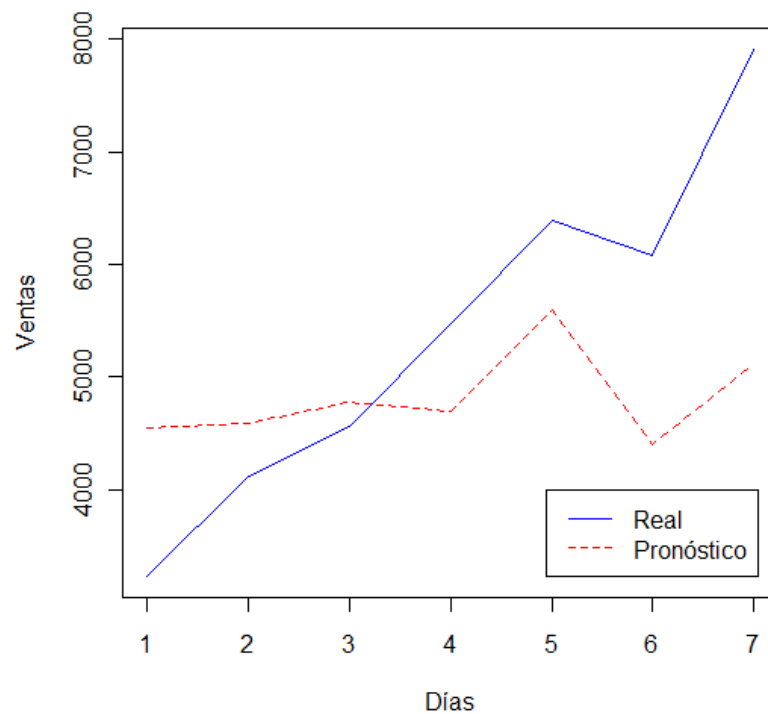


Figura 38. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1200

Referencia 1400

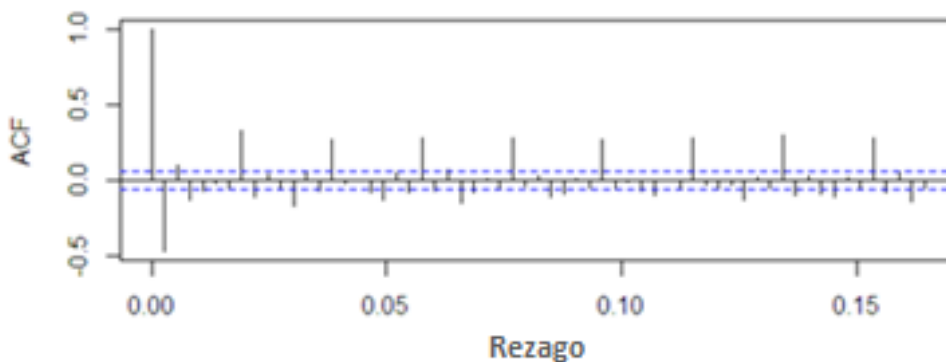


Figura 39. Función de autocorrelación diferenciada, referencia 1400

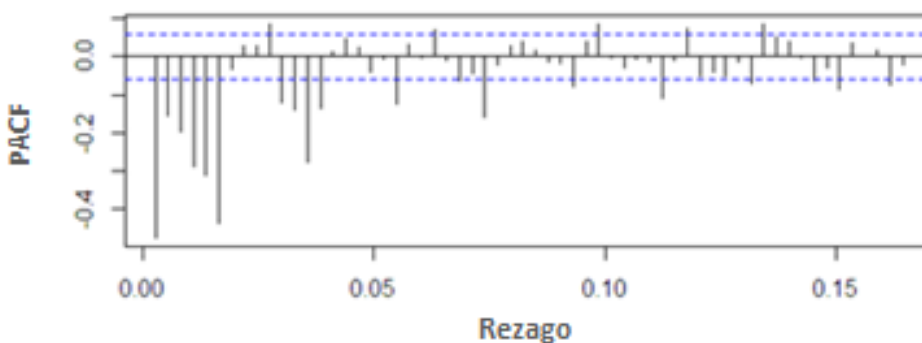


Figura 40. Función de autocorrelación parcial diferenciada, referencia 1400

En las figuras 39 y 40 se observan muchos valores de autocorrelación por fuera de las bandas de confianza, similar a las anteriores referencias, se presenta un patrón estacional cada 7 periodos.

Analizando la ACF en el componente no estacional (antes del periodo 7) se ve la dependencia de orden 1. En la PACF los 7 primeros valores están por fuera de las bandas de seguridad, es decir significativamente diferentes de cero y los últimos valores son los más alejados de las bandas de seguridad. Se puede evidenciar un patrón senoidal en ambas, por lo tanto se sugiere un modelo ARMA (4, 1).

Para la parte estacional, en la ACF las autocorrelaciones en los rezagos 7,14,..., 56 decrecen lentamente después del orden 1L. En la PACF los valores de

autocorrelación $\rho_{77}, \rho_{1414}, \dots, \rho_{5656}$ decrecen lentamente después del orden 2L. El comportamiento es de tipo SARMA (2, 1).

El análisis de las gráficas sugiere un modelo tentativo SARIMA (4, 1, 1) x (2, 0, 1) [7].

La prueba de normalidad de Jarque-Bera arroja un valor p de 2.2×10^{-16} , lo cual afirma que los residuales no provienen de una distribución normal. La prueba de Box Pierce con un valor p de 0.000131926, al ser menor que el nivel de significancia del 5%, indica que no hay información en los residuales. Se obtuvo un MAPE de ajuste de 11.69%, de pronóstico de 22.77% y un MSE de 269678.25.

La función automática de *Statgraphics* estima un modelo SARIMA (2, 0, 1) x (0, 1, 1) [7], con un MAPE de ajuste del 9.69% y un MSE de 235987.41.

El modelo no cumple con la prueba de normalidad de Jarque-Bera, ni con la prueba de incorrelación de Box-Pierce, ya que arroja un valor p de 2.2×10^{-6} y 0.000173317 respectivamente, inferiores al grado de significancia del 5%. Por lo tanto los residuales no provienen de una distribución normal y están autocorrelacionados.

El MAPE de pronóstico fue del 21.31%. Debido a que éste es inferior al del modelo anterior y ninguno de los modelos cumplió con las pruebas de validación, se elige este último como modelo para la referencia 1400: SARIMA (2, 0, 1) x (0, 1, 1) [7], cuyos coeficientes se muestran en la tabla 20.

Tabla 20. Coeficientes del modelo Box-Jenkins para la referencia 1400

Parámetro	Estimado	Error estándar	T
AR(1)	0.613109	0.077518	7.90926
AR(2)	0.186139	0.0338586	5.49754
MA(1)	0.622257	0.0761952	8.16661
SMA(1)	0.969133	0.00779879	124.267
Media	5.18917	1.19557	4.34034
Constante	1.04174		

Fuente: Programa estadístico R

La siguiente ecuación representa el modelo general estimado.

$$Z_t = -1.04174 - 0.613109Z_{t-1} - 0.186139Z_{t-2} - 0.622257a_{t-1} - 0.969133a_{t-7}$$

Los pronósticos para la semana del 25 al 31 de diciembre del 2012 se muestran en la siguiente tabla:

Tabla 21. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1400

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	2984.25	3449.19	16%	21.31%
26/12/2012	3979	3224.36	19%	
27/12/2012	3780.05	3426.15	9%	
28/12/2012	5045.52	3461.2	31%	
29/12/2012	5465.98	4330.15	21%	
30/12/2012	4625.06	3773.37	18%	
31/12/2012	5886.44	3844.86	35%	

Fuente: Elaboración propia

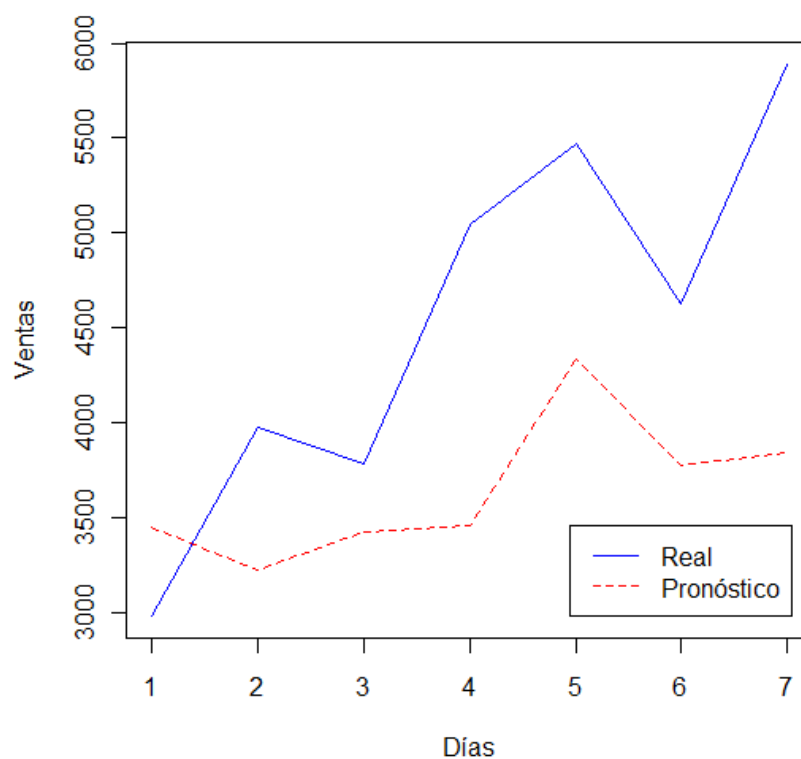


Figura 41. Demanda real vs pronóstico con modelo Box Jenkins, referencia 1400

Referencia 4100

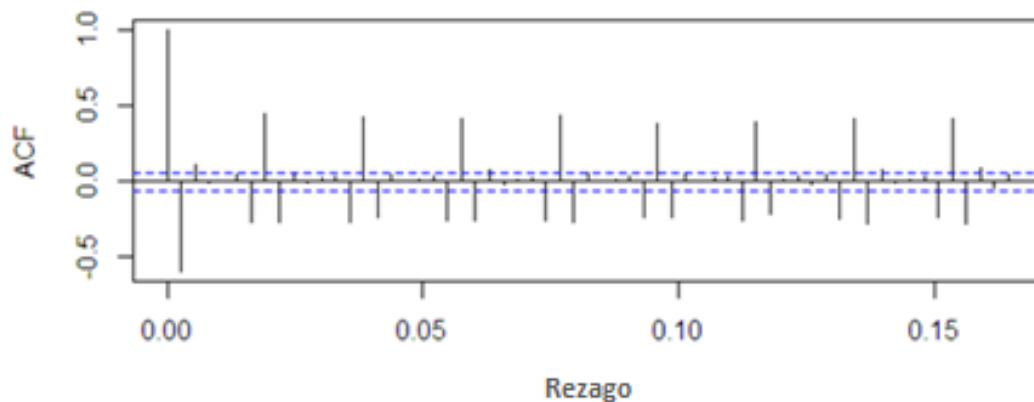


Figura 42. Función de autocorrelación diferenciada, referencia 4100

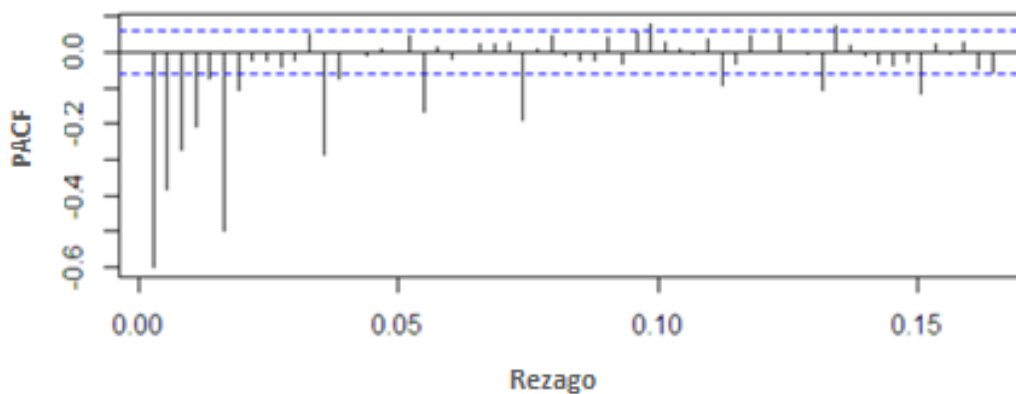


Figura 43. Función de autocorrelación parcial diferenciada, referencia 4100

La autocorrelación de orden 7 es muy fuerte por que se sale de las bandas de confianza, tal como se muestra en las figuras 42 y 43, lo que evidencia la estacionalidad de la serie. Debido a esto se estima que el modelo debe contener una parte estacional.

El análisis de la parte no estacional (antes del rezago 7), los rezagos 1,2 y 5 son significativamente diferentes de cero, ya que están por fuera de las bandas de la ACF. En la PACF el decrecimiento es lento, los valores de autocorrelación son grandes debido a que los primeros 7 rezagos están por fuera de las bandas.

En ambas se puede evidenciar una dependencia de orden 1, ya el primer valor es el más alejado de las bandas de seguridad y un patrón senoidal, por lo tanto el comportamiento es tipo ARMA (1,1).

Para la parte estacional, en la ACF los valores de autocorrelación de orden 7 decrecen lentamente después del orden 1L, en la PACF después del orden 2L. Por lo tanto el comportamiento es tipo SARMA (2, 1).

El modelo tentativo producto del análisis de las gráficas ACF y PACF, es un SARIMA (1, 1, 1) x (2, 0, 0) [7].

La prueba de normalidad Jarque-Bera arroja un valor p de 2.2×10^{-16} , lo cual afirma que los residuales no provienen de una distribución normal, pero la prueba de incorrelación Box Pierce con un valor p de 0.0728988, al ser mayor que un nivel de significancia del 5%, indica que los residuales no están autocorrelacionados. El MAPE de ajuste fue de 10.4%, el de pronóstico 14.29% y el MSE 3526803.96.

Finalmente, el mejor modelo encontrado fue un SARIMA (2, 0, 1) x (0, 1, 2) [7] por medio de la función automática de *Statgraphics*, con un MAPE de ajuste del 9.11% y un MSE de 2978745.95.

El modelo cumple con la prueba de incorrelación de Box-Pierce con un valor p de 0.358431, superior a un grado de significancia del 10%, pero no con la prueba de normalidad Jarque-Bera ya que el valor p es de 2.2×10^{-16} , por lo tanto los residuales no provienen de una distribución normal.

El MAPE de pronóstico es de 18.59%, superior al del modelo anterior. Debido a esto se elige el primero como modelo Box-Jenkins para esta referencia.

Tabla 22. Coeficientes del modelo Box-Jenkins para la referencia 4100

Parámetro	Estimado	Error estándar	t
AR(1)	-0.157894	0.0300541	-5.25366
MA(1)	0.992474	0.00343972	288.534
SAR(1)	0.310788	0.0296075	10.497
SAR(2)	0.284313	0.0295876	9.60919
Media	1.17204	1.63933	0.714948
Constante	0.549487		

Fuente: Programa estadístico R

$$Z_t = -0.549487 + 0.613109Z_{t-1} - 0.992474a_{t-1} - 0.310788Z_{t-7} - 0.284313Z_{t-14}$$

Los pronósticos para la semana del 25 al 31 de diciembre del 2012 se muestran en la siguiente tabla:

Tabla 23. Demanda real vs pronóstico con modelo Box Jenkins, referencia 4100

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	11017.4175	14011.9	0.733027588	14.29%
26/12/2012	16123.05	14318.2	0.609318132	
27/12/2012	13104.815	15339.6	0.682461472	
28/12/2012	18004.0725	15209.4	0.563759722	
29/12/2012	14532.2424	15583.1	0.647891006	
30/12/2012	15048.18	15803.6	0.635398929	
31/12/2012	19657	16354.1	0.523743761	

Fuente: Elaboración propia

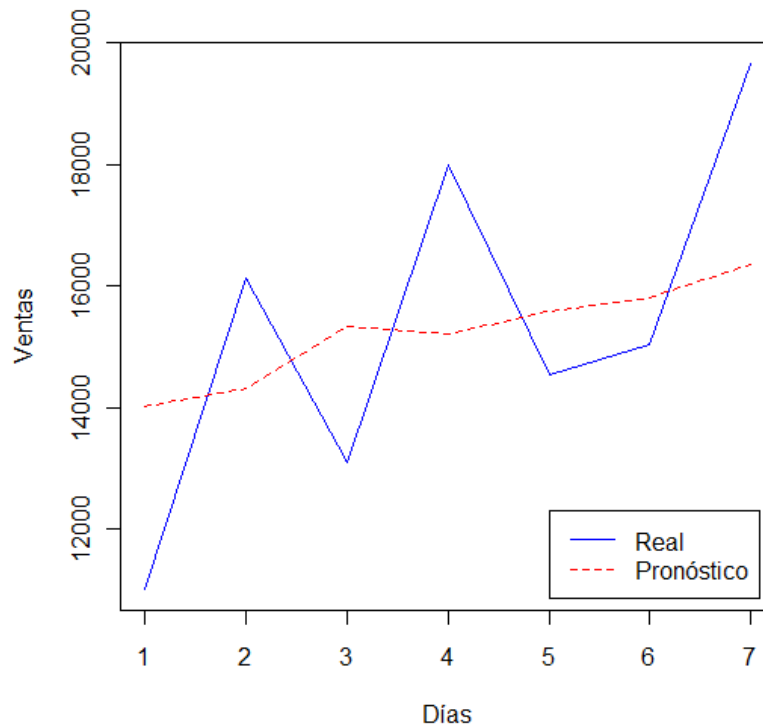


Figura 44. Demanda real vs pronóstico con modelo Box Jenkins, referencia 4100

Debido a los altos valores de MSE y MAPE de pronóstico se concluye que los modelos Box Jenkins son poco precisos y eficientes a la hora de pronosticar. Lo cual se corrobora en las figuras 38, 41 y 44 donde se evidencia la falta de capacidad del modelo para estimar pronóstico (línea punteada) que imite un comportamiento similar a la serie real (línea continua).

5.2.4 Modelos de Redes Neuronales

Referencia 1200

Las variables significativas, que constituyen las entradas de la RNA, para esta referencia son:

- Tiempo.
- Variable autorregresiva de orden 6.
- Variables exógenas de nivel y día de la semana.

La salida es la variable respuesta con transformación logarítmica.

En la creación de la red se asignó una función de entrenamiento lineal, de aprendizaje de momentos, de transferencia logística y de rendimiento el MSE. El número de capas ocultas es 2 y el de neuronas es 5, ya que son las variables de entrada más la de salida.

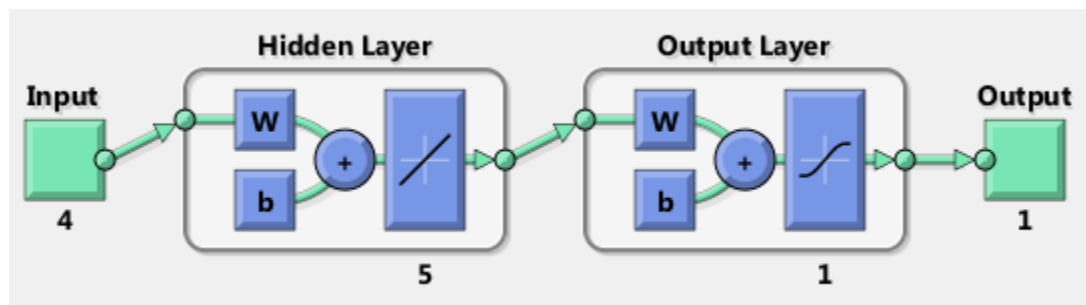


Figura 45. Estructura de la RNA, referencia 1200

Se entrenó la red, con un número de iteraciones igual a 1000 y se obtuvo un MSE de 0.0013, debido a que éste es inferior al del modelo dinámico para esta referencia, se prosiguió con la simulación de la red.

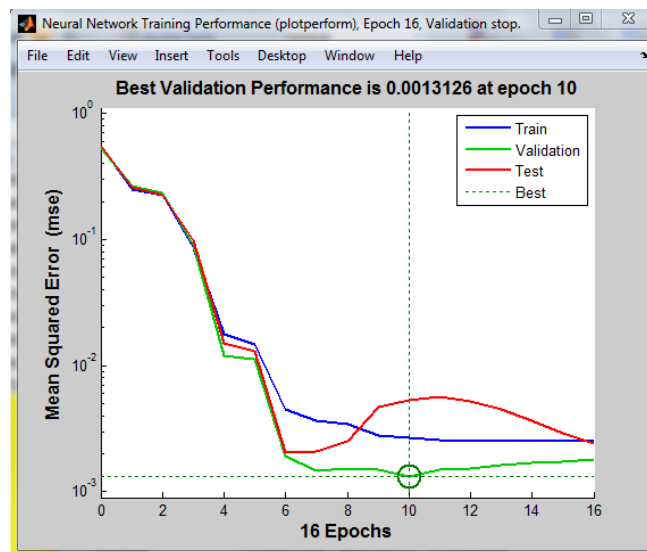


Figura 46. Rendimiento del modelo de RNA, referencia 1200

Se simuló la red y se obtuvo un MAPE de ajuste de 3.42%. Los pronósticos generados por la simulación para la semana del 25 al 31 de diciembre del 2012 se muestran en la siguiente tabla:

Tabla 24. Demanda real vs pronóstico mediante RNA, referencia 1200

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	3233	3241.37811	0.26%	3.93%
26/12/2012	4120	4204.60683	2.05%	
27/12/2012	4561.35	4662.79773	2.22%	
28/12/2012	5473.62	5914.02499	8.05%	
29/12/2012	6385.89	6818.02713	6.77%	
30/12/2012	6081.8	5821.99732	4.27%	
31/12/2012	7906.34	7598.97006	3.89%	

Fuente: Elaboración propia

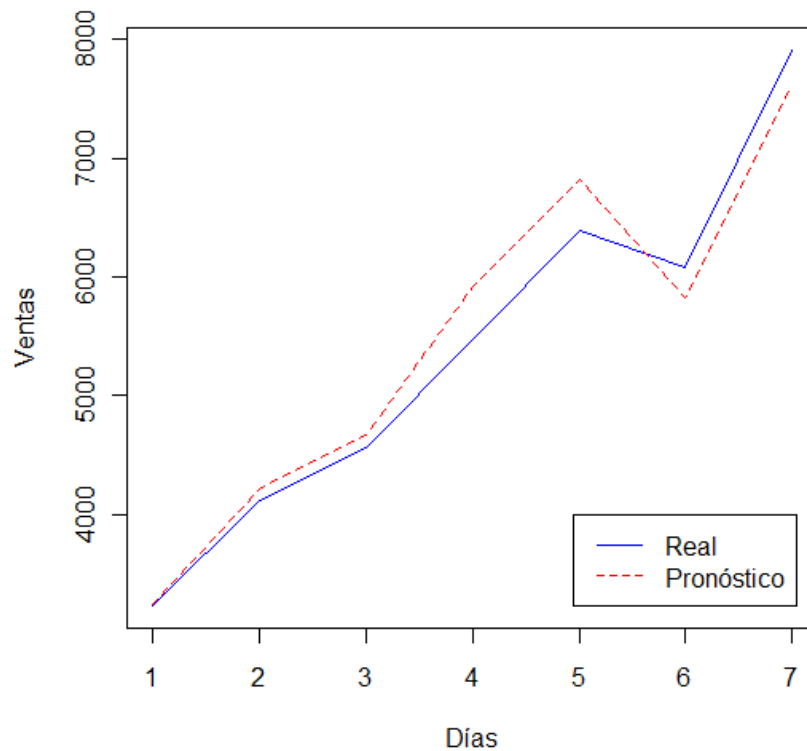


Figura 47. Demanda real vs pronóstico mediante RNA, referencia 1200

Referencia 1400

Las variables significativas, que constituyen las entradas de la RNA, para esta referencia son:

- $\cos(4\pi t/7)$, donde t es el tiempo.
- Variables exógenas de nivel y día de la semana.

La salida es la variable respuesta con transformación de raíz cuadrada.

Se creó una RNA con una función de entrenamiento lineal, de aprendizaje de momentos, de transferencia logística y de rendimiento el MSE. El número de capas ocultas es 2 y el de neuronas es 4.

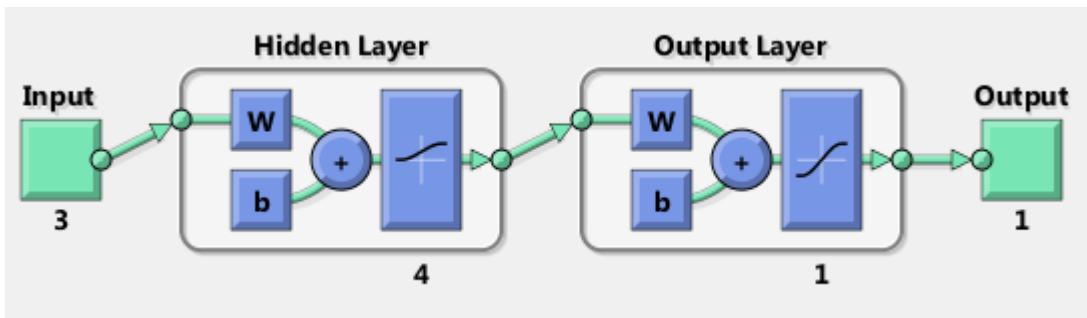


Figura 48. Estructura de la RNA, referencia 1400

En el entrenamiento de la red se asignó un número de iteraciones igual a 1000 y se obtuvo un MSE de 1.7325, debido a que éste es inferior al del modelo dinámico para esta referencia, se prosiguió con la simulación de la red.

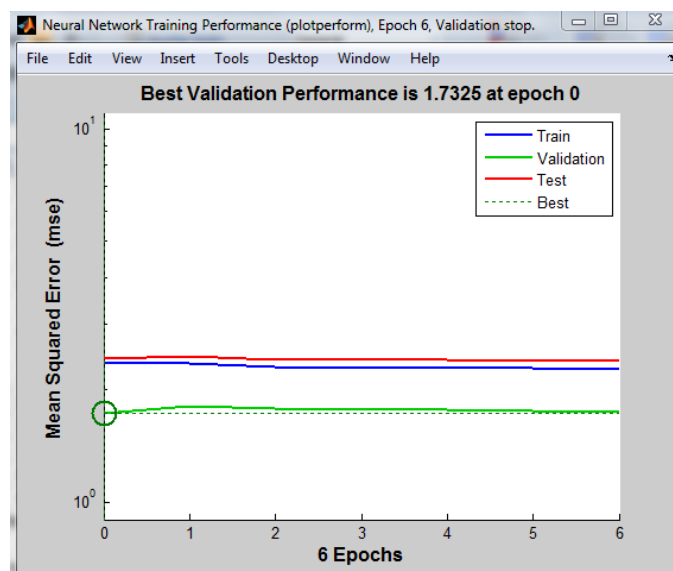


Figura 49. Rendimiento del modelo de RNA, referencia 1400

Se simuló la red y se obtuvo un MAPE de ajuste de 4.29%. Los pronósticos generados por la simulación para la semana del 25 al 31 de diciembre del 2012 se muestran en la siguiente tabla:

Tabla 25. Demanda real vs pronóstico mediante RNA, referencia 1400

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	2984.25	2902.06917	2.75%	6.99%
26/12/2012	3979	3854.93637	3.12%	
27/12/2012	3780.05	3868.21316	2.33%	
28/12/2012	5045.52	4792.68237	5.01%	
29/12/2012	5465.98	4804.43112	12.10%	
30/12/2012	4625.06	4815.74596	4.12%	
31/12/2012	5886.44	4739.29742	19.49%	

Fuente: Elaboración propia

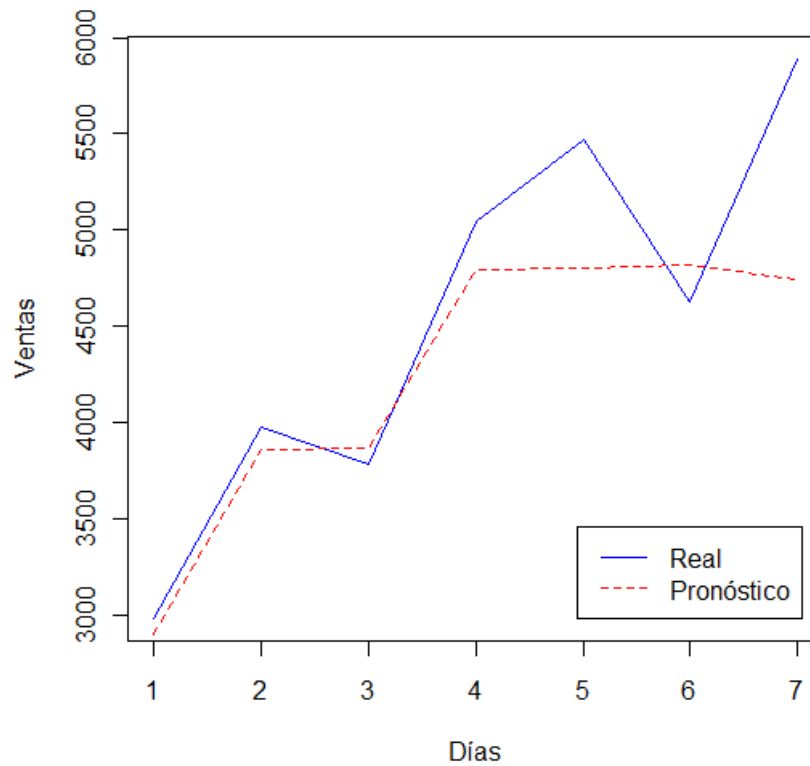


Figura 50. Demanda real vs pronóstico mediante RNA, referencia 1400

Referencia 4100

Las variables significativas, que constituyen las entradas de la RNA, para esta referencia son las variables exógenas de nivel y día de la semana. La salida es la variable respuesta con transformación logarítmica.

En la creación de la red se asignó una función de entrenamiento lineal, de aprendizaje de momentos, de transferencia lineal y de rendimiento el MSE. El número de capas ocultas es 2 y el de neuronas es 3.

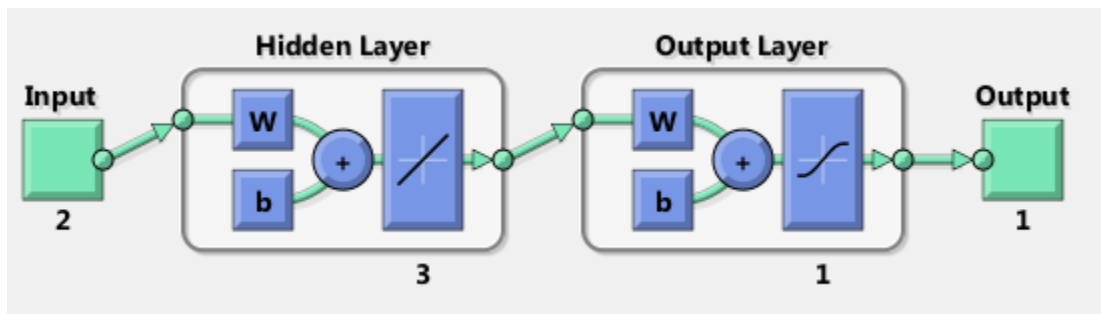


Figura 51. Estructura de la RNA, referencia 4100

Se entrenó la red, con un número de iteraciones igual a 1000 y se obtuvo un MSE de 0.0018, debido a que éste es inferior al del modelo dinámico para esta referencia, se prosiguió con la simulación de la red.

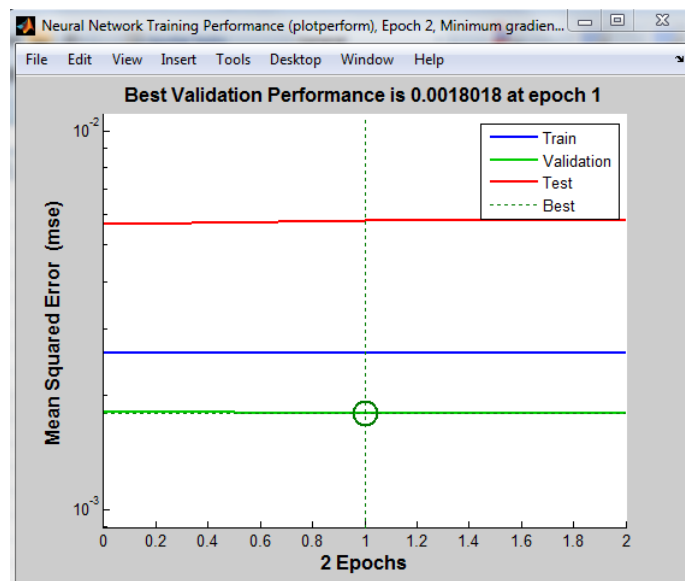


Figura 52. Rendimiento del modelo de RNA, referencia 4100

Se simuló la red y se obtuvo un MAPE de ajuste de 3.77%. Los pronósticos generados por la simulación para la semana del 25 al 31 de diciembre del 2012 se muestran en la siguiente tabla:

Tabla 26. Demanda real vs pronóstico mediante RNA, referencia 4100

Fecha	Real	Pronóstico	Error relativo	MAPE
25/12/2012	11018	11540.1744	4.74%	2.86%
26/12/2012	16124	16475.0331	2.18%	
27/12/2012	13105	13054.0221	0.39%	
28/12/2012	18005	17878.5616	0.70%	
29/12/2012	14533	14661.3523	0.88%	
30/12/2012	15049	14610.7575	2.91%	
31/12/2012	19657	18039.6591	8.23%	

Fuente: Elaboración propia

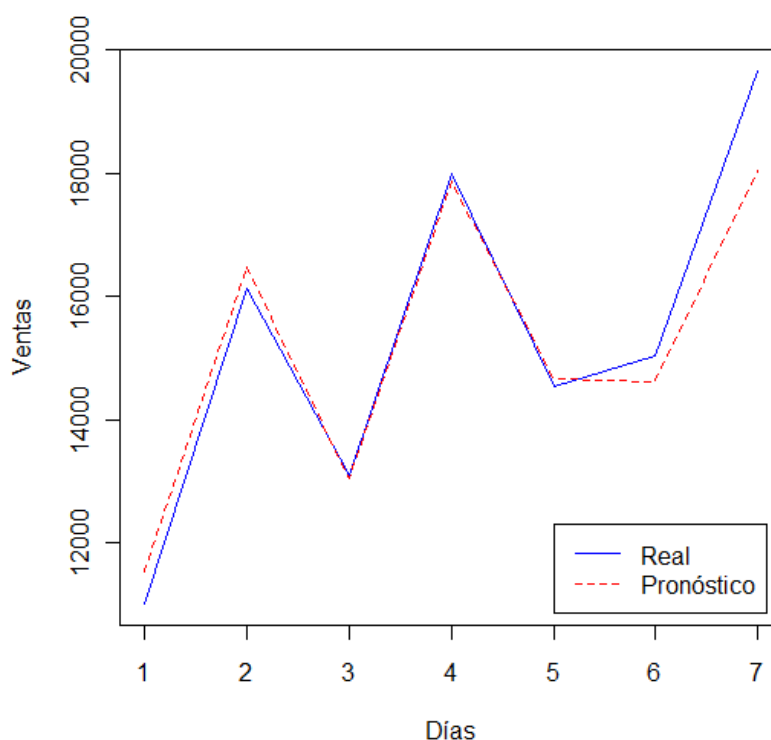


Figura 53. Demanda real vs pronóstico mediante RNA, referencia 4100

Los bajos valores de MSE y MAPE de pronóstico indican una gran precisión y eficiencia de los modelos de RNAs para pronosticar, lo que se corrobora en las figuras 47, 50 y 53, en las cuales se observa que el pronóstico (línea punteada) tiene un comportamiento muy similar a la serie real (línea continua).

Tabla 27. Resumen de valores de MSE, MAPE de pronóstico y cumplimiento de supuestos de validación

Modelo	MAPE Pronóstico (%)			MSE			Prueba de Normalidad			Prueba de Incorrelación		
	1200	1400	4100	1200	1400	4100	1200	1400	4100	1200	1400	4100
Suavización exponencial simple	20.47	22.92	20.53	433014.26	242877.19	3067585.64	x	x	x	x	x	x
Suavización exponencial Holt	20.45	22.04	20.63	435022.12	245292.02	3086120.83	x	x	x	x	x	x
Dinámico	5.97	7.44	3.27	0.0025	2.0670	0.0022	x	x	x	√	√	√
Box-Jenkins	20.92	21.31	14.29	417304.11	235987.41	3526803.97	x	x	x	√	x	√
RNA	3.93	6.99	2.86	0.0013	1.7325	0.0018	N/A	N/A	N/A	N/A	N/A	N/A

Fuente: Elaboración propia

En la Tabla 27 se observa cómo los MAPE de pronóstico y el MSE de los modelos de RNAs dan valores inferiores a los obtenidos con todos los modelos clásicos evaluados. Por lo tanto se concluye que los modelos de RNAs son los de mayor precisión y eficiencia para la estimación del pronóstico de ventas de las tres referencias para los siguientes siete días.

5.3. Modelo de optimización de inventarios

El modelo basado en revisión periódica, acorde con lo propuesto en la sección 4.2, debe cumplir con la demanda requerida X_{ij} , para cada uno de los productos ($i=1,2,3$), para los 7 días de una semana ($j=1,2,...7$). En este caso ésta es la pronosticada por el modelo de RNAs, por ser los más eficientes.

Tabla 28. Límite superior e inferior de demanda por producto y día

Pedido	Demanda		
	X_{1j}	X_{2j}	X_{3j}
1	3241.378	2902.07	11540.17
2	4204.607	3854.94	16475.03
3	4662.798	3868.21	13054.02
4	5914.025	4792.68	17878.56
5	6818.027	4804.43	14661.35
6	5821.997	4815.75	14610.76
7	7598.97	4739.3	18039.66

Fuente: Elaboración propia

Tabla 29. Costos asociados a los inventarios y utilidad por referencia

Producto	Costo de mantener (CM_i)	Costo de producción (CP_i)	Utilidad (U_i)
1	327.08	959.12	294.78
2	577.84	1704.32	322.61
3	226.33	663.67	224.40

Fuente: Elaboración propia

Dadas las utilidades por producto y las demandas pronosticadas se estima una utilidad total de la producción de \$45.270.00 lo cual da por arriba del mínimo estipulado (45 millones), con el sobrante de \$ 270.066.

De acuerdo con lo anterior se plantea, y resuelve en *LINGO*, el modelo de revisión periódica que emita las cantidades óptimas a producir de manera que se

minimicen los costos que acarrean los inventarios, asegurando que se cumpla con la utilidad operacional mínima esperada por la empresa.

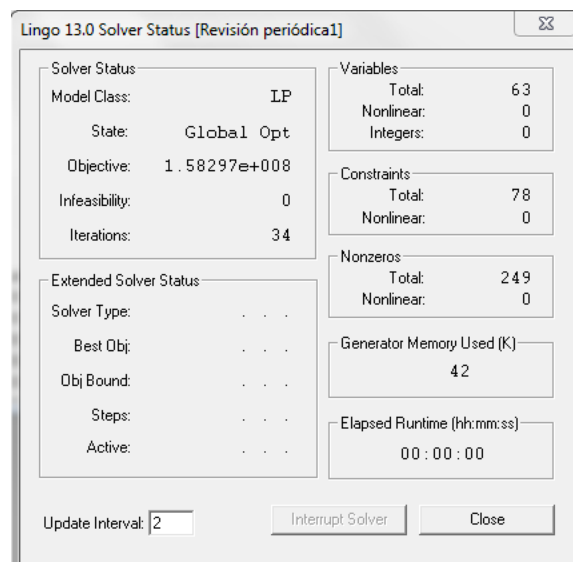


Figura 54. Solución del sistema de inventarios en LINGO

Los costos totales mínimos alcanzados, de la producción más los inventarios, a una utilidad mínima esperada de 45 millones, fue de \$158.297.000.

Los valores de la solución de pedidos e inventarios estimados por el software *LINGO* se muestran para cada una de las referencias en las siguientes Tablas 30, 31 y 32.

Debido a que se está hablando de unidades de producto se redondean los resultados hacia arriba, en dirección contraria a cero y sin cifras decimales.

Referencia 1200

Tabla 30. Valores de la solución de pedidos e inventarios, referencia 4100

Periodo	Producción (P_{ij})	Inventario (I_{ij})
0		1128
1	2114	0

2	4205	0
3	10559	5896
4	19	0
5	9653	2834
6	2988	0
7	7599	0

Fuente: Elaboración propia

Referencia 1400

Tabla 31. Valores de la solución de pedidos e inventarios, referencia 4100

Periodo	Producción (P_{ij})	Inventario (I_{ij})
0		835
1	2068	0
2	4108	253
3	3616	0
4	4793	0
5	4805	0
6	4816	0
7	4740	0

Fuente: Elaboración propia

Referencia 4100

Tabla 32. Valores de la solución de pedidos e inventarios, referencia 4100

Periodo	Producción (P_{ij})	Inventario (I_{ij})
0		3412
1	8129	0
2	17553	1077
3	11978	0

4	21339	3460
5	11694	492
6	18347	4228
7	13812	0

Fuente: Elaboración propia

De los resultados mostrados en las Tablas 30, 31 y 32 la empresa puede realizar la planeación de la producción, de tal forma que los costos sean mínimos y la empresa reciba la utilidad mínima esperada.

Tabla 33. Valores de costo reducido

Variable	Costo reducido
I11	577.84
I17	4426.16
I21	577.84
I27	5171.36
I31	577.84
I37	4130.71

Fuente: Elaboración propia

En la Tabla 33 se muestran las cantidades de inventario que tienen valores positivos de costo reducido, indicando que hay pérdidas por el almacenamiento de inventario de los 3 productos en el primer y último periodo. Ninguna variable estructural tiene costo reducido.

Tabla 34. Excedentes y precios sombra

Fila	Holgura o excedente	Precio sombra
1	1.58E+08	-1.00
2	0	-960
3	0	-960
4	0	-1537
5	0	-2115

Fila	Holgura o excedente	Precio sombra
6	0	-2693
7	0	-3271
8	0	-3849
9	1128	0
10	0	-1705
11	0	-1705
12	0	-2283
13	0	-2860
14	0	-3438
15	0	-4016
16	0	-4594
17	835	0
18	0	-664
19	0	-664
20	0	-1242
21	0	-1820
22	0	-2398
23	0	-2976
24	0	-3553
25	3412	0
26	0	251
27	4696	0
28	0	352
29	13841	0
30	285	0
31	0	578
32	0	1156
33	0	1734
34	0	2312

Fila	Holgura o excedente	Precio sombra
35	0	2890

Fuente: Elaboración propia

Para la primera y tercera referencia la capacidad de almacenamiento semanal llega al tope, para la segunda referencia hay un sobrante de 4696 debido a que solo se almacenan 253 unidades del producto en la semana.

En todos los períodos se llega al límite de la capacidad de producción, excepto para el primer y segundo periodo, para los cuales hay sobrantes de 13841 y 285 respectivamente.

Los precios sombra negativos indican el grado en que disminuyen los costos por cada valor que aumentan los recursos de las restricciones de pedidos e inventarios; precisamente, al aumentar las unidades de inventarios, disminuye la función objetivo en estas cantidades, aspectos que se veían reflejados en los costos reducidos.

Los precios sombra positivos muestran cómo aumentarían los costos por cada unidad adicional que se pida o almacene de cada producto en todos los periodos.

CONCLUSIONES

El modelo de pronóstico con redes neuronales frente a los demás modelos clásicos usados para la estimación de la demanda fue el más eficiente bajo la comparación del error MAPE y MSE, pero el acierto de este modelo se debe en gran medida a la inclusión de variables significativas realizadas durante la estimación del modelo dinámico. Es importante resaltar el alto peso que tienen los niveles de demanda y los días de la semana, este último permite comprobar la estacionalidad de cada siete días que se presenta en la serie temporal para las tres referencias.

El tiempo de cómputo para procesar las iteraciones requeridas para lograr una minimización de los errores del pronóstico realizado bajo el modelo de redes neuronales, fue muy corto, lo cual se convierte en una gran ventaja frente a los demás modelos; además no es necesaria una re-estimación de los coeficientes y el peso que las variables poseen en las ecuaciones del pronóstico, tal como ocurre en los modelos dinámicos. Las pruebas de validación de correlación y normalidad que se deben realizar sobre los residuales en los modelos tradicionales es una razón más para preferir el modelo de RNAs.

Es importante destacar la eficiencia de las RNAs para procesar grandes cantidades de información y variables asociados a ellas, como se evidenció en el desarrollo del modelo para las tres referencias.

El modelo de optimización de inventarios basado en revisión periódica permitió proponer una planeación de la producción que garantice el cumplimiento de la demanda real, generar un plan de pedidos y almacenar inventarios a costo mínimo durante los 7 días de una semana, lo cual es de gran importancia para las empresas que se dedican a la producción de derivados lácteos, ya que al ser productos perecederos, con una corta vida útil, su sobreproducción puede generar grandes pérdidas debido a los altos costos de almacenamiento y dadas de baja.

RECOMENDACIONES

La estimación de demanda, producción y principalmente de inventarios en las empresas distribuidoras y para nuestro caso productoras, se realizan generalmente a través de la experimentación o juicio de expertos; aunque estos pronósticos cualitativos son muy importantes para realizar estas actividades, el nivel de servicio y de respuesta serían mucho mejor si además se integran modelos cuantitativos.

La organización de la información es muy importante para esta empresa láctea, por su alta variación en demandas y producción, esto exige que tenga una sistematización constante y verídica, para mejorar las bases de toma de decisiones. Se sugiere que la demanda sea ingresada diariamente, evitando efectos atípicos o extremos que puedan afectar la planeación.

El modelo propuesto de optimización permitió asignar las restricciones según las capacidades de la empresa e incluso, presupuesto y mínimas utilidades, de esta forma se tiene un modelo, que además de minimizar los costos, permite flexibilidad, tal como sucede en la realidad con la programación de la producción.

Las propuestas metodológicas planteadas en este trabajo, pueden de ser empleadas en empresas del mismo sector, e incluso extrapolarse para empresas que quieran o necesiten un apoyo para la planeación de su producción y deseen una reducción en los niveles de inventarios, lo cual sería un buen comienzo si en un futuro se desean implementar mejores técnicas estadísticas para su optimización.

La propuesta de RNAs es de fácil estimación en el software MATLAB, bajo compra de licencia, sin embargo, la empresa podría apoyarse en software libre que permite la construcción de otros modelos estadísticos, entre ellos el software *R*; su utilización es especial para aquellas empresas que desean implementar estos modelos sin los altos niveles de inversión en programas licenciados.

REFERENCIAS BIBLIOGRÁFICAS

- Adam, E. E., & Ebert, R. J. (1991). *Administración de la producción y las operaciones: conceptos, modelos y funcionamiento*. Madrid: Pearson Educación.
- Aguirre, A. (1994). *Introducción al tratamiento de series temporales: aplicación a las ciencias de la salud*. Madrid: Ediciones Díaz de Santos.
- Alonso, P., & Villegas, E. (2005). La importancia de los criterios cualitativos de los pronósticos en los Agronegocios. *Revista Mexicana de Agronegocios* , 464 - 477.
- Álvarez, R. (1994). *Estadística multivariante y no paramétrica con SPSS: Aplicación a las ciencias de la salud*. Madrid: Ediciones Díaz de Santos.
- Anderson, D. R., & Sweeney, D. J. (2008). *Estadística para administración y economía*. México: Cengage Learning Editores.
- Au, K. F., Choi, T.M., & Yu, Y. (2008). *Fashion retail forecasting by evolutionary neural networks*. Hong Kong: International Journal of Production Economics.
- Benseman, B. R. (1986). Production Planning in the New Zealand Dairy Industry. *The Journal of the Operational Research Society*, Vol. 37, No. 8 .
- Bowerman, B. L., & O'Connell, R. T. (2007). *Pronósticos, series de tiempo y regresión: Un enfoque aplicado*. Cengage Learning Latin America.
- Caridad y Ocerín, J. M. (1998). *Econometría: Modelos económicos y series temporales*. Barcelona: Editorial Reverté.
- Chiu, Y. P. (2003). *Determining the optimal lot size for the finite production model with random defective rate, the rework process, and backlogging*. Taiwan: Engineering Optimization.
- Correa, A., & Gómez, R. (2009). Cadena de Suministro para el sector minero como una estrategia para la productividad. *Boletín Ciencias de la Tierra* .
- Devore, J. L. (2004). *Probabilidad y estadística para ingeniería y ciencias*. México: Cengage Learning Editores.
- Doganis, P., Alexandridis, A., Patrinos, P., & Sarimveis, H. (2006). Time series sales forecasting for shelf-life food products based on artificial neural networks and evolutionary computing. *Journal of food Engineering* 75 , 196-204.
- Flórez, R., & Fernández, J. M. (2008). *Las Redes Neuronales Artificiales*. España: Netbilo S.L.

- Gutierrez, R. S., Solis, A. O., & Mukhopadhyay, S. (2008). *Lumpy demand forecasting using neural networks*. El paso, Texas: International Journal of Production Economics.
- Hanke, J. E., & Wichern, D. (2006). *Pronósticos en los negocios*. México: Pearson Education.
- Hillier, F. S., & Lieberman, G. J. (2002). *Teoría de Inventarios*. México: McGraw Hill.
- Izar, J. M. (1996). *Fundamentos de investigación de operaciones para administración*. San Luis Potosí: Editorial Universitaria Potosina.
- Koh, S. G., Hwang, H., Sohn, K. I., & Ko, C. S. (2002). *An optimal ordering and recovery policy for reusable items*. Korea del sur: Computers and Industrial Engineering.
- Lotfi, K. G., & Choueiki, M. H. (2000). *A neural network model for solving the lot sizing Problem*. Kuwait: Omega 28.
- Medina, S., & García, J. (2005). Predicción de demanda de energía en Colombia mediante un sistema de inferencia difuso neuronal. *Energética* 33 , 15-24.
- Montgomery, D. C., Peck, E., & Vining, G. G. (2006). *Introducción al análisis de regresión lineal*. México: CECSA.
- Pankratz, A. (1983). *Forecasting with Univariate Box - Jenkins Models: Concepts and Cases*. Canadá: John Wiley & Sons, Inc.
- Partovi, F. Y., & Anandarajan, M. (2002). *Classifying inventory using an artificial neural network approach*. Philadelphia: Computers and Industrial Engineering Vol.41.
- Pepió Viñals, M. (2001). *Series Temporales*. Cataluña: Ediciones Universit.
- Ramírez, S. (2013). *Comparación de metodologías estadísticas para el pronóstico de demanda de productos de difícil estimación*. (T. d. Industrial, Ed.) Medellín: Universidad Pontificia Bolivariana.
- Render, B. (2004). *Principios de administración de operaciones*. Madrid: Pearson Educación.
- Rob A.C.M. Broekmeulen, K. H. (2009). A heuristic to manage perishable inventory with batch ordering, positive lead-times, and time-varying demand. *Computers & Operations Research* 36 .
- Rueda, V. M., Velásquez, J. D., & Franco, C. J. (2011). *AVANCES RECIENTES EN LA PREDICCIÓN DE LA DEMANDA DE ELECTRICIDAD USANDO MODELOS NO LINEALES*. Bogotá D.C.

Ruíz, Á., & Morillo, L. E. (2004). *Epidemiología clínica: investigación clínica aplicada*. Bogotá, D.C, Colombia: panamericana.

Simchi-Levi, D., Kaminsky, P., & Simchi-Levi, E. (2008). *Designing and Managing the Supply Chain*. California: McGraw-Hill.

Srivaree-ratana, C., Konak , A., & Smith, A. E. (2002). *Estimation of all-terminal network reliability using an artificial neural network*. Auburn, Alabama: Computer and operations Research.

Tabares, H., & Hernandez, J. (Octubre de 2006). *Pronostico puntos críticos de la serie temporal "Consumo de energía eléctrica del sector industrial en la ciudad de Quito" usando algoritmos genéticos*. Recuperado el 4 de Septiembre de 2013, de <http://ingenieria.udea.edu.co/grupos/revista/buscador/main.php>

Taha, H. A. (2004). *Ivestigación de operaciones*. Arkansas: Pearson Educación.

Vaklieva-Bancheva, N., Shopova, E., Espuña, A., & Puigjaner, L. (2006). Product portafolio optimization for dairy industry. *International Mediterranean modelling multiconference*, (págs. 100-109). Bcelona.

Vidal Holguín, C. J., & Londoño Ortega, J. C. (2004). Aplicación de modelos de inventarios en una cadena de abastecimiento de productos de consumo masivo con una bodega y N puntos de venta. *Revista Ingeniería y Competitividad Vol 6* , páginas 35-52.

Vidal, C. J., & Gutiérrez, V. (2008). Modelos de Gestión de Inventarios en Cadenas de Abastecimiento. *Revista Facultad de Ingeniería Universidad de Antioquia* , 134 - 149.

Watson, R. B. (1987). The Effects of Demand-Forecast Fluctuations on Customer Service and Inventory Cost When Demand is Lumpy. *The Journal of the Operational Research Society, Vol. 38, No. 1* .

Wee, H. M., & Chung, C. J. (2009). *Optimizing replenishment policy for an integrated production inventory deteriorating model considering green component-value design and remanufacturing*. Pingtung, Taiwan: International Journal of Production Research.