

**OPTIMIZACIÓN DEL ABASTECIMIENTO ENERGÉTICO DE COLOMBIA,
USANDO UNA TÉCNICA DE PRONÓSTICO EFICIENTE PARA LA
DEMANDA**

JOSE FERNANDO TABARES MUÑOZ

CARLOS ANDRÉS VELÁSQUEZ GALVIS

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA DE INGENIERÍAS

FACULTAD DE INGENIERÍA INDUSTRIAL

INGENIERÍA INDUSTRIAL

MEDELLÍN

2013

**OPTIMIZACIÓN DEL ABASTECIMIENTO ENERGÉTICO DE COLOMBIA,
USANDO UNA TÉCNICA DE PRONÓSTICO EFICIENTE PARA LA
DEMANDA**

JOSE FERNANDO TABARES MUÑOZ

CARLOS ANDRÉS VELÁSQUEZ GALVIS

Trabajo de grado para optar al título de Ingeniería Industrial

Directora

MARISOL VALENCIA CÁRDENAS

Ingeniera Industrial, Magister en Ciencias-Estadística

UNIVERSIDAD PONTIFICIA BOLIVARIANA

ESCUELA DE INGENIERÍAS

FACULTAD DE INGENIERÍA INDUSTRIAL

INGENIERÍA INDUSTRIAL

MEDELLÍN

2013

NOTA DE ACEPTACIÓN

Firma
Nombre
Presidente del jurado

Firma
Nombre
Presidente del jurado

Firma
Nombre
Presidente del jurado

Medellín, 15 de agosto 2013

Este trabajo de grado está dedicado a Dios y a nuestros padres.

AGRADECIMIENTOS

Primero le agradecemos a Dios por permitirnos terminar exitosamente este trabajo de investigación, segundo a nuestros familiares por todo el apoyo que nos brindaron en todo este proceso de formación humana y profesional.

Agradecemos a nuestros docentes que aportaron mucho a nuestra formación, en especial a nuestra directora Marisol Valencia Cárdenas, quien nos guio en este difícil camino académico por medio de sus enseñanzas, y estuvo incondicionalmente para construir esta investigación de la mejor manera. Finalmente, queremos agradecerles a todas las personas que nos ayudaron dentro de XM Expertos en Mercado filial de ISA, que de alguna forma fueron claves para el desarrollo y ejecución del trabajo de grado.

CONTENIDO

1. OBJETIVOS	16
1.1. GENERAL	16
1.2. ESPECÍFICOS	16
2. ANTECEDENTES.....	17
2.1 Antecedentes del sector energético colombiano	17
2.2 Modelos de pronósticos.....	20
2.2.1 Suavización exponencial	21
2.2.2 Modelos ARIMA y SARIMA.....	25
2.2.3 Metodología Delphi.....	26
2.2.4 Modelo Econométrico Dinámico.....	28
3. MARCO TEÓRICO	33
3.1 Pronósticos	33
3.2 Definición del método Delphi.....	35
3.3 ARIMA.....	39
3.4 Suavización exponencial simple.....	43
3.5 Modelo suavización exponencial lineal: método Holt.....	44

3.6 Definición modelo dinámico.....	45
3.7 Pruebas de Hipótesis	46
3.7.1 Prueba de normalidad de Shapiro-Wilk.....	47
3.7.2 Prueba de Jarque Bera.....	48
3.7.3 Prueba de Box-Pierce.....	49
3.8 Técnica de Monte Carlo por Cadenas de Markov (MCMC)	50
4. METODOLOGÍA.....	53
4.1 Criterio de Evaluación de los modelos.....	55
4.2 Metodología Delphi	56
4.2.1 Planteamiento del objetivo del pronóstico.....	56
4.2.2 Composición del grupo coordinador	56
4.2.3 Algoritmo para el análisis de la distribución a posteriori.....	59
5. RESULTADOS	60
5.1 Suavización exponencial simple.....	60
5.2 Suavización exponencial lineal de holt	64
5.3 Modelo SARIMA.....	68
5.4 Modelo Dinámico.....	70
5.5 Método Delphi	81

5.6 Modelo de regresión Gaussiana vía Monte Carlo por Cadenas de Markov.	92
CONCLUSIONES	99
RECOMENDACIONES.....	101
BIBLIOGRAFÍA.....	102

LISTA DE FIGURAS

Figura 1. Comportamiento de la demanda energética en Colombia.....	53
Figura 2. Valor real vs valor pronosticado por Suavización Exponencial Simple	60
Figura 3. Autocorrelación de los residuales del modelo de suavización exponencial simple	61
Figura 4. Autocorrelación parcial de los residuales del modelo de suavización exponencial simple	62
Figura 5. Gráfica de pronóstico del modelo de suavización exponencial simple	63
Figura 6. Autocorrelacion de los residuales del modelo de suavización exponencial lineal	65
Figura 7. Autocorrelacion parcial de los residuales del modelo de suavización exponencial lineal	65
Figura 8. Gráfica del pronóstico del modelo de suavización exponencial lienal.	67
Figura 9. Gráfica de la tendencia de la serie.....	70
Figura 10. Autocorrelación de los residuales del modelo dinámico	76
Figura 11. Autocorrelación parcial de los residuales de los errores del modelo dinámico	77
Figura 12. Gráfico de frecuencia de los datos de lunes a viernes	84
Figura 13. Gráfico de frecuencia de los datos del fin de semana	84
Figura 14. Gráfico de frecuencia para el pronóstico de lunes a viernes	86
Figura 15. Gráfico de frecuencia para el pronóstico del fin de semana	86

Figura 16. Gráfico de los intervalos de alta densidad del pronóstico de lunes a viernes.....	88
Figura 17. Gráfico de los intervalos de alta densidad del pronóstico del fin de semana	89

LISTA DE TABLAS

Tabla 1. Resumen de los antecedentes del sector energético	32
Tabla 2 Resumen de la clasificación.....	33
Tabla 3. Resumen métodos y modelos estadísticos utilizados en el trabajo.	52
Tabla 4. Pronostico suavización exponencial simple	62
Tabla 5. Pronostico suavización exponencial Holt	66
Tabla 6. Coeficientes SARIMA.....	68
Tabla 7. Pronostico SARIMA	69
Tabla 8. Coeficientes Estimados del Modelo Dinámico.....	71
Tabla 9. Analisis de Varianza (ANOVA).....	74
Tabla 10. Residuales	75
Tabla 11. Coeficientes Estimados finales del modelo dinámico final.....	75
Tabla 12. Coeficientes ARIMA con errores del Dinámico.....	77
Tabla 13. Pronostico Dinámico con errores SARIMA.....	78
Tabla 14. Resumen pronosticos metodologías clásicas.....	80
Tabla 15. Resultados 2da. Ronda método Delphi (semana)	82
Tabla 16. Resultados 2da. Ronda método Delphi (semana)	83
Tabla 17 Estadísticas básicas para el consumo en los días de lunes a viernes.....	85

Tabla 18 Estadísticas básicas para el consumo en los días del fin de semana.....	85
Tabla 19. Muestreo para pronóstico en semana	87
Tabla 20. Muestreo para pronóstico del fin de semana.....	87
Tabla 21. Pronóstico para una semana no atípica	90
Tabla 22. Porcentaje disminución de energía	91
Tabla 23 Pronostico de demanda de energía para días atípicos.....	91
Tabla 24. Coeficientes Estimados.....	93
Tabla 25. Pruebas de estacionariedad e incorrelación.....	94
Tabla 26 Pronostico de demanda de energía	94
Tabla 27. Comparación de resultados de los modelos Bayesianos para el pronóstico energético diario en Colombia.	95
Tabla 28. Comparación del modelo clásico más eficiente con el modelo Bayesiano más eficiente.	96
Tabla 29. Síntesis de resultados.....	97

GLOSARIO

ELICITAR: Preguntar e indagar acorde una evaluación subjetiva, el comportamiento de una variable o factor determinado.

ESTACIONARIEDAD: Usado en series temporales, la cual describe a una serie que posee media constante. (Mahia, 2001)

ESTACIONALIDAD: En estadística, se dice que la demanda de un determinado producto muestra estacionalidad cuando la serie de tiempo subyacente atraviesa una variación cíclica predecible, dependiendo de la época del año. La estacionalidad es uno de los patrones estadísticos más utilizados para mejorar la precisión de los pronósticos de demanda. (Vermorel, 2011)

SIMULACIÓN MONTECARLO: es una técnica que combina conceptos estadísticos (muestreo aleatorio) con la capacidad que tienen los ordenadores para generar números pseudo-aleatorios y automatizar cálculos. (Anderson, Sweeney, & Williams, 2009)

MCMC (MÉTODOS DE CADENA DE MARKOV MONTE CARLO): son métodos de simulación para generar muestras de las distribuciones a posteriori y estimar cantidades de interés a posteriori. (Conchi, 2012)

Las siguientes definiciones fueron tomadas de Simon, D. & Larget, B. (2000):

CADENA DE MARKOV: Una secuencia de las variables al azar en las cuales la distribución de cada variable al azar depende solamente del valor de su precursor.

DISTRIBUCIÓN A PRIORI: La distribución de la probabilidad de unos o más parámetros antes de la observación de datos. Idealmente, representa la creencia anterior del investigador.

DISTRIBUCIÓN A POSTERIORI: La distribución condicional de la probabilidad de unos o más parámetros, después de la observación de datos. La distribución a posteriori es proporcional al producto de la probabilidad y de la distribución anterior. Una distribución a posteriori común describe la distribución más que una variable al azar.

MCMCpack: Es un paquete de software diseñado para permitir a los usuarios realizar inferencia bayesiana a través de Monte Carlo por cadenas de Markov (Universidad de Washington, 2013). El programa R es un lenguaje estadístico muy potente, libre, diseñado en un entorno para el diseño de funciones versátiles para programar y solucionar diferentes tipos de problemas matemáticos y estadísticos.

MAPE: Es el error absoluto porcentual promedio, y mide la exactitud que tuvo el pronóstico con relación a la serie de tiempo (porcentaje de error). (Mynsbrugge, 2010)

RESUMEN

En el presente trabajo de grado, se pretende mostrar una optimización del abastecimiento energético de Colombia, a partir de la exploración, estimación y análisis de varios modelos, obteniendo la técnica que se considera más eficiente para pronosticar el consumo energético diario típico y atípico, lo cual permite tener conocimiento previo y efectivo de dicho consumo.

Fue necesario realizar la estimación, validación y comparación de eficiencia en el acierto de los siguientes modelos: Suavización exponencial simple, Suavización exponencial lineal Holt, modelo ARIMA con coeficiente de estacionalidad, modelos dinámicos con variable endógena y exógena, Regresión gaussiana por Monte Carlo con Cadenas de Markov y Elicitación Bayesiana con simulación de pronósticos.

Se presentan dichos modelos, con el proceso necesario para su análisis y comparaciones respectivas con el fin de efectuar la caracterización del comportamiento y pronóstico de la demanda energética diaria del país; mostrando los desaciertos de algunos modelos como el de suavización exponencial, y las bondades de métodos como: la elicitación Bayesiana y el modelo dinámico, con el resultado más eficiente. Se presentan alternativas de pronósticos usando teoría bayesiana que no requieren las mismas premisas que muchos de los modelos clásicos ya conocidos, y que no necesariamente son muy eficientes.pp

PALABRAS CLAVE: Demanda, Energía, Pronóstico, Modelos Clásicos, Modelos Bayesianos.

INTRODUCCIÓN

El objetivo de esta investigación es hallar una técnica estadística más eficiente que permita optimizar la predicción de demanda de energía diaria de Colombia, con el mínimo error posible, mediante la comparación de diferentes modelos de series de tiempo. La idea surge a partir de la necesidad de encontrar un modelo, con la capacidad de explicar comportamientos típicos y atípicos del consumo energético del país, lo que proporcionaría al Centro Nacional de Despacho (CND) alternativas diferentes a las que usualmente ha manejado para tomar decisiones más adecuadas.

Existen diversas investigaciones sobre el comportamiento habitual de la energía en Colombia, que se han tenido en cuenta para este trabajo, sin embargo, se encontraron en ellas algunas falencias, que podrían afectar el error en el pronóstico, como por ejemplo, la falta de integración de variables significativas dentro de los modelos, que pueden ayudar a captar cambios estructurales, que ocurren en las series de energía diaria, ya que se encontraron evidencias de que algunos de los días de las últimas semanas del año, presentan una conducta diferente, que no podía ser explicada con los mismos argumentos utilizados a lo largo del año según los datos proporcionados por XM expertos en mercados (filial de ISA). Para llevar a cabo este trabajo se realizó un análisis estadístico, haciendo estimación, caracterización y análisis de seis métodos estadísticos para encontrar el que optimizara el pronóstico de la demanda energética diaria de Colombia lo más acertado posible. Estos fueron: los modelos de suavización exponencial, ARIMA, y regresión, dos métodos bayesianos: regresión gaussiana con simulación Monte Carlo por Cadenas de Markov (MCMC), y elicitación bayesiana con simulación de pronóstico. A partir de todo el análisis, se dio una explicación más adecuada del comportamiento de la demanda energética diaria del país.

1. OBJETIVOS

1.1. GENERAL

Optimizar el abastecimiento de energía eléctrica en Colombia, a partir de la estimación eficiente de la demanda, para garantizar adecuados niveles en el manejo de costos y tiempos para el abastecimiento

1.2. ESPECÍFICOS

1. Realizar un análisis de los datos históricos de demanda de energía, en las diferentes épocas del año, identificando variables exógenas que puedan ser significativas para los modelos.
2. Revisar los modelos bayesianos existentes en la literatura, especificando los criterios existentes de estimación.
3. Comparar algunas técnicas clásicas, como los modelos de suavización exponencial, ARIMA, y regresión, con las bayesianas, con el fin de concluir cual es la herramienta más adecuada que permita optimizar recursos para la compañía.
4. Encontrar los niveles adecuados de energía diaria para el país, para la planeación adecuada del abastecimiento con anticipación.

2. ANTECEDENTES

2.1 Antecedentes del sector energético colombiano

El Centro Nacional de Despacho (CND), departamento de XM Compañía de Expertos en Mercados S.A.E.S.P, filial de ISA, es el encargado de la operación y administración de todo el Sistema Interconectado Nacional de Colombia (SIN). Es decir, su tarea es planear, supervisar y controlar los recursos de generadores, transmisores, distribuidores y comercializadores de energía. El CND debe realizar un plan de maniobras para las empresas generadores, indicando la cantidad de potencia que deben producir diariamente, por este motivo, el pronóstico de demanda energética es uno de los suministros más importantes para este proceso, por lo que es realmente necesario una predicción efectiva, garantizando calidad, seguridad y confiabilidad en el servicio de los usuarios. Al respecto, Rueda et al (2010) expresan: *“La predicción de la demanda es un problema de gran importancia para el sector eléctrico, ya que a partir de sus resultados, los agentes del mercado de energía toman las decisiones más adecuadas para su labor”*.

Y precisamente, apuntando al mejoramiento de pronósticos, se han utilizado diferentes tipos de modelos estadísticos clásicos e incluso métodos de inteligencia artificial como las redes neuronales, tratando de garantizar una eficiencia en la predicción de dicha demanda energética. A continuación se hace referencia a los modelos utilizados recientemente y que se consideran más relevantes.

En XM, se utiliza un modelo SARIMA para el pronóstico de demanda diaria de energía, pero sus características son confidenciales. Por otro lado, los modelos utilizados por Codensa S.A ESP, empresa distribuidora y comercializadora de energía eléctrica en Colombia, son: Regresión lineal, suavización exponencial y media móvil, según lo expone Gómez et al (2011), quienes además, realizaron un análisis en el que discriminan cada operador de red por la calidad de pronósticos que realizan teniendo como criterio el MAPE, dando como resultado: Pronóstico con oportunidad de mejoramiento y Pronósticos con indicadores de mala calidad.

En la última década se han realizado investigaciones muy importantes, con relación al pronóstico de energía en Colombia. En el año 2003, Murillos et al, llevaron a cabo un proyecto llamado: “Estudio del pronóstico de la demanda de energía eléctrica, utilizando modelos de series de tiempo”. Esta investigación tenía como finalidad predecir la demanda de energía eléctrica, utilizando modelos ARIMA, basado en datos históricos que van desde el primero de enero de 2001, hasta el 31 de diciembre de 2001 y el paquete estadístico SPSS (*Statistical package for the social sciences*). Su modelo final, luego de realizar los análisis pertinentes fue un SARIMA (1,0,0)(2,1,0).

Por otra parte, en 2010 Arturo et al, realizaron una investigación, la cual consistió en establecer y utilizar un modelo dinámico para el pronóstico de energía horaria para Codensa S.A. ESP, La intención con este modelo era mejorar el margen de error en un 2% aproximadamente, con respecto al modelo utilizado con anterioridad en esta compañía. La metodología utilizada, fue un poco similar a la que se utilizó en este proyecto, los investigadores compararon varias técnicas y modelos estadísticos, con datos históricos desde el 2004, con el fin de encontrar el más indicado para la demanda energética diaria. Analizaron la serie de tiempo, identificando

comportamientos atípicos e intentaron normalizarla ingresando variables ficticias.

Los modelos que utilizaron fueron: Modelo de promedios móviles, Modelo potencial, Modelo ARMAX (Autoregressive moving average model with exogenous inputs model).

Rueda et al (2010), llevaron a cabo, una investigación en la que presentaron un análisis de las técnicas y modelos más utilizados para el pronóstico de demanda de electricidad en varios países y varias regiones de Colombia, planteando problemas o dificultades a las se han enfrentado los investigadores de esta variable. En dicha investigación se obtuvo que los modelos más utilizados son: ARIMA y redes neuronales artificiales (RNA). Los autores concluyen: *“se encontró poca claridad sobre cuál modelo es más adecuado y en qué casos, adicionalmente, los estudios no presentan una recomendación específica para desarrollar modelos de pronóstico de demanda, específicamente en el caso colombiano”*, para finalizar, proponen que se realice un estudio sistemático, para determinar los modelos más apropiados para el pronóstico de demanda de energía.

Medina & García (2004), realizaron una investigación en la que buscaron pronosticar la demanda mensual de energía eléctrica en Colombia. Exploraron dos técnicas de inteligencia artificial, para posteriormente compararlas con otros modelos tradicionales de series de tiempo, consiguiendo errores muy bajos, alrededor de 1,108% con la técnica de Redes Neuronales. A continuación, se revisan antecedentes asociados con modelos de pronósticos.

2.2 Modelos de pronósticos

Los modelos de pronóstico tradicionales han sido ampliamente usados para la toma de decisiones a nivel de producción, finanzas y la energía como se ha mencionado, tal es el caso de los modelos ARIMA, los cuales fueron desarrollados en los años 70 por George Box y Gwilym Jenkins, con la intención de crear una metodología destinada a identificar, estimar y diagnosticar modelos dinámicos de series temporales, involucrando la variable tiempo como la más importante.

Según Maté (2006) la principal ventaja de esta metodología es que permite elegir entre distintos modelos dependiendo del comportamiento de los datos (el que los represente mejor). Sin embargo, la determinación del modelo que mejor se adecua a la serie de datos no es trivial y, por tanto, se requiere que la persona que realice predicciones tenga amplios conocimientos sobre esta metodología. El autor sostiene: *“Esto ha inhibido el uso de esta metodología para realizar predicciones en el mundo de la empresa, ya que el aumento de precisión de las mismas no compensaba el coste de implantación. No obstante, es posible manejar algoritmos automáticos, que permiten que la persona que utilice estas técnicas no tenga que tener conocimientos extensos sobre esta materia”*

Frente a este tipo de dificultades en la modelación, siempre se buscan alternativas que permitan optimizar los pronósticos, como se ha observado en los diferentes trabajos del sector energético. Es así como la combinación de pronósticos cualitativos con los cuantitativos, puede aportar en este sentido.

Los modelos presentados en este trabajo van desde los de suavización exponencial, SARIMA, modelos dinámicos, hasta los modelos bayesianos como el de regresión gaussiana con Monte Carlo por Cadenas de Markov

(MCMC) y simulación por elicitación bayesiana usando el método Delphi.

A continuación se introducen los antecedentes asociados a dichas técnicas.

2.2.1 Suavización exponencial

Según Fornero (2007) el método de suavización exponencial, surgió de la marina de los Estados Unidos durante la segunda guerra mundial, por Robert G. Brown cuando trabajó para el apoyo antisubmarino, desarrollando un modelo para el control de la ubicación de los submarinos y el seguimiento de la información. Durante la década del cincuenta, amplió la suavización exponencial simple de datos discretos y desarrolló métodos para obtener tendencias y estacionalidad. *"El ahorro en el almacenamiento de datos a través de medias móviles condujo a la adopción de la suavización exponencial a través de los sistemas de inventario Marina durante la década de 1950"*. Su primera aplicación fue el pronóstico de demanda de piezas de repuesto con el fin de controlar el inventario en la Marina.

Gardner (2006), realizó un estudio con el fin de determinar el estado en el que se encuentra la técnica de suavizamiento exponencial, afirmando que se ha avanzado poco en la identificación y selección de los métodos de suavización exponencial y que es difícil aplicar adecuadamente de una tendencia amortiguada para cada serie de tiempo.

2.2.1.1 Suavización exponencial simple

La suavización exponencial simple es una técnica útil para suavizar una serie de tiempo, además es utilizada para conseguir predicciones a corto plazo. Se caracteriza por dar un mayor peso a los últimos valores de la serie y un menor valor a los primeros. (Aguilera, 2009)

Bellas & Martinez (2005) desarrollaron una metodología de series de tiempo para el área de la salud, cuyo objetivo es analizar el número de consultas externas realizadas en la institución prestadora de servicios de salud de la universidad de Antioquia (IPS Universitaria), debido a que la institución ha incrementado su prestación de servicios como también el número de usuarios en los últimos años, por lo que es importante saber que tanto ha cambiado el comportamiento de estas variables en el tiempo.

Para este caso de estudio el número de consultas externas no es muy cambiante, por esto se procedió a estimar el modelo mediante el método de suavización. La suavización exponencial simple mostró que el pronóstico es muy influenciado por el último mes de la serie, donde las consultas han incrementado, sin embargo se llega a la conclusión de no realizar proyecciones de más de 6 meses ya que se introducirían muchos datos y tienden a robustecer el modelo.

En Venezuela se hizo un estudio por parte de Guerra, Sánchez, & Reyes, (1997) con el objetivo de predecir mediante modelos de series de tiempo la inflación en dicho país. Por lo que se eligieron períodos de un mes con el propósito de hacer predicciones a corto plazo. Mediante la suavización exponencial se obtiene una ecuación que permite describir el proceso estocástico que ha sufrido la variable en un pasado reciente como un

promedio ponderado. Luego se presentan los resultados mediante un modelo ARIMA, y finalmente se estima un modelo que combina diversas técnicas de series de tiempo para estimar la inflación.

En un modelo de suavización exponencial simple es importante definir un valor adecuado de la constante de suavización (α); Castro & Botero (2012) propusieron una metodología para la selección del parámetro α en un modelo de suavización exponencial, con el fin de que el modelo escogido realice optimas predicciones. Además de aplicar la metodología en un caso industrial con datos reales.

La metodología propuesta fue la siguiente (pasos):

1. Calcular coeficiente de variación y ordenar
2. Pronosticar utilizando el modelo SES (Suavización exponencial simple)
3. Calcular medidas para posibles valores de α
4. Buscar “óptimos”
5. Generar rangos de α alrededor del optimo
6. Separar series
7. Graficar cada grupo y determinar rango común
8. Emplear los rangos encontrados en la validación

Aplicando esta metodología a 58 series mensuales cada una con 133 observaciones del sector industrial, se pudo identificar que a medida que los coeficientes de variación aumentan, los rangos de α que disminuyen el MAPE incrementan también.

2.2.1.2 Suavización exponencial lineal: método Holt

Algunas series temporales, a ser estudiadas, no son estacionales pero con tendencia lineal, que puede ser creciente o decreciente. La metodología ideal para estos casos es la de Holt (Esparza, 2000).

Zuluaga , Molina, & Guisao, (2011) realizaron un estudio sobre planeación de la demanda, como requisito para la gestión de las cadenas de suministro de las empresas en Colombia. Partiendo del hecho que la logística y las cadenas de suministro de una empresa deben ser tomadas como ventajas competitivas, si las empresas quieren agregar valor a sus procesos y productos, se debe comenzar por proponer buenas estrategias de planeación. El estudio presenta el análisis de diferentes modelos, entre ellos el de suavización exponencial lineal, y como estos deben ser elegidos dependiendo del comportamiento de los mercados. Para seleccionar el modelo más adecuado, se debe definir claramente cuál va a ser la serie a estudiar, posteriormente se ejecutan los modelos mediante un software (El método Holt está entre los que se ejecutaron). Se deben caracterizar los atributos de los modelos, según su dificultad de implementación, horizonte de tiempo y características básicas. Una vez estudiadas las características se elige el mejor modelo. Llegando a la conclusión *“El uso de pronósticos y el análisis de sus históricos, permite la planificación del aprovisionamiento: La integración de los procesos, la alineación de terceros y proveedores, culminando con la estrategia de planeación de ventas y operaciones”*

Cabanzo (2011) realizó un plan de requerimiento de distribución para una empresa de plásticos, con el fin de desarrollar un sistema de planeación para ayudar a disminuir los pedidos rechazados por los clientes e implementar un modelo de producción para tal empresa. Se pretende lograr que la empresa no base sus tiempos de respuesta a los clientes en la experiencia de los dueños, sino que se basen en un modelo ajustado a las condiciones de la compañía.

Aplicando los métodos de suavización exponencial simple y suavización exponencial lineal, se obtuvieron diferentes pronósticos de demanda. Para la suavización exponencial simple se presentó el error medio en desviaciones

medias absolutas en promedio por periodo. Se llegó a la conclusión de que este método no se ajustó a los datos históricos de demanda de la compañía. *“Dado que el pronóstico de demanda obtenido con el método de suavización exponencial doble tiene una tendencia lineal y presenta un error muy alto, esta técnica no se ajusta a los datos históricos de la demanda”*. (Cabanzo, 2011)

2.2.2 Modelos ARIMA y SARIMA

Los modelos ARIMA, incorporan características del pasado de la misma serie. Lo que pretende generalmente este modelo es convertir la serie en una estacionaria por medio de unas diferenciaciones sobre la misma (Chavez, 1997). Dichos modelos tienen aplicación en numerosos campos, como la industria manufacturera, usados para el análisis de variables como la demanda de producto terminado, similar al estudio de demanda energética, en especial si ambas son continuas o tienen tamaños que permiten su aproximación a la distribución normal, como se relaciona en las siguientes aplicaciones.

En la universidad de los Andes, Medina & Mutis (2004) usaron la metodología ARIMA para pronosticar la demanda de productos de consumo masivo para la limpieza del hogar en el mercado colombiano, encontrándolo adecuado para pronosticar el consumo mensual de estos productos, usando los patrones: diario, semanal y mensual.

Dentro de estos modelos también se conoce el modelo ARIMA estacional, o modelo SARIMA, cuyo objetivo es captar el comportamiento estacional de una serie (Chavez, 1997). En el trabajo de Contreras et al (2009) se propone un modelo SARIMA para predecir la tasa de desempleo en Chile, ya que para

cualquier país es de suprema importancia pronosticarla y desarrollar estrategias que permitan disminuir el índice. Luego de hacer varios intentos se seleccionó el mejor modelo SARIMA, con el criterio de selección BIC (criterio de información Bayesiana) y se concluye que las fluctuaciones que puede tener el entorno laboral son significativas (Contreras & Idrovo , 2009).

2.2.3 Metodología Delphi

Dentro del campo de la estadística existen diversos métodos probabilísticos que permiten predecir el comportamiento de variables continuas en un tiempo determinado, facilitando o ayudando a la oportuna toma de decisiones dentro del ámbito empresarial. Las empresas manufactureras tienen variables como la demanda, con características muy similares a las del sector energético, por ejemplo su continuidad en el tiempo y algunos patrones estacionales. Desde hace algunas décadas se viene implementando y demostrando la utilidad en ciertas estimaciones de técnicas y modelos estadísticos basados en el conocimiento subjetivo a priori, que no requieren la toma de muchos datos, en comparación con los modelos clásicos donde sólo permite probabilidades fundamentadas en experimentos repetibles y validaciones, pero que igual se aproxima a fenómenos reales; un ejemplo de esto es la estadística bayesiana (Valencia & Sierra, 2009).

Uno de los métodos que basan su inferencia en datos netamente subjetivos es el Método Delphi. Según Pettyjohnn (2012) en una investigación asociada a la empresa RAND Corporation, pionera en el desarrollo de este método, su desarrollo e iniciativa se dio por el interés de pronosticar el impacto de la tecnología en la guerra y consiste en que un grupo de expertos respondan a un cuestionario de forma anónima, entregando con la “respuesta colectiva”

resultados cuantitativos y cualitativos, para después representarlos estadísticamente. Según Landeta et al (2011) el método Delphi es una herramienta que ayuda a la investigación para obtener datos confiables de un grupo de expertos y en varias ocasiones puede proporcionar solución a problemas complejos.

El desarrollo del método Delphi se ha trabajado en varias áreas del conocimiento. Bijak & Wiśniowsk (2010) realizaron una investigación en siete países europeos con el fin de pronosticar, mediante modelos Bayesianos, el número de inmigrantes que estos países podrían tener en el año 2025. Se basaron en datos cuantitativos y conocimientos cualitativos partiendo de la elicitación a expertos en el tema de cada país, y argumentaron: *“Bajo la no estacionariedad de los procesos migratorios, el horizonte para pronósticos demasiado largos son insuficientes, lo que es un desafío serio para los pronósticos de población en general”*. Pero, encontraron útil la información de expertos para pronósticos a corto plazo.

Ji & Chen (2012) realizaron una investigación que buscaba evaluar el riesgo en la cadena de suministro y un análisis empírico para las empresas procesadoras de ropa; recolectaron opiniones de expertos a través del método Delphi, y basándose en la teoría de riesgo de Markowitz, los resultados mostraron que el sistema de evaluación de riesgos no sólo se puede utilizar para identificar y evaluar el riesgo de esta empresa a disposición, sino que también se pudo descubrir, que el método de riesgo de dicha teoría puede ser cuantificado con los sentimientos intuitivos (Delphi), y que combinado con la teoría de la utilidad, refleja realmente el comportamiento de los tomadores de decisiones y la preferencia de riesgo, lo que podría ser utilizado como referencia para los costos de insumos de la gestión del riesgo de la empresa.

En Colombia, García (2008) hizo una aplicación de éste, formulando una metodología para la estimación de demanda en accesorios de joyería, de manera que usa la variable: ventas y su distribución de probabilidad elicitada para efectuar el pronóstico, técnica similar a la propuesta en este trabajo. Uno de los objetivos era usar el método Delphi para indagar la demanda con expertos, ya que no se conocen datos históricos de dicha variable. Fue necesario entonces realizar un estudio detallado del comportamiento de los consumidores con respecto a varios productos y a su vez realizar distintos cuestionarios.

Los resultados mostraron además, que el método Delphi sí permite incorporar variables de tipo cualitativo en el estudio de la demanda, que es cuantitativa, por medio de la interacción del grupo de expertos sobre temas como los gustos, preferencias y precios de los consumidores.

Y además resalta que *“Los sistemas clásicos requieren un número mínimo de datos para la previsión, pero teniendo en cuenta que un producto tiene una frecuencia de aparición inferior al número de campañas por año, la información se hace insuficiente. Así el método Delphi se justifica como herramienta de previsión, pues ante una situación de escasez de información, se recomienda el uso del juicio de los expertos”* (García, 2008).

2.2.4 Modelo Econométrico Dinámico

El estudio de la econometría tiene como principal objetivo predecir acontecimientos económicos y sociales en el tiempo, por ejemplo, permite

pronosticar la tasa de empleo de un país y la productividad de una empresa entre otros. Por medio de la estimación de modelos econométricos dinámicos, también llamados modelos regresivos dinámicos, o modelos dinámicos con variable endógena y exógena, se caracterizan variables respuesta o dependientes con respecto a sus valores pasados en el tiempo (Andrade et al, 2004), o variables exógenas a esta.

Marshall (1986) planteó un modelo econométrico dinámico para la inflación mensual chilena en el periodo de 1977-1983. Se parte del supuesto que la economía chilena es abierta y pequeña, que en todos sus procesos de producción tienen competencia, y poseen dos bienes: uno transable y uno no transable. El bien transable parte del hecho de que se tiene un único precio, mientras que el bien no transable se determina de acuerdo a las condiciones de oferta y demanda. Para la construcción del modelo dinámico *“se supone que los salarios son determinados exógenamente por la autoridad económica”* es decir con el bien no transable. Se definió la variable respuesta como inflación interna, analizando el modelo estimado se llega a la conclusión que la inflación depende de los excesos de demanda, cambios en los salarios e inflación externa.

También en Chile se construyó un modelo dinámico por parte de Gonzáles et al (2007) para explicar las diferencias en los precios de una canasta de alimentos, tomando una muestra de 24 ciudades y analizando el impacto que tiene la industria de los supermercados sobre los precios. El modelo estimado tiene como variable respuesta el precio de equilibrio, que depende de unas variables exógenas como lo son la ciudad o el periodo de tiempo. Por lo que se ajustan los precios en cada instante de tiempo.

Los resultados mostraron que los precios de los productos alimenticios si se ven afectados por el tipo de ciudad, incluso en algunas hasta de un 8%. *“La*

entrada de un hipermercado en una ciudad disminuye los precios de los alimentos. Nuestros resultados son perfectamente consistentes con este resultado. La entrada de un hipermercado aumenta el número de competidores en el mercado local, disminuyendo la concentración y los precios". (González & Gómez, 2007)

El estudio solo permitió estudiar los productos alimenticios ya que solo se disponía de esta información, sin embargo, estos productos representan la mayoría de la oferta de un supermercado.

Roca & Pereira (1998), estimaron un modelo econométrico dinámico de vectores autorregresivos, que permitiera evaluar el impacto de la inversión en infraestructuras de transportes y comunicaciones, sobre el comportamiento de la economía española. El autor lo denomina con un "shock", es decir el impacto que tienen este tipo de inversiones sobre variables como: producto interno bruto, capital de trabajo y capital privado.

La serie utilizada cubre el periodo de 1970-1991 de la economía española, y analiza el impacto sobre las variables ya mencionadas anteriormente. Los resultados arrojados permitieron concluir que la inversión pública en España sí afecta de manera positiva el crecimiento del producto interno bruto, mientras que el crecimiento de la variable inversión privada a largo plazo no parece ser muy relevante.

Si bien los modelos y métodos presentados son apropiados para realizar pronósticos de variables cuyas características continuas y distribucionales son muy similares, en este trabajo se considerarán sólo seis de estos que se muestran en la tabla 1. Además, en el estado del arte se presentaron técnicas como las Redes Neuronales, aplicadas para pronosticar energía por los autores Medina & García (2004), que son muy complejas y no constituyen

parte de los objetivos de este trabajo de grado.

En la tabla 1 se muestra una síntesis de las fortalezas y debilidades encontradas en los antecedentes sobre los modelos y métodos de estimación de pronósticos de demanda en general, algunos específicos alrededor de la energía, sin embargo, se presentan los utilizados para el logro de los objetivos de este trabajo.

Tabla 1. Resumen de los antecedentes de modelación y métodos de pronóstico

Método	Fortalezas encontradas	Debilidad encontradas
Suavización Exponencial	Es una técnica útil para conseguir predicciones a corto plazo, se caracteriza por dar mayor peso a los últimos valores de la serie y menor valor a los primeros (Aguilera, 2009)	Se ha avanzado muy poco en la identificación y selección de métodos de suavización exponencial. Es difícil su aplicación para series de tiempo (Gardner, 2006)
SARIMA	Es una técnica que tiene aplicaciones en distintos campos como: la industria manufacturera, la tasa de desempleo de un país y su crecimiento económico (Chávez, 1997)	Exige el cumplimiento de premisas de incorrección y normalidad, en muchos casos no se cuenta con la cantidad de datos requeridos para el desarrollo de estos modelos (Cohen, 1986)
	Convierte una serie de tiempo en una serie estacionaria, además capta su comportamiento estacional (Chávez, 1997)	
Modelo Económico Dinámico	Permite incorporar variables exógenas y endógenas, con respecto a sus valores pasados en el tiempo (Andrade et al, 2004)	Exige el cumplimiento de premisas de incorrección y normalidad, en muchos casos no se cuenta con la cantidad de datos requeridos para el desarrollo de estos modelos (Cohen, 1986)
	Es una técnica que permite predecir acontecimientos económicos y sociales en el tiempo, como la tasa de empleo de un país y la productividad de una empresa. (Andrade et al, 2004)	
Método Delphi	Es una técnica basada en el conocimiento subjetivo a priori, lo que no requiere la toma de muchos datos como si lo exigen los modelos clásicos (Valencia & Sierra, 2009)	El uso de esta técnica para predecir acontecimientos a largo plazo no es muy confiable, lo cual se convierte en un reto para el desarrollo de pronósticos. (Bijak & Wiśniowski, 2010)
	Soporta la investigación para obtener datos confiables de un grupo de expertos y puede proporcionar solución a problemas complejos (Landeta, 2011)	

3. MARCO TEÓRICO

A continuación se hará referencia a las definiciones utilizadas en este trabajo de tesis, con el fin de dar un mejor entendimiento técnico para el buen desarrollo de la investigación.

3.1 Pronósticos

Es la estimación del valor de una variable, realizándose de forma anticipada en un determinado lapso de tiempo (Hanke & Arthur, 1996).

Vera & Bustamante (2007) afirman que: *“el pronóstico no es una predicción de lo que irremediablemente pasará en el futuro, sino una información que se obtiene con cierto grado de error (dado por una probabilidad) de lo que pudiera pasar”*.

Los pronósticos se pueden caracterizar o clasificar de varias maneras como se muestra en la tabla 2.

Tabla 2 Resumen de la clasificación

Plazo	Corto plazo
	Largo plazo
Según el tipo de proceso a seguir (datos a utilizar)	Cuantitativo
	Cualitativo
	Mixto
Entorno	Micro
	Macro

A continuación se dará una pequeña explicación (Hanke & Arthur, 1996):

1. Dependiendo del plazo en el tiempo. Es decir, si se desean conocer valores a corto o a largo plazo. Los pronósticos a corto plazo se realizan por lo general para menos de 3 meses y son utilizados para realizar estrategias que requieren ser diseñadas en un lapso inmediato, y así ejecutar rápidas tomas de decisiones, y por otro lado, los pronósticos a largo plazo, son predicciones en un margen aproximado de 5, 10 y hasta 20 años, con fines de establecer el curso de una organización de una forma general, que permita una planeación para estar preparados a eventuales cambios.
2. Si son cualitativos o cuantitativos, o ambos. Cuando se realizan predicciones netamente cualitativas, no es necesaria una gran manipulación de datos, se utiliza básicamente el juicio de un experto, a partir de una manipulación mental. Por otro lado, al realizar técnicas netamente cuantitativas, se prescinde de los juicios por parte de expertos, y su análisis por lo general, es matemático. Pero también, hay un punto medio, pronósticos en los cuales se utilizan datos cualitativos y datos cuantitativos, por medio de la elicitación a expertos, la cual da como resultado, observaciones cuantitativas para el cálculo estadístico, y las cualitativas para justificar sus criterios (Bravo & Arrieta, 2005).
3. Dependiendo del entorno que se desea pronosticar, (micropronóstico o macropronóstico) Hanke (1996) planteó un ejemplo para una mejor comprensión: *“El gerente de una planta pudiera estar interesado en*

pronosticar el número de trabajadores que requerirá en los próximos meses (un micropronóstico), mientras que el gobierno federal está pronosticando el número total de personas empleadas en toda la nación (un macropronóstico)”.

3.2 Definición del método Delphi

El método Delphi es un proceso que consiste en la selección de un grupo de expertos sobre un tema específico, en el cual se hacen preguntas o se piden opiniones sobre posibles hechos que puedan suceder en un futuro no muy lejano. Es una técnica de investigación social que tiene como objeto la obtención de una opinión grupal.

Este método arroja resultados cuantitativos y cualitativos, dependiendo de la variable a ser estudiada, y por lo tanto puede ser aplicado en estudios exploratorios, predictivos y de consenso. Las apreciaciones de los expertos se realizan en diferentes rondas anónimas por medio de cuestionarios, con la característica que, a partir de la segunda ronda en adelante, los resultados se entregan, es decir, se dan a conocer a los expertos para que les sirvan de retroalimentación en el proceso.

Al transcurrir una ronda, los expertos pueden discutir sobre sus respuestas, llegando a la próxima ronda influenciados por sus compañeros, haciendo del Delphi un proceso de comunicación, ya que la información puede estar dispersa o incompleta.

La calidad de los resultados depende mucho de cómo se haya realizado la metodología, de la correcta selección de los expertos, de la formulación de las preguntas en el cuestionario, haciendo entender al elicitado qué es exactamente lo que se pretende conocer, y de la retroalimentación entre los expertos que permita generar nuevos conocimientos.

Una definición corta pero precisa es la que proponen Liston y Turoff citados por Landeta (2008): *“El Método Delphi es un método de estructuración de un proceso de comunicación grupal que es efectivo a la hora de permitir a un grupo de individuos, como un todo, tratar un problema complejo”*.

Otra definición es la que propone Parisca (1995) considera que el método Delphi: *“Se basa en el principio de la inteligencia colectiva y que trata de lograr un consenso de opiniones expresadas individualmente por un grupo de personas seleccionadas cuidadosamente como expertos calificados en torno al tema, por medio de la iteración sucesiva de un cuestionario retroalimentado de los resultados promedio de la ronda anterior, aplicando cálculos estadísticos”*.

Sin embargo para Konow y Pérez (1990) establecer una definición sobre el método es limitar su alcance, por tanto, ellos aconsejan que es mejor realizar una descripción de sus características, restricciones y aplicaciones.

Dalkey propuso cuatro tipologías para establecer al método Delphi como una técnica con características propias:

1. Proceso iterativo: Los participantes del proceso, es decir, los expertos, tendrán que dar su opinión en varias ocasiones, con el fin de estabilizar las respuestas. Esto permite que el experto reflexione y tenga la

oportunidad de reconsiderar su opinión, basándose en las opiniones de otros participantes.

2. Anonimato: El objetivo del anonimato es que ningún miembro del grupo sabe a quién corresponde una respuesta específica, esto ayuda a evitar influencias negativas entre los expertos. El control y manejo de la información siempre tiene que estar manipulado por el grupo encargado, y se recomienda que los cuestionarios se ejecuten de manera escrita, telefónica o vía internet.
3. Realimentación o *feedback* controlado: Antes de comenzar cada ronda, el grupo encargado da a conocer la posición de los expertos, destacando algún aporte que se considere relevante. Así, la realimentación por parte del grupo encargado, permite la circulación de la información y ayuda a configurar un lenguaje común entre los expertos.
4. Respuesta estadística de grupo: Dado el caso en que al grupo se le haya solicitado una estimación numérica, se maneja la mediana de las respuestas individuales. “Aunque se promueva un consenso, este no es siempre el objetivo último y no necesariamente se alcanza. El rango intercuartílico de las estimaciones será el indicador de dispersión de las respuestas, que se considera un nivel de consenso.” (Valencia & Sierra, 2009).

El grupo encargado, o grupo coordinador, debe estar conformado por un reducido número de integrantes conocedores del Delphi y del caso de estudio. Con el fin de poder ejecutar el método correctamente, interpretando las opiniones y respuestas brindadas por parte del grupo de expertos.

Premisas:

- En situaciones de incertidumbre o cuando no se dispone de suficiente información objetiva, es conveniente el juicio subjetivo de expertos en disciplinas que no son exactas.
- Contar con un solo juicio subjetivo de un experto, es arriesgado porque se puede sesgar la información, además, al aferrarse a la experiencia de una sola persona puede arrojar estimaciones imprecisas.

“En algunos casos, los juicios de expertos individuales pueden conducir a sesgos, imprecisiones, o malas estimaciones por la falta de conocimiento de la persona” (Valencia & Sierra, 2009)

- El juicio subjetivo grupal presenta mayor calidad que el juicio subjetivo individual, debido a que el grupo, en conjunto, dispone de más información.

Tipos de pronósticos:

- Delphi convencional: Está compuesto por un cuestionario enviado a un grupo de expertos, posteriormente se envía un segundo cuestionario basándose en los resultados del primero. Para terminar los cuestionarios se complementan midiendo la exactitud en el criterio del grupo de expertos.
- Delphi de tiempo de real: Este se identifica porque su desarrollo es muy corto, se ejecuta en el transcurso de una reunión, y por medio de unos elementos, se resumen las respuestas dadas por los expertos rápidamente.

- Delphi político: “Es una herramienta de análisis de políticas alternativas y no un mecanismo de toma de decisiones” (Valencia & Sierra, 2009).

En este tipo de pronóstico no se toma una decisión, el grupo de expertos debate sus opiniones respecto a un problema, y se aseguran que todas las opiniones se hayan tenido en cuenta para medir el impacto que puedan generar.

3.3 ARIMA

La expresión general de un modelo ARIMA significa (*Autoregressive Integrated Moving Average*) que traduce: modelo autorregresivo integrado de media móvil. Para establecer la expresión algebraica de un modelo ARIMA se debe partir de la expresión algebraica de un modelo ARMA (*Autoregressive Moving Average models*).

Los modelos ARIMA o modelos Box y Jenkins proponen varios métodos para estimar los inconvenientes que se mencionaron anteriormente, aportando soluciones óptimas a corto plazo, permitiendo elegir entre diferentes modelos el que mejor muestre el comportamiento de los datos. (De Arce & Mahía , 2007)

Un modelo ARMA tiene la característica de ser lineal, esto quiere decir que la variable que define la serie temporal depende de una constante. Sea $Y(t)$ una serie temporal su expresión algebraica está dada por (Maté, 2006) como :

$$Y_t = C + \phi_1 * Y_{t-1} + \dots + \phi_P * Y_{t-P} + \theta_1 * \varepsilon_{t-1} + \dots + \theta_q * \varepsilon_{t-q} + \varepsilon_t$$

La serie temporal tiene una dependencia con los valores pasados de la misma serie, esto en el modelo se conoce como: componente autorregresiva (AR), el número de datos pasados que se incluyen en la serie temporal $Y(t)$ se conocen como: orden autorregresivo y se simboliza mediante la letra p . así mismo la serie temporal tiene una dependencia con los valores pasados de los errores, en el modelo se conoce como: componente de media móvil(MA), el número de datos de errores pasados que se incluyen en $Y(t)$ se conocen como: orden de media móvil y se simboliza mediante la letra q . Así un modelo ARMA se escribe abreviadamente: $ARMA(p,q)$. (Robalino , 2005)

Conociendo como es la composición de un modelo ARMA, se puede definir la de un modelo ARIMA, que se construyen a partir de estos, pero teniendo en cuenta que para la serie ser estacionaria, tiene que diferenciarse una serie d veces.

Un modelo ARIMA es un modelo $ARMA(p,q)$ sobre la serie diferenciada d veces. Se escribe abreviadamente $ARMA(p,d,q)$ y su expresión algebraica está dada por:

$$Y_t^{(d)} = C + \phi_1 * Y_{t-1}^{(d)} + \dots + \phi_p * Y_{t-p}^{(d)} + \phi_1 * \varepsilon_{t-1}^{(d)} + \dots + \phi_q * \varepsilon_{t-q}^{(d)} + \varepsilon_t^{(d)}$$

El parámetro d indica el número de veces que una serie debe ser diferenciada con el fin de volverla estacionaria.

Donde $Y(t)^{(d)}$ es la serie de las diferencias de orden d , y $E(t)^{(d)}$ es la serie de los errores que se cometieron en la serie anterior o los residuos del modelo. (Maté, 2006)

La parte más importante de estos modelos, es la identificación de los parámetros p, d, q , que pueden tomar cualquier valor, casi siempre entre 0 y 1, y en muy raras ocasiones 2. Lo primero en identificar es el grado de diferenciación para que la serie sea estacionaria (d). Se analiza entonces, gráficamente si la serie es estacionaria, es decir que esta tenga la misma media y varianza, de no ser así, se diferencia la serie y se vuelve a analizar gráficamente. Por tanto el valor d será el número de veces que se haya diferenciado (Robalino , 2005)

Según Villavicencio (2010), para conocer los valores de los parámetros p y q se tiene en cuenta la función de autocorrelación, esta mide la dependencia de las variables sobre valores pasados, ya que en una serie de tiempo en muchas ocasiones los valores que toma cierta variable no son independientes entre sí. Se recurre entonces a la función de autocorrelación ACF y a la función de autocorrelación parcial PACF.

La autocorrelación ACF mide la correlación que existe entre dos variables que están separadas por n periodos.

La autocorrelación parcial PACF mide la correlación que existe entre dos variables que están separadas por n periodos cuando no se considera la dependencia creada por los retrasos intermedios existentes entre ambas.

El componente u orden autoregresivo (p) viene ligado a la definición de autocorrelación. Si dos momentos continuos están correlacionados se dice que existe correlación de orden 1, y se simboliza ARIMA(1,0,0). Si se presenta el caso de que la serie depende de los dos momentos continuos se obtiene un modelo ARIMA(2,0,0).

El valor del componente de media móvil (q) de la serie, es establecido por los componentes aleatorios. El parámetro q expresa la cantidad de componentes aleatorios que forman el valor presente en la serie temporal.

Por otra parte Maté (2006) sugiere que un modelo ARIMA óptimo debe de cumplir con ciertas características que permiten identificar, estimar y verificar el comportamiento del modelo:

- Parquedad: En el modelo no se deben incluir coeficientes innecesarios. Es importante porque permite obtener mejores predicciones, y ayuda a estimar un modelo óptimo que se aproxime estadísticamente a la serie que se está evaluando.
- Estacionariedad: Al asegurar que la serie sea estacionaria, ósea que su media es constante, se podrán utilizar N observaciones para estimarla.
- Estimación de buenos coeficientes: Una buena estimación es que los coeficientes de la componente autorregresiva (AR) y la componente de media móvil (MA) sean significativamente diferentes de cero.
- Buen ajuste de los datos: El modelo se debe ajustar en lo posible a los datos de los cuales ha sido generado. No existe un criterio que indique el buen ajuste, pues todas las series se comportan de maneras diferentes.
- Se obtengan buenos pronósticos: Si bien el modelo se ha ajustado lo suficientemente bien a los datos del pasado, lo que se espera de este es que haga predicciones del futuro consistentes, para esto se recomienda evaluar el modelo varias veces.

Box y Jenkins también diseñaron un modelo que permite captar el comportamiento estacional de una serie. Este tipo de modelo se conoce como modelo SARIMA. Según Arce & Mahía (2003), cuando se desea

analizar la estacionariedad y la estacionalidad de una serie se procede a estimar un modelo multiplicativo ARIMA, que también se conoce o denota como modelo multiplicativo ARIMA(p,d,q)*SARIMA(P,D,Q)_s y está definido algebraicamente por:

$$Y_t = \varphi Y_{T-1} + \varphi Y_{T-2} + \dots + \varphi_{P_S+p+D_S+d} Y_{T-P_S-p-S_D-d} + \delta + U_T + \theta_1 U_{T-1} + \dots + \theta_{Q_S+q} U_{T-S_Q-q}$$

3.4 Suavización exponencial

La suavización exponencial, es una técnica de análisis de pronóstico de series de tiempo, que tiene como característica la predicción de los valores futuros en función de la ponderación exponencial de los periodos anteriores, teniendo mayor peso los periodos recientes que los antiguos (Gardner, 2006). Además, es un modelo en el que se puede incorporar fácilmente el nivel, la tendencia y la estacionalidad que presentan los históricos de la serie temporal (Hyndman et al, 2008).

Según Bowerman et al. (2007) la suavización exponencial es más eficaz cuando la tendencia y la variación estacional de las series pueden manifestar cambios en el tiempo.

3.4.1 Suavización exponencial simple

Modelo de suavización simple (Hanke & Reitsch, 1996)

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

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

α = constante de suavización.

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

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

3.4.2 Modelo suavización exponencial lineal: método Holt

Existen otros métodos de suavización, que en algunos casos, son más efectivos para pronosticar, como el Suavizamiento Exponencial Lineal (Holt), este involucra tendencias lineales locales que se van desarrollando dentro de la serie de tiempo. Además, suaviza directamente el nivel y la pendiente, usando diferentes constantes de suavizamiento para cada una, por lo tanto, tienen la ventaja de adaptarse y ser flexible a nuevas observaciones que se van encontrando en la serie temporal. En este sentido, el suavizamiento exponencial simple se puede ver en desventaja, ya que en las series de tiempo en ocasiones puede cambiar ciertos niveles, entonces su modelación puede llevar un poco más de tiempo. (Ramírez, González & Cardona, 2012).

Modelo de suavización exponencial de Holt (Hanke & Reitsch, 1996)

$$A_t = \alpha Y_t + (1 - \alpha)(A_{t-1} + T_{t-1})$$

Tasa de crecimiento o estimación de la tendencia

$$T_t = \beta(A_t - A_{t-1}) + (1 + \beta)T_{t-1}$$

El pronóstico de P periodos en el futuro

$$\hat{y}_{t+P} = A_t + PT_t$$

En donde:

A_t = nuevo valor atenuado

α = constante de suavización de los datos ($0 < \alpha < 1$)

Y_t = nueva observación o valor real de la serie, en el periodo t

β = constante de suavización de la estimación de la tendencia ($0 < \beta < 1$)

T_t = estimación de la tendencia

P = periodos a pronosticar en el futuro

\hat{y}_{t+P} = Pronóstico de p periodos en el futuro

3.4 Definición modelo dinámico

Los modelos dinámicos son muy utilizados para analizar la evolución del empleo, la productividad en las empresas y otras variables que tienen gran impacto en el desarrollo económico de un país.

Según Caridad & Ocerin (1998) un modelo econométrico está formado por distintas ecuaciones, en la que la variable endógena, es decir la variable explicada, depende de una o más variables explicativas que también se nombran exógenas. El estudio de modelos dinámicos se caracterizan por porque analizan la trayectoria de las variables en distintos instantes de tiempo; las interrelaciones entre estas, casi siempre involucran retardos o desfases. Estos lapsos de tiempo también son conocidos como rezagos. Su estimación comprende la teoría de máxima verosimilitud estudiada para los modelos de regresión comunes (Montgomery, Peck & Vining, 2012)

A continuación se definirán dos tipos de modelos con retardos según (Guisán, 2007):

- Modelos dinámicos de rezagos distribuidos: En este modelo la variable a rezagar es una variable exógena. Definición algebraica:

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \varepsilon_t$$

- Modelos dinámicos autorregresivos: En este modelo la variable a rezagar es la variable endógena. Definición algebraica:

$$Y_t = \alpha + \beta_0 X_t + \gamma Y_{t-1} + \varepsilon_t$$

Modelo de rezagos distribuidos en el tiempo

$$Y_t = \alpha + \beta_0 X_t + \beta_1 X_{t-1} + \beta_2 X_{t-2} + \dots + \beta_k X_{t-k} + \varepsilon_t$$

Dónde:

β_0 es el multiplicador que mide el impacto a corto plazo

$\beta_1 + \beta_2 + \beta_3$ Indican el impacto en un plazo intermedio

$\sum \beta_i = \beta_0 + \beta_1 + \beta_2 + \dots + \beta_k$ indica el multiplicador de rezagos distribuidos a largo plazo o total.

γ es el efecto del retardo de la variable endógena sobre la respuesta en el tiempo t.

3.5 Pruebas de Hipótesis

Según Hanke & Reitsch (1996) en diversos problemas estadísticos, su resultado se centra en afirmaciones sobre los datos observados, y no por la estimación de algunos parámetros. A este procedimiento se le denomina prueba de hipótesis y comprende los siguientes pasos:

1. Establecer la hipótesis que se desea probar, esta es la hipótesis nula y se simboliza con H_0 y luego se establece la hipótesis alterna, a la que se llega en conclusión si se descarta H_0 , el símbolo es H_1 .
2. Suponer que la hipótesis nula es cierta y luego se consulta la distribución muestral de la cual se extrajeron las estadísticas de la muestra bajo esta suposición.
3. Se reúne la muestra aleatoria de elementos de la población, o se establece la serie de datos históricos (observaciones), se realizan las mediciones y posteriormente se calcula las estadísticas correspondientes a la muestra.
4. Calcular la probabilidad de que tal estadística de muestra pudiera haber sido extraída de esta distribución muestral.
5. Dependiendo de la probabilidad que se obtenga y estableciendo el nivel de significancia (α), se rechaza o se acepta H_0 .

3.5.1 Prueba de normalidad de Shapiro-Wilk

La prueba de Shapiro-Wilk determina si los residuales se pueden modelar adecuadamente por una distribución normal. La prueba se basa en la comparación de los cuantiles de la distribución normal ajustada a los cuantiles de los datos.

El estadístico del test es (S. Shapiro & M. Wilk, 1965):

$$W = \frac{(a' * y)^2}{S^2} = \frac{\sum_i (a_i * y_i)^2}{\sum_i (y_i - \bar{y})^2}$$

donde y_i es el número que ocupa la i -ésima posición en la muestra, \bar{y} es media muestral y las constantes a_i están dadas por:

$$a' = (a_1, \dots, a_n) = \frac{m'V^{-1}}{(m'V^{-1}V^{-1}m)^{\frac{1}{2}}}$$

$$m = (m_1, \dots, m_n)^T$$

m_1, \dots, m_n son valores medios del estadístico, de variables muestreadas de distribuciones normales. V matriz de covarianzas del estadístico.

Ho: La distribución de la población de la cual proviene la muestra es normal.

H1: La distribución de la población de la cual proviene la muestra no es normal.

Si el p-valor es menor a alfa (nivel de confianza) entonces la hipótesis nula es rechazada (se concluye que los datos no vienen de una distribución normal).

3.5.2 Prueba de Jarque Bera

Esta es una prueba de bondad de ajuste, la cual determina si los datos de muestra tienen la asimetría y la curtosis que coincide con una distribución normal.

Jarque & Bera (1987) dicen: “Las pruebas sugieren óptimas propiedades asintóticas de la recolección y el buen desempeño de la muestra finita. Debido a su simplicidad, debe llegar a ser una herramienta útil en el análisis estadístico”. El estadístico es:

$$JB = \frac{n}{6} \left(S^2 + \frac{1}{4} (K - 3)^2 \right)$$

Donde, n son los grados de libertad, S muestra de la asimetría y K muestra curtosis

$$S = \frac{\hat{\mu}_3}{\hat{\sigma}^3} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^3}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^{3/2}} \quad K = \frac{\hat{\mu}_4}{\hat{\sigma}^4} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^4}{\left(\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 \right)^2}$$

3.7.3 Prueba de Box-Pierce

Esta prueba tiene como finalidad probar en forma conjunta que todos los coeficientes de autocorrelación son simultáneamente iguales a cero, es decir, que se presenta independiencia en la serie hasta el resago H (Villavicencio, 2010).

El contraste de las hipótesis de la prueba lo presentan los autores Franco & González (2012):

$H_0: \rho_1 = \rho_2 = \dots \rho_H = 0$ La serie no está Correlacionada

$H_1: \text{Al menos un } \rho_i \neq 0$ La serie está Correlacionada

El estadístico de prueba (Box & Ljung, 1978) es el siguiente:

$$Q_{BP} = n \sum_{k=1}^h \hat{\rho}_k^2$$

Donde, n es el tamaño de la muestra, $\hat{\rho}_k^2$ es la autocorrelación en los rezagos de la muestra k, y h es el número de rezagos que se están probando.

Para nivel de significancia α , la región crítica para el rechazo de la hipótesis de aleatoriedad está dada por: $Q > X_{1-\alpha, h}^2$

Donde $X_{1-\alpha, h}^2$ es el cuantil de la distribución chi-cuadrado con grados de libertad h .

3.8 Técnica de Monte Carlo por Cadenas de Markov (MCMC)

Para poder hablar de esa técnica, tendremos que explicar la forma de operar en la inferencia bayesiana, su razonamiento, por medio del modelo probabilístico.

Según Rodríguez (2003) el modelo probabilístico de la inferencia bayesiana es $f(y|\theta)$, $\theta \in \Omega$ es un espacio n dimensional de parámetros (n parámetros en el vector θ), lleva a observaciones muestrales de la forma $y = (y_1, y_2, \dots, y_n)$ y teniendo una densidad de probabilidad a priori $\rho(\theta)$, se consigue la densidad de probabilidad a posteriori de la siguiente manera:

$$\rho(\theta|y) = \frac{f(y|\theta)\rho(\theta)}{\int_{\theta \in \Omega} f(y|\theta)\rho(\theta)} \propto f(y|\theta)\rho(\theta) = \ell(\theta|y)\rho(\theta)$$

Por medio de esta densidad a posteriori podemos realizar inferencias como:

Estimación de parámetros: $\min_{\hat{\theta}} E[\ell(\theta, \hat{\theta})] = \min_{\hat{\theta}} [\int_{\theta \in \Omega} \ell(\theta, \hat{\theta})\rho(\theta|y)d\theta]$

Predicción: $p(x) = \int p(x|\theta) \rho(\theta)d\theta$

Evaluación de hipótesis: $1 - \alpha \leq P(C|x) = \int_C dF^{\pi(\theta|x)} d\theta$

Por otro lado, las distribuciones a priori para el parámetro no siempre se conocen, algunas veces deben elegirse sin información alguna. Por ello se crearon las distribuciones no informativas, la más común, la distribución uniforme, que da muy poca o casi ninguna explicación acerca del comportamiento del parámetro (Gelman et al, 2004). Para este trabajo, se

considerará en la técnica de elicitación bayesiana, una a priori no informativa y para el caso de la regresión por MCMC, la normal para los parámetros.

Rodríguez (2010) muestra la regresión bayesiana de la siguiente manera:

$$y = \beta + \beta_2 X_2 + \beta_3 X_3 + \epsilon_i, \epsilon_i \sim iid N(0, \sigma^2)$$

$$p(B, \sigma) \propto \sigma^{-m} \exp\left\{\frac{-1}{2\sigma^2} [n + (\beta - \mu)' \Psi^{-1} (\beta - \mu)]\right\}$$

$$p(\sigma) \sim Z^{-1/2} \text{ Donde, } Z \sim \Gamma(m, n) \quad p(B|\sigma) \sim N_{mult}(\mu, \sigma^2 \Psi)$$

Martin et al (2011) en su artículo nos introducen al uso del paquete de simulación, para el lenguaje R, y por medio de este da una explicación precisa de la función y la esencia del modelo. Los autores manifiestan que al utilizar MCMCpack, el vector β asume una distribución normal. Además señalan que al realizar esta simulación se garantiza estacionariedad en la serie, lo cual lleva a un cumplimiento de las pruebas de incorrelación para cada factor de la regresión. Respecto a la prueba de normalidad, no es necesaria su realización para este modelo, porque desde su estructura, se puede garantizar el cumplimiento de este supuesto, ya que al tener valores a priori normales y multiplicarlo por la verosimilitud normal, dará como resultado un factor a posteriori normal.

En la tabla 3 se muestra un resumen del marco teórico de los métodos y modelos a utilizar en este trabajo.

Tabla3. Resumen métodos y modelos estadísticos utilizados en el trabajo.

Técnica	Características Principales	Premisas
Suavización Exponencial	Estima valores del lapso del ajuste y pronostica valores futuros en función de la ponderación exponencial de periodos anteriores.	El modelo debe permitir la incorrelación final de residuales.
	Tiene la ventaja de adaptarse y ser flexible a observaciones que se encuentran en la serie.	
SARIMA	Ajusta la variabilidad con términos rezagados de la serie, ajusta la estacionalidad de esta.	El modelo no debe incluir coeficientes innecesarios, esto permite obtener mejores predicciones.
		Asegurar que la serie es estacionaria, es decir que su media es constante.
Modelo Econométrico Dinámico	La variable endógena (respuesta) depende de una o más variables exógenas (explicativas), retardos de cualquiera de estas y variables estacionales.	Debe cumplir supuesto de normalidad e incorrelación en los residuales.
	Se caracteriza por analizar la trayectoria de las variables en distintos instantes de tiempo.	
Método Delphi	Es una técnica de investigación social, que permite la obtención de una opinión grupal	El juicio subjetivo grupal presenta mayor calidad que el juicio subjetivo individual.
	Se mantiene el anonimato de los expertos durante todo el proceso, y se realizan retroalimentaciones de las opiniones encontradas en cada ronda.	Contar con un sólo juicio subjetivo de un experto, puede sesgar la información.
MCMC	Es una simulación Monte Carlo vía Cadenas de Markov	La prueba de incorrelación se le realiza a los parámetros simulados.
	Realiza una simulación para los coeficientes de cada variable, asumiendo las distribuciones a priori de los parámetros y de verosimilitud de los datos.	

4. METODOLOGÍA

Se utilizaron datos históricos de la demanda energética diaria en Colombia, desde el 15 de diciembre de 2010 hasta 31 de diciembre de 2012, facilitados por el departamento encargado de pronósticos de demanda de energía en XM filial de ISA, información de carácter público, lo que permite la posibilidad de mostrar los resultados expuestos a continuación. En la figura 1 se puede apreciar el comportamiento de la demanda diaria energética del país, por facilidad y para una mayor comprensión de la figura, sólo se tomaron los datos de tres semanas de junio de 2012, que son muy similares a las del periodo completo de análisis.

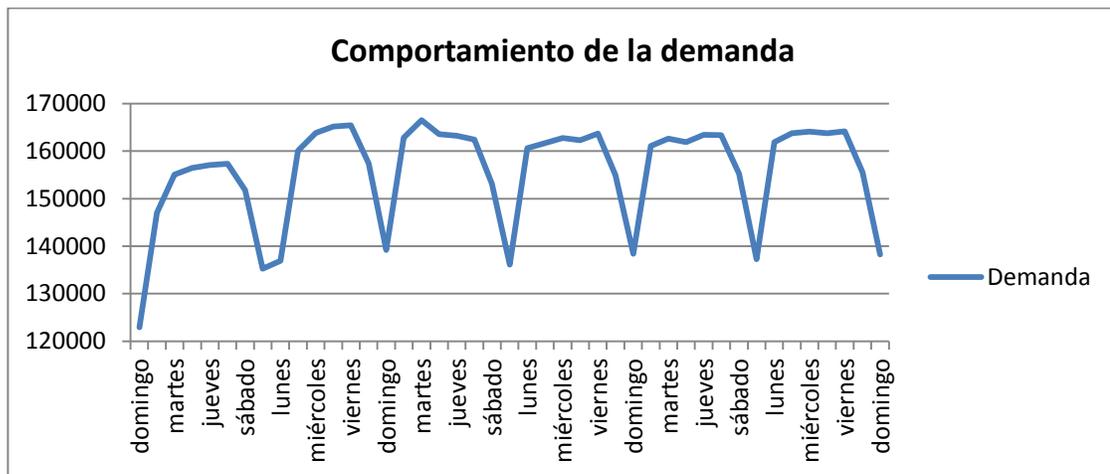


Figura 1. Comportamiento de la demanda energética en Colombia

Se realizó un análisis explorativo a priori a la serie, identificando picos energéticos significativos, así como días en donde la demanda era atípica, información que además sirvió de insumo para la construcción de otros modelos presentados en el trabajo.

Dentro de la serie de tiempo se incorporó la variable Temperatura, tenida en cuenta por el criterio de algunos expertos de la compañía XM por su posible impacto, a pesar de nunca haber sido utilizada para este tipo de predicciones.

Posteriormente, se estimaron modelos por medio de programas de lenguaje estadístico R y Statgraphics¹, con el respectivo análisis de forma, que caracterizara el comportamiento energético diario con todos los efectos importantes sobre éste. Se validará los supuestos de algunos modelos, se evaluará su eficiencia, conforme con el acierto por la disminución del error y se extraerá toda la información a priori, necesaria para los modelos bayesianos presentados también en este trabajo.

Para la estimación del modelo dinámico con variables endógenas y exógenas, fue necesario incluir retardos en la variable dependiente, así como también **variables ficticias**, para determinar posibles efectos cualitativos o niveles determinados de energía. Hanke et al. (1996) dice: *“En ocasiones es necesario determinar cómo se relaciona una variable dependiente con una independiente cuando un factor cualitativo ejerce influencia en la situación. Esta relación se logra creando una variable ficticia. Existen muchas formas de identificar cuantitativamente las clases de una variable cualitativa.”* En síntesis, las variables ficticias son utilizadas en el campo de la estadística para mejorar la explicación de la relación entre la variable dependiente y variables independientes. En resumen las variables más importantes consideradas para la modelación dinámica y bayesiana fueron:

- Retardos en la variable endógena
- Variables ficticias de nivel: día de la semana, nivel de consumo de energía (bajo, medio, alto), tipo de día (festividad, común).

¹ El software Statgraphics tiene aplicaciones interactivas y simples para el usuario, sin embargo el software R presenta más versatilidad para programar y para estimar algunas pruebas que Statgraphics no tiene, y que son necesarias para el desarrollo de este trabajo.

- Temperatura
- Tendencia temporal

4.1 Criterio de Evaluación de los modelos

En las estimaciones, validaciones y pronósticos realizados con los múltiples modelos clásicos utilizados, los datos usados comprenden la energía diaria de Colombia entre Diciembre de 2010 y diciembre de 2012, dejando los días de las últimas dos semanas de este mes, para pronosticar y evaluar el error de pronóstico del modelo.

Además de la validación de supuestos de los modelos estimados, el criterio final de evaluación del nivel óptimo que permitió comparaciones y elección del más indicado para el pronóstico de demanda diaria de energía, fue el MAPE (Mean Absolute Percentage Error), el cual mide la exactitud que tuvo el pronóstico con relación a la serie de tiempo (porcentaje de error) y se expresa en forma de porcentaje de error y se calcula de la siguiente manera:

$$MAPE = \sum_{t=1}^n \left| \frac{P_t - V_t}{V_t} \right|$$

Donde,

P_t=Valor pronosticado en el periodo t

V_t=Valor real en el periodo t

4.2 Metodología Delphi

El objetivo principal del desarrollo de la técnica Delphi para este trabajo es recopilar información valiosa por parte de un grupo de expertos, que permitirá pronosticar la demanda diaria del consumo energético en Colombia. A continuación se describirá paso a paso el desarrollo de la técnica a este caso de estudio.

4.2.1 Planteamiento del objetivo del pronóstico

Se pretende estimar la demanda diaria del consumo energético en Colombia, por medio de una empresa encargada de transportar la energía en todo el país, haciendo énfasis en días en que la demanda no tiene un comportamiento típico.

4.2.2 Composición del grupo coordinador

El grupo coordinador debe estar conformado por personas conocedores de la aplicación de la metodología Delphi, estas personas son las encargadas de interpretar las respuestas de los expertos y realizar ajustes. Los miembros de este grupo deben poseer características como liderazgo, escucha, comprensión y paciencia, con el fin de realizar una buena gestión

El grupo coordinador está integrado por la Magister en Estadística Marisol Valencia y los estudiantes de ingeniería industrial Jose Tabares y Carlos Velásquez que han hecho énfasis en esta técnica.

El grupo coordinador tiene estas funciones:

- **Establecer cronograma de trabajo:**

En esta etapa se definieron cada uno de los tiempos que lleva cada proceso, dentro de las actividades programadas estuvieron:

- Escoger un lugar adecuado donde se llevan a cabo las reuniones con los expertos. Se tiene que tener en cuenta las condiciones del espacio, comodidad y cualquier tipo de recursos (Computadores, impresoras) que se puedan requerir para llevar a cabo la metodología sin inconvenientes

- **Establecer criterio de selección de los participantes:**

Se eligieron los expertos en distintas áreas dentro de la compañía, con el fin de obtener resultados de distintos puntos de vista, esto permite que la información no se sesgue ni tampoco arroje datos atípicos que hacen del pronóstico poco confiable. También se tiene en cuenta la experiencia en años que poseen los participantes dentro de la compañía, permitiendo que la información sea precisa.

- **Grupo de expertos**

La selección del grupo de expertos es sin duda la parte más importante para el buen resultado del proceso. Para este caso de estudio los expertos son claramente cómo se menciona anteriormente, personas que llevan ya un recorrido dentro de la compañía y que tienen criterio y argumentos para responder el cuestionario.

La invitación a los participantes se les hace llegar oficialmente por medio de un correo electrónico donde se especifica que es lo que se pretende lograr:

- El objetivo de estudio
- Una descripción de cómo se llevara a cabo la metodología
- Explicación del porque esta persona ha sido considerada como experta
- Numero de cuestionarios y de rondas a llevar a cabo
- Duración del proceso

- **Obtención y transferencia de la información**

Este ítem tiene como objetivo el diseño del cuestionario, y el proceso de retroalimentación de la información.

Para el diseño del cuestionario, después de analizar el comportamiento de la demanda energética, se tuvo en cuenta que el comportamiento de los días en la semana es diferente a los días del fin de semana, por lo tanto se aclaró en el cuestionario qué preguntas correspondían a cada aspecto identificado. Por otra parte, también se tuvo en cuenta que para el transporte de energía influyen variables exógenas, para este caso: Tipo de día, época del año, características del mes, y temperatura del clima. Para el pronóstico de demanda atípica en diciembre fue pertinente realizar una pregunta más a los expertos, con el fin de establecer la diferencia porcentual que tienen estos días con respecto a la demanda normal diaria.

En los resultados se mostrará el cuestionario para semana de lunes a viernes y para el fin de semana (sábado y domingo)

- **Integración de las opiniones individuales**

El grupo coordinador lleva a cabo la ejecución de dos rondas sobre el cuestionario elaborado.

En la primera ronda el grupo analiza los datos, y realiza una retroalimentación a los expertos sobre las respuestas de los demás, siempre manteniendo el anonimato. Para la próxima ronda los expertos darán un juicio más realista basándose en las opiniones logrando resultados más aplomados.

4.2.3. Algoritmo para el análisis de la distribución aposteriori

A continuación se muestran los pasos para realizar un análisis de los resultados que permite estimar el pronóstico, con estadística bayesiana:

- Realizar una descripción general de las estadísticas básicas del consumo energético, para ambos periodos: lunes-viernes, fin de semana.
- Estimar intervalos de alta densidad de ambas distribuciones.
- Usar el peso del cambio del día para crear el efecto de aumento o disminución en el consumo energético, como vector de información a priori.
- Usar una apriori no informativa para la media como la distribución uniforme, y como verosimilitud, la elicitada, lo cual hace constante la uniforme para la aposteriori. Finalmente, se estima la media de la distribución uniforme aposteriori, que se basa en un intervalo de alta densidad del 50% encontrado, para generar la predicción media.
- Incorporando la información a priori del peso y la media m estimada de la distribución uniforme, la ecuación usada de pronóstico es $Y_t = m + m * peso_t$

5. RESULTADOS

En este capítulo se presentarán las estimaciones de los modelos estadísticos usados, con la respectiva caracterización del consumo energético diario de Colombia, usando las pruebas de validación y criterios de comparación necesarios para un adecuado análisis y elección del que permite optimizar de manera más integral, el pronóstico al mostrar menor nivel de error absoluto relativo.

5.1 Suavización exponencial simple

Los modelos de suavización simple estimados no reflejaban muy bien la volatilidad de la serie, como se observa en la figura 2, donde se muestra un lapso del valor real (azul) versus el pronosticado (rojo), evidenciando un mal ajuste.

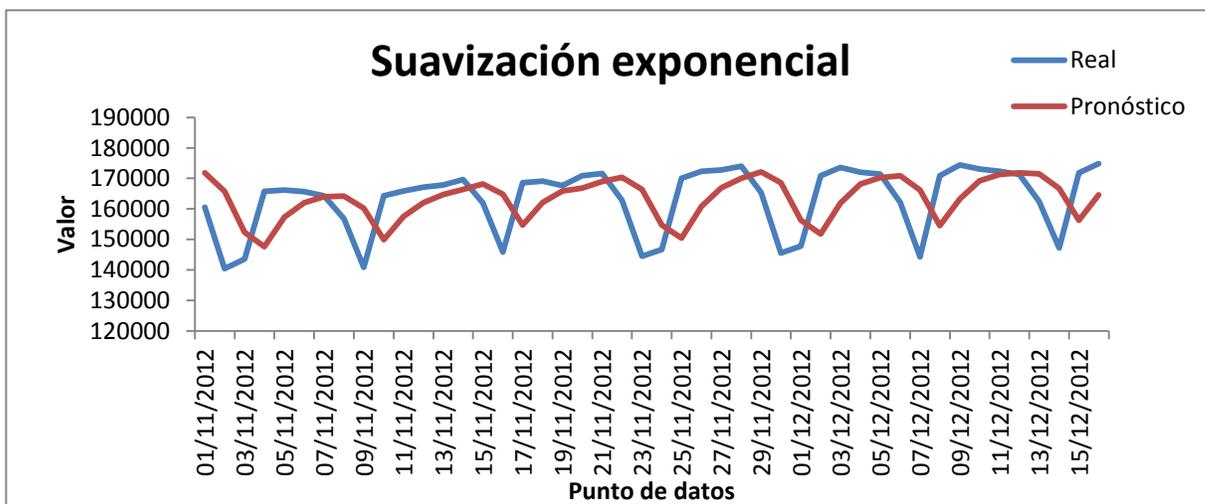


Figura 2. Valor real vs valor pronosticado por Suavización Exponencial Simple

La estimación se basa en una optimización del parámetro que realiza el software Statgraphics. Dicho software permite evaluar la autocorrelación en los residuales, y estimar valores promedios del error absoluto relativo (MAPE), y predecir valores futuros.

El mejor modelo estimado con la técnica de Suavización exponencial simple, se obtuvo a partir de un coeficiente de suavización $\alpha = 0.4678$

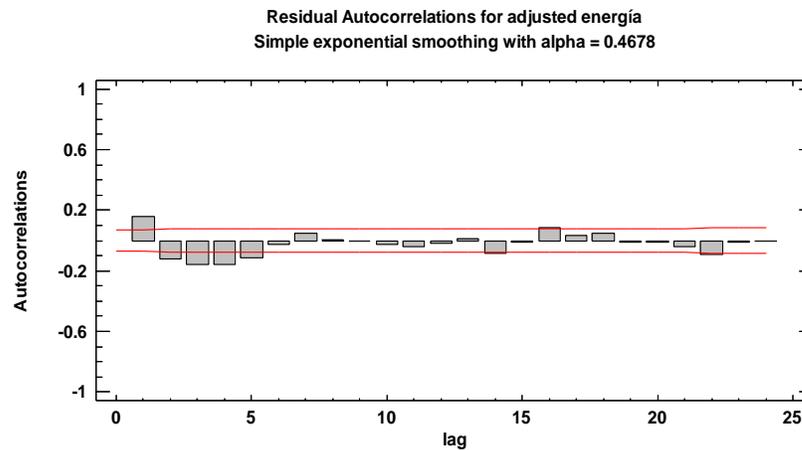


Figura 3. Autocorrelación de los residuales del modelo de suavización exponencial simple

En la figura 3. se evidencia una alta correlación en los residuales hasta el rezago 5, se puede intuir que no pasará la prueba de Box Pierce, mostrando el no cumplimiento de la premisa de incorrelación.

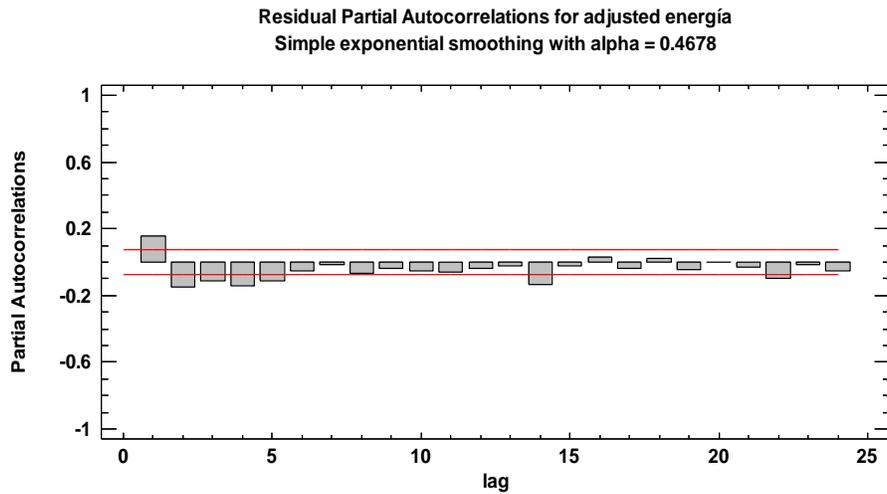


Figura 4. Autocorrelación parcial de los residuales del modelo de suavización exponencial simple

Validación de los residuales del modelo estimado:

- Prueba de Incorrelación de Box-Pierce:

Basada en los primero 24 rezagos se obtuvo un estadístico de Large sample test statistic = 98,3511 y un Valor P = 2,70787E-11, por lo tanto se corrobora que hay correlación en los residuales, el modelo no es confiable.

Tabla 4. Pronostico suavización exponencial simple

Fecha	Demanda Real (Mw)	Pronóstico	Error relativo
17/12/2012	169756,92	167770	1,17%
18/12/2012	170427,43	174092	2,15%
19/12/2012	173748,52	175176	0,82%
20/12/2012	173399,14	174696	0,75%
21/12/2012	170600,50	174003	1,99%
22/12/2012	160288,86	164050	2,35%
23/12/2012	145036,17	146561	1,05%

Fecha	Demanda Real (Mw)	Pronóstico	Error relativo
24/12/2012	147566,05	167770	13,69%
25/12/2012	131777,18	174092	32,11%
26/12/2012	160614,76	175176	9,07%
27/12/2012	162983,71	174696	7,19%
28/12/2012	163891,65	174003	6,17%
29/12/2012	155519,14	164050	5,49%
30/12/2012	143156,82	146561	2,38%
31/12/2012	143094,86	167770	17,24%

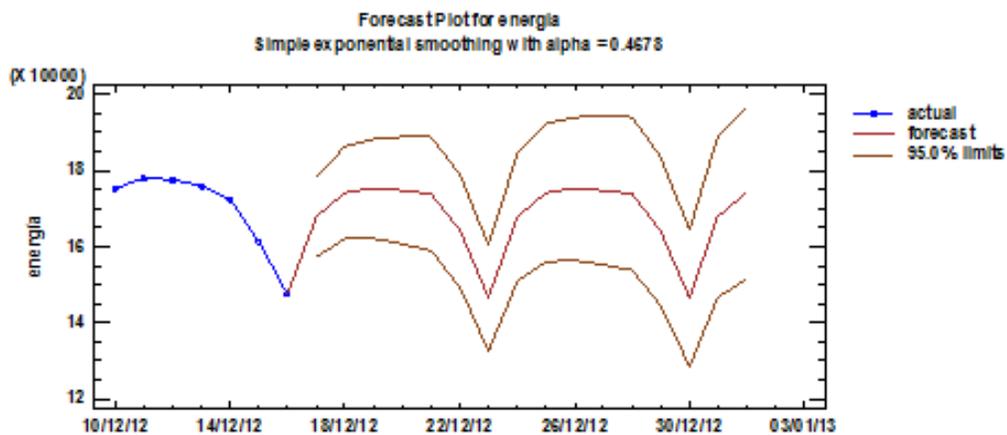


Figura 5. Gráfica de pronóstico del modelo de suavización exponencial simple

Análisis de resultados

En la tabla 4, se observa que el valor medio del error absoluto de los datos ajustados desde el 10 de diciembre del 2010 hasta el 15 de diciembre de 2012 fue de 2,17% (MAPE de ajuste) y el valor medio del error absoluto del pronóstico de la demanda energética comprendida entre los días 16 y 31 de diciembre del 2012 fue 9,70% (MAPE de pronóstico).

Aunque el modelo tiene un MAPE de ajuste bajo, no cumple la prueba de incorrelación de Box Pierce, por lo tanto el modelo no es muy confiable, lo cual además se corrobora con el alto valor del MAPE de pronóstico (9.7%), indicando desacierto en las predicciones.

Como se puede observar, entre los MAPES hay una diferencia de 7 puntos aproximadamente, esto se debe a que el modelo no está en la capacidad de captar los días atípicos, como 24, 25, y 31 de diciembre, donde se presenta mayor error.

5.2 Suavización exponencial lineal de holt

En este modelo, se utilizaron el mismo lapso de datos de demanda energética del modelo anterior. Con la ayuda del software Statgraphics, se realizaron las estimaciones y se encontraron dos coeficientes de suavización, uno para la media y otra para la tendencia: $\alpha = 0,3239$ y $\beta = 0,0036$ respectivamente. Se encontró una desviación muy similar de los ajuste versus los valores pronosticados.

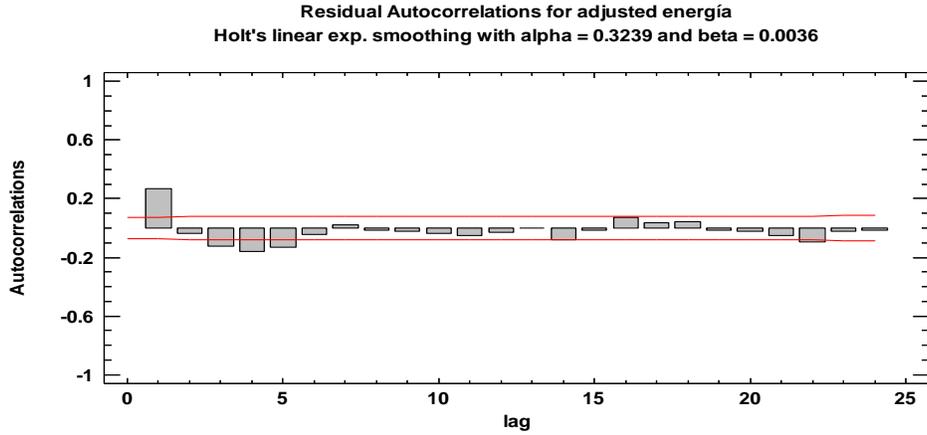


Figura 6. Autocorrelacion de los residuales del modelo de suavización exponencial lineal

En la figura 6 se evidencia una posible correlación de los datos, puesto que varias barras están por fuera de las bandas. Hay una alta correlación en los rezagos 1 y 5, por lo que se intuye un incumplimiento de la prueba de Box Pierce.

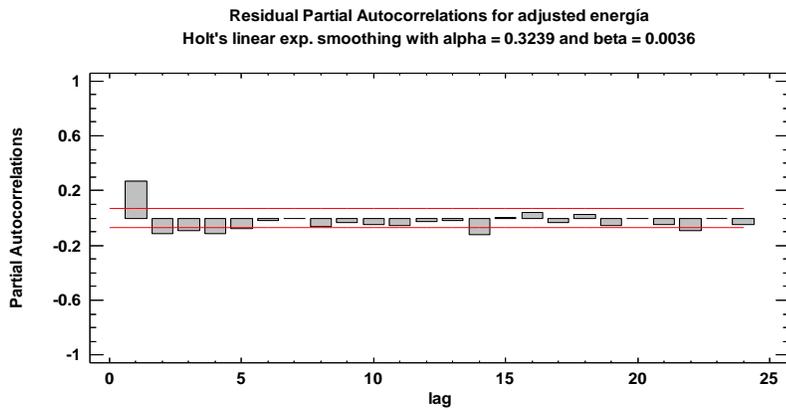


Figura 7. Autocorrelacion parcial de los residuales del modelo de suavización exponencial lineal

Validación de los residuales del modelo estimado:

- Prueba de Incorrelación de Box-Pierce:

Basada en los primeros 24 rezagos se obtuvo un estadístico de Large sample test statistic = 98,3511 y un Valor P = 2,70787E-11, por lo tanto se corrobora que hay correlación en los residuales, el modelo no es confiable.

Dado que el valor p de la prueba de Box Pierce es muy pequeño, se observa que los residuales tienen autocorrelación, este modelo no es confiable, por generar sesgos en las predicciones.

Tabla 5. Pronostico suavización exponencial Holt

Fecha	Demanda Real (Mw)	Pronóstico	Error relativo
17/12/2012	169756,92	168095,0	0,98%
18/12/2012	170427,43	174476,0	2,38%
19/12/2012	173748,52	175609,0	1,07%
20/12/2012	173399,14	175175,0	1,02%
21/12/2012	170600,50	174527,0	2,30%
22/12/2012	160288,86	164588,0	2,68%
23/12/2012	145036,17	147080,0	1,41%
24/12/2012	147566,05	168410,0	14,13%
25/12/2012	131777,18	174803,0	32,65%
26/12/2012	160614,76	175938,0	9,54%
27/12/2012	162983,71	175503,0	7,68%
28/12/2012	163891,65	174854,0	6,69%
29/12/2012	155519,14	164897,0	6,03%
30/12/2012	143156,82	147356,0	2,93%

Fecha	Demanda Real (Mw)	Pronóstico	Error relativo
31/12/2012	143094,86	168726,0	17,91%

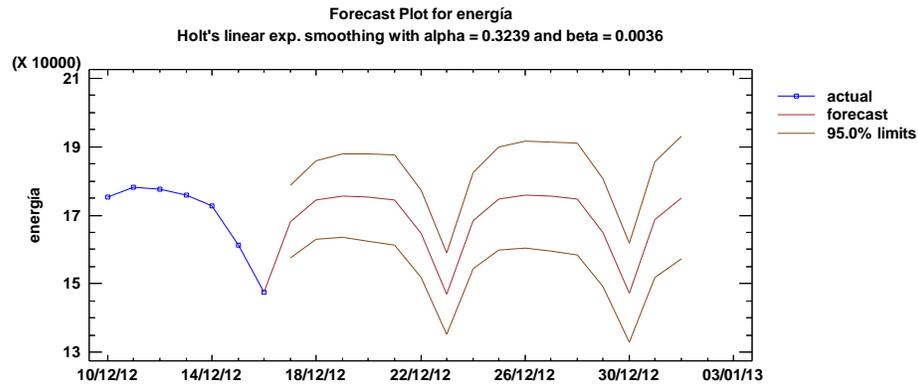


Figura 8. Gráfica del pronóstico del modelo de suavización exponencial lienal.

En la figura 8 se puede observar el comportamiento de los datos y como el modelo captó la estacionalidad para el pronóstico.

En la tabla 5 se muestra el valor medio del error absoluto de los datos ajustados desde el 10 de diciembre del 2010 hasta el 15 de diciembre de 2012 fue de 2,155% (MAPE de ajuste), y el valor medio del error absoluto del pronóstico de la demanda energética comprendida entre los días 16 y 31 de diciembre del 2012 fue 7,19% (MAPE de pronóstico).

La diferencia entre los MAPES es de 5 puntos aproximadamente, esto se debe a que el modelo no está en la capacidad de captar los días atípicos, como 24, 25, y 31 de diciembre, donde se presenta mayor error.

El modelo no cumple la prueba de incorrelación de Box Pierce, además, se obtuvo un error de pronóstico poco aceptable, por lo tanto, no se considera confiable para la predicción de semanas atípicas.

5.3 Modelo SARIMA

Al analizar los valores de autocorrelación de la serie de consumo energético, se detectó una estacionalidad de orden 7 muy evidente, ya que los valores de autocorrelación en los periodos 7 y 14 fueron muy altos (0,7639 y 0,7070 respectivamente). Así mismo, la autocorrelación de orden 1 es también alta (0,4527). Esto indica que los retardos más evidentes que tienen efectos sobre la serie son Z_{t-1} , Z_{t-7} y Z_{t-14} . Además de estas evidencias, se usó el método automático de R para detectar de mejor forma los parámetros indicados para el modelo. Luego de estimar varios modelos, se obtuvo que el ARIMA (1,1,1)X(2,0,1) fue el de menor error

Tabla 6. Coeficientes SARIMA

Coeficientes				
ar1	ma1	sar1	sar2	sma1
0,4950	-0,9321	0,9842	0,0075	-0,8028
0,0416	0,0196	0,0407	0,0402	0,0275

Se estimó un modelo ARIMA con coeficiente estacional, por medio del método automático de R, también se estimaron modelos ARIMA pero no resulta conveniente por la estacional de la serie, este modelo no la capta y arroja un valor de pronóstico para todos los días, por lo tanto no es indicado.

Validación de los residuales del modelo estimado:

- Prueba de normalidad de Jarque Bera

$\chi^2=1956,265$, $df = 2$, $p\text{-value} < 2,2e-16$

- Prueba de incorrelación de Box-Pierce

$\chi^2=2,079$, $df = 5$, $p\text{-value} = 0,8381$

No cumple normalidad pero sí incorrelación

Tabla 7. Pronostico SARIMA

Fecha	Demanda Real (Mw)	Pronóstico	Error relativo
17/12/2012	169756,92	157213	7,39%
18/12/2012	170427,43	162075	4,90%
19/12/2012	173748,52	162372	6,55%
20/12/2012	173399,14	162127	6,50%
21/12/2012	170600,5	160839	5,72%
22/12/2012	160288,86	149750	6,57%
23/12/2012	145036,17	133938	7,65%
24/12/2012	147566,05	153945	4,32%
25/12/2012	131777,18	160745	21,98%
26/12/2012	160614,76	161646	0,64%
27/12/2012	162983,71	161712	0,78%
28/12/2012	163891,65	160668	1,97%
29/12/2012	155519,14	149600	3,81%
30/12/2012	143156,82	133598	6,68%
31/12/2012	143094,86	153791	7,47%

MAPE ajuste=2,05%

MAPE pronóstico=6,20%

Además de no cumplir todas las pruebas de validación, el error relativo absoluto aumenta mucho a futuro, lo cual deja este modelo en un estado poco confiable en relación a su capacidad de pronóstico, en especial a largo plazo.

5.4 Modelo Dinámico

En este modelo, se realizaron alrededor de ocho estimaciones, pero por practicidad, sólo encontrarán el modelo final y uno previo, para mostrar brevemente el proceso que se realizó.

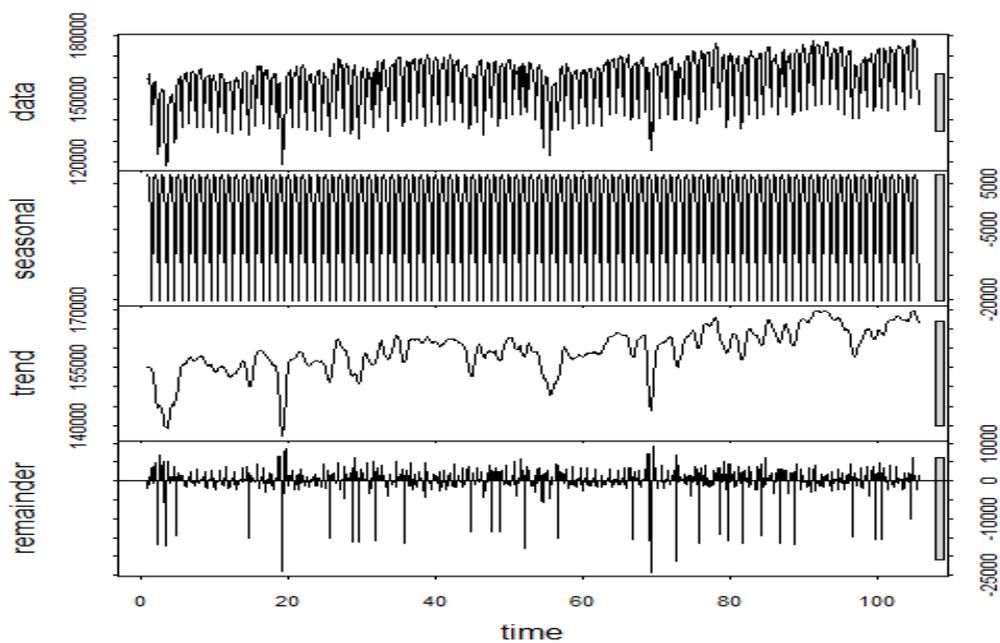


Figura 9. Gráfica de la tendencia de la serie

En la figura 9 no se observa una clara tendencia en los datos, pero se puede incorporar varias variables de tiempo dentro del modelo dinámico, con el fin de captarla y contribuir a la reducción del error.

En la siguiente estimación se involucraron variables cómo:

- Tiempo
- Día
- Tendencia senoidal
- Tres rezagos (variables del pasado)

Pero la variable senoidal y la tercera del pasado, tenían un valor p (vp) mayor a 0.05 por lo tanto no eran significativas, por lo que se procedió a excluirlas del modelo, como se aprecia en la Tabla 8.

Tabla 8. Coeficientes Estimados del Modelo Dinámico

Coeficientes	Estimado	Error estándar	t_0^*	Valor P
Intercepto	4,05E+04	5,75E+03	7,039	4,53e-12
T	5,03E+01	8,25E+00	6,092	1,82e-09
Día jueves	1,88E+04	1,10E+03	17,083	< 2e-16
Día lunes	2,75E+04	1,13E+03	24,467	< 2e-16
Día martes	2,21E+04	1,07E+03	20,621	< 2e-16
Día miércoles	1,98E+04	1,11E+03	17,870	< 2e-16
Día sábado	1,01E+04	8,98E+02	11,269	< 2e-16
Día viernes	1,85E+04	1,09E+03	16,987	< 2e-16
L(energíats, 1)	5,43E-01	3,13E-02	17,336	< 2e-16
L(energíats, 7)	8,17E-02	3,14E-02	2,606	0,00936

* t_0 : es el valor del estadístico de prueba de cada parámetro.

Según la tabla 6, el Valor P inferior a 0,05 indica que todos los factores son relevantes para el modelo.

La ecuación para este modelo es:

$$y = 4,05E + 04 + 50,3E + 01 * T + 0,543 * Z_{t-1} + 0,0817 * Z_{t-2} + 27500 * \text{lunes} + 22100 * \text{martes} + 19800 * \text{miércoles} + 18800 * \text{jueves} + 18500 * \text{viernes} + 10100 * \text{sábado} + \xi$$

Validación de los residuales del modelo estimado:

1. Prueba de normalidad de Jarque Bera

$$\chi^2 = 4593,097, \text{ df} = 2, \text{ p-value} < 2,2e-16$$

2. Prueba de Incorrelación de Box-Pierce

$$\chi^2 = 372,5744, \text{ df} = 5, \text{ p-value} < 2,2e-16$$

Los residuales no cumple las pruebas de normalidad e incorrelación, ya que el VP es menor que 0,05, por lo tanto no se considera un pronóstico confiable a largo plazo. El MAPE de ajuste del modelo es de 2,485%, aunque da pequeño, el modelo generaría sesgo de pronóstico a el largo plazo, además pueden usarse unos métodos alternativos para mejorar el cumplimiento de estos supuestos.

Para el siguiente modelo fue pertinente investigar y realizar un análisis de la temperatura promedio en Colombia, puesto que esta fue una variable que en XM nunca se había tenido en cuenta y según los expertos, es una variable que influye mucho en el comportamiento de la demanda de energía.²

² Valores consultados en la base de datos de: www.tutiempo.net

Luego de una exploración de la base y de conversaciones preliminares con los expertos de XM, se notaron unas variables de nivel que pueden justificar posibles cambios estructurales en la serie y que si se predicen, pueden disminuir los sesgos o errores en los pronósticos. Estas son:

1. En la variable "Nivel", se catalogaron la demanda en 4 intervalos (1, 2, 3, 4) de menor a mayor respectivamente. Estos intervalos tienen un rango de más o menos de 14000 Mwh cada uno.
2. En la variable Nivel2, se catalogaron las fechas especiales, que son: 24, 25 y 31 de diciembre. Esto se hizo con el fin de que el modelo identificara los picos de la demanda energética Colombiana.
3. En Nivel3 se identificaron los picos más bajos de los datos, generalmente lo representó el día domingo.

Estos niveles se involucran en el modelo, con el fin de poder captar picos y datos atípicos a la hora de pronosticar la demanda.

Para llegar al siguiente modelo (final) se realizaron varias transformaciones, cómo logarítmica, raíz cuadrada, raíz cúbica, y se realizaron las pruebas de hipótesis pertinentes. También intentó ingresar variables de seno y coseno en función del tiempo, intentando identificar la tendencia. Y se estimó un modelo para los residuales.

En este modelo se realizó una transformación de raíz cuadrada y se involucraron las siguientes variables:

- Tiempo
- Transformación cúbica del tiempo
- Día
- Temperatura

- Nivel
- Nivel2
- Nivel3
- Dos variables del pasado (2,7)

En el análisis previo, la variable Nivel2 no resultó significativa para el modelo, además aumentaba un poco el MAPE, por lo que se decidió sacarla y así mejorar el modelo. Como se ve en la tabla 9; en la cual se muestran la suma de cuadrados, los grados de libertad el valor F y su significancia.

Tabla 9. Analisis de Varianza (ANOVA)

	Suma de cuadrados	Grados de libertad	Valor F	Valor P
Intercepto	13923,3	1	1300,7231	< 2,2e-16
T	555,1	1	51,8609	1,52E-12
T3	159,4	1	14,8959	0,000124
Día	1041,3	6	16,2125	< 2,2e-16
Temperatura	118,5	1	11,0678	0,0009237
Factor (nivel)	7836,4	3	244,0283	< 2,2e-16
Nivel3	5051,2	2	235,9456	< 2,2e-16
L(sqrt(energy), 7)	41,3	1	3,8583	0,0498903
L(sqrt(energy), 2)	294	1	27,467	2,11E-07
Residuales	7589,3	709		

En la tabla 9 se ve que con una significancia del 5% todas las variables son significativas en el modelo. El supuesto de que la temperatura influye en la demanda es corroborado.

Tabla 10. Residuales

Residuales:				
Min	1Q	Median	3Q	Max
-13,9005	-1,9844	0,1109	2,2875	11,8034

Tabla 11. Coeficientes Estimados finales del modelo dinámico final.

Coeficientes:				
	Estimado	Error estándar	t ₀	Valor P
Intercepto	2,99E+02	8,30E+00	36,066	< 2e-16
T	8,08E-02	1,12E-02	7,201	1,52E-12
T3	-3,88E-06	1,01E-06	-3,86	0,000124
Día jueves	4,64E+00	8,84E-01	5,249	2,03E-07
Día lunes	3,65E+00	8,01E-01	4,555	6,16E-06
Día martes	7,15E+00	1,11E+00	6,465	1,88E-10
Día miércoles	5,06E+00	9,25E-01	5,473	6,15E-08
Día sábado	1,46E-01	7,54E-01	0,193	0,847106
Día viernes	4,34E+00	8,69E-01	4,994	7,46E-07
Temperatura	2,83E-01	8,51E-02	3,327	0,000924
factor(nivel)2	1,63E+01	1,09E+00	15,032	< 2e-16
factor(nivel)3	2,35E+01	1,29E+00	18,309	< 2e-16
factor(nivel)4	3,28E+01	1,39E+00	23,606	< 2e-16
Factor-m Nivel3	1,27E+01	9,21E-01	13,785	< 2e-16
Factor-s Nivel3	1,98E+01	9,88E-01	20,054	< 2e-16
L(sqrt(energy), 7)	3,10E-02	1,58E-02	1,964	0,04989
L(sqrt(energy), 2)	8,56E-02	1,63E-02	5,241	2,11E-07

En la tabla 11 se observa que todos los factores exceptuando el día sábado es significativo para el modelo, como en el análisis de varianza observamos

que la variable Día era significativa, entonces podemos dejar en el modelo el día sábado.

Sin embargo, los residuales aún no son incorrelacionados (no son ruido blanco) y se hace necesaria una corrección, como modelarlos con otra técnica. En este caso se utilizará un modelo ARIMA para ajustar los residuales y con base en esto unificarlo al modelo dinámico estimado.

Por ahora, el MAPE ajustado es 1,29%, aunque es pequeño, se verá su disminución posterior a la modelación ARIMA para los residuales.

Ahora, se hará una exploración de la posible dependencia de los residuales del modelo dinámico estimado.

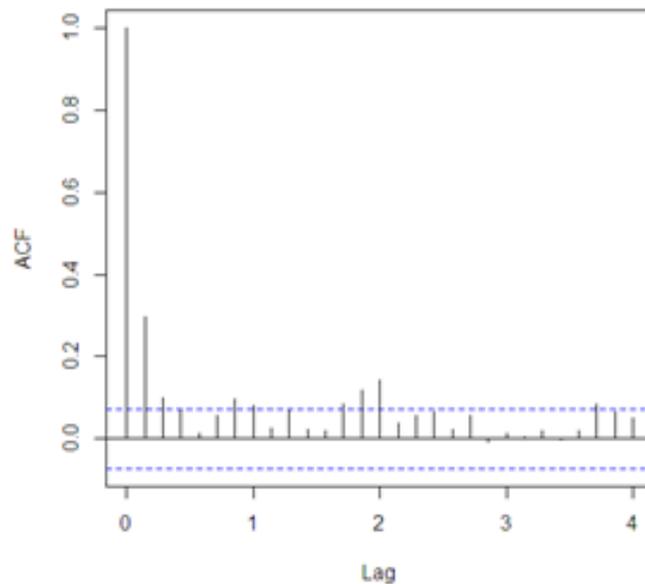


Figura 10. Autocorrelación de los residuales del modelo dinámico

En la figura 10 se observa que los rezagos se comportan de manera sinodal, y se nota la alta correlación que tiene el rezago 1, pero también en el orden estacional con $L=7$.

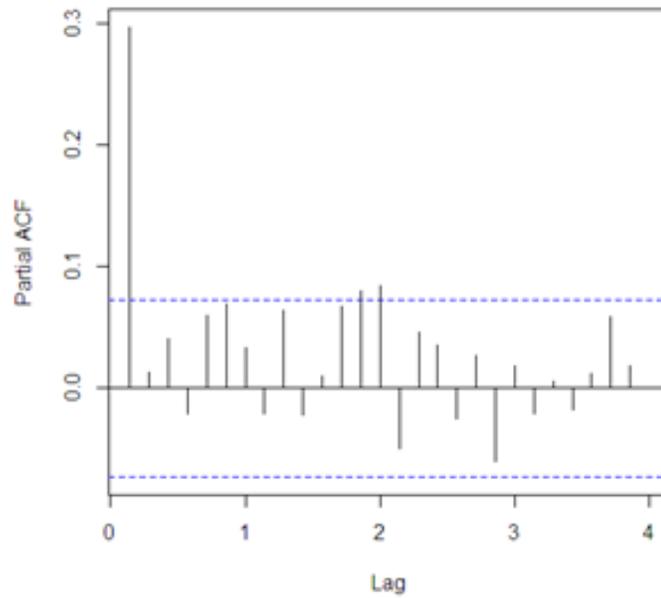


Figura 11. Autocorrelación parcial de los residuales de los errores del modelo dinámico

El comportamiento de dependencia de orden 1 y de orden 7 (estacional) fue tomado en cuenta para ajustar el mejor modelo para estos residuales, encontrando como el mejor, un modelo SARIMA con coeficiente estacional, por medio del método automático de R.

Modelo: SARIMA (1,0,0)x(2,0,1) para los residuales del modelo dinámico.

Tabla 12. Coeficientes del modelo SARIMA

Coeficientes:				
	ar1	sar1	sar2	sma1
	0,2876	-0,1040	0,1309	0,1555
S.E	0,0360	0,3958	0,0411	0,3998

Validación de los residuales del modelo estimado:

- Prueba de normalidad de Box Pierce

X-cuadrado = 7,5156, df = 7, p-value = 0,3772

Cumple el supuesto con una significancia del 5%, los residuales de los errores no están correlacionados.

Al unir ambos modelos, se estima el calcula el error del modelo Dinámico con errores SARIMA como se evidencia en la Tabla 13.

Tabla 13. Pronostico Dinámico con errores SARIMA

Fecha	Demanda Real (Mw)	Pronóstico	Erro relativo
17/12/2012	169756,92	169577,047	0,11%
18/12/2012	170427,43	172507,292	1,22%
19/12/2012	173748,52	173107,07	0,37%
20/12/2012	173399,14	173129,296	0,16%
21/12/2012	170600,5	172197,649	0,94%
22/12/2012	160288,86	162291,04	1,25%
23/12/2012	145036,17	148032,433	2,07%
24/12/2012	147566,05	150608,785	2,06%
25/12/2012	131777,18	131452,115	0,25%
26/12/2012	160614,76	162301,458	1,05%
27/12/2012	162983,71	161744,319	0,76%
28/12/2012	163891,65	163645,38	0,15%
29/12/2012	155519,14	153983,187	0,99%
30/12/2012	143156,82	147318,848	2,91%
31/12/2012	143094,86	143032,345	0,54%

Fecha	Demanda Real (Mw)	Pronóstico	Erro relativo
		MAPE	0,99%

MAPE ajustado: 1,23%

MAPE pronóstico: 0,99%

El ajuste baja de 1,29 a 1,23%, pero además, encuentra un error muy bajo futuro (0,99%), indicando con ello que este es el que minimiza más el error de pronóstico de los que van explorados hasta ahora.

La ecuación final para el modelo dinámico con errores ARIMA es:

$$y = 2,99E + 02 + 8,08E - 02 * T - 3,88E - 06 * T^3 + 8,56 * Zt - 2 + 0,031 * Zt - 7 + 3,65 * \text{lunes} + 7,15 * \text{martes} + 5,06 * \text{miércoles} + 4,64 * \text{jueves} + 4,34 * \text{viernes} + 0,146 * \text{sábado} + 0,283 * \text{Temperatura} + 16,3 * \text{Factor 2 (Nivel)} + 23,5 \text{Factor 3 (Nivel)} + 32,8 * \text{Factor 4 (Nivel)} + 12,7 * \text{Nivel m} + 19,8 \text{Nivel s} + \text{SARIMA}(1.1.1) \times (2.0.1)$$

Se encuentra incorrelación en los residuales y se logra un MAPE bastante pequeño. Sin embargo, se evidencia no normalidad, razón por la cual, el modelo es adecuado sólo a corto plazo y se presentarán otras alternativas de pronóstico.

Al caracterizar el comportamiento de la demanda energética diaria de Colombia, se encontraron diferencias entre las semanas comunes de trabajo, estudio, cotidianas, y semanas atípicas, como la comprendida entre 24 y 31 de diciembre, que requieren ser discriminadas por medio del análisis de datos. Precisamente, los días como el 24, 25, y 31 de diciembre no cumplen

con la misma estacionalidad semanal que se tiene la serie, por lo que su dependencia no estaba muy ligada al factor día (lunes, martes, miércoles, etc.), pero sí a una disminución de consumo energético, lo que se refleja en el las variables de nivel consideradas en el modelo dinámico: una debida al tipo de demanda, en relación con la particularidad del día (festividades de fin de año) y otra con el nivel de consumo (bajo, medio y alto).

Estos resultados se ven ilustrados en la tabla 13, en la cual se evidencia un muy buen modelo para este tipo de demanda, logrando errores relativos bastante bajos. En la tabla 14 se realiza una comparación parcial de errores en los modelos analizados (clásicos), se logra apreciar el buen desempeño del Dinámico en días críticos como el 24, 25 y 31 de diciembre, este resultado se debe a la capacidad y la ventaja que tiene dicha técnica en tener niveles que permiten capturar los cambios estructurales más pronunciados.

Tabla 14. Resumen pronosticos metodologías clásicas

Fecha	Demanda Real (Mw)	Error relativo			
		Dinámico	SARIMA	Suavización Holt	Suavización simple
17/12/2012	169756.92	0,11%	7,39%	1,17%	0,98%
18/12/2012	170427.43	1,22%	4,90%	2,15%	2,38%
19/12/2012	173748.52	0,37%	6,55%	0,82%	1,07%
20/12/2012	173399.14	0,16%	6,50%	0,75%	1,02%
21/12/2012	170600.5	0,94%	5,72%	1,99%	2,30%
22/12/2012	160288.86	1,25%	6,57%	2,35%	2,68%
23/12/2012	145036,17	2,07%	7,65%	1,05%	1,41%
24/12/2012	147566,05	2,06%	4,32%	13,69%	14,13%
25/12/2012	131777,18	0,25%	21,98%	32,11%	32,65%
26/12/2012	160614,76	1,05%	0,64%	9,07%	9,54%
27/12/2012	162983,71	0,76%	0,78%	7,19%	7,68%
28/12/2012	163891,65	0,15%	1,97%	6,17%	6,69%
29/12/2012	155519,14	0,99%	3,81%	5,49%	6,03%

Fecha	Demanda Real (Mw)	Error relativo			
		Dinámico	SARIMA	Suavización Holt	Suavización simple
30/12/2012	143156,82	2,91%	6,68%	2,38%	2,93%
31/12/2012	143094,86	0,54%	7,47%	17,24%	17,91%
	MAPE	0,99%	6,20%	6,91%	7,29%

La presentación de técnicas bayesianas inicia en la siguiente sección, donde se introduce un método de elicitación bayesiana para extraer de los expertos información acerca de la variación de la serie del comportamiento energético diario del país.

5.5 Método Delphi

A continuación, en las dos siguientes tablas (tabla 15 y 16) se muestra la segunda ronda de elicitación que se realizó a los expertos, en las que se evaluó el comportamiento común de la demanda en una semana de lunes a viernes y del fin de semana respectivamente. Los valores están en Mwh, exceptuando el ajuste del rango intercuartil, el cual está en GWh.

Tabla 15. Resultados 2da. Ronda método Delphi (semana)

De 1 a 5, cómo califica su conocimiento sobre la demanda de energía diaria eléctrica en Colombia.	¿Cuál es la demanda media de energía diaria en una semana de lunes a viernes? (Mw)	¿Cuál cree que es un valor de consumo de energía diaria muy bajo, sin que sea el mínimo de lunes a viernes? (Mw)	¿Cuál cree que es el valor de consumo de energía diaria mínimo de lunes a viernes? (Mw)	¿Cuál cree que es el valor de consumo de energía diaria por debajo del cual está el 25% del consumo energético de lunes a viernes? (Mw)	¿Cuál cree que es el valor de consumo de energía diaria máximo de lunes a viernes? (Mw)	¿Cuál cree que es el valor de consumo de energía diaria alto, sin que sea máximo?(Mw)	¿Cuál cree que es el valor de consumo de energía diaria que deja el 75% por debajo?	Ajuste rango intercuartil en GWh
4	174000	172000	170000	173000	179000	178000	176000	(173-176)
4	170000	162000	157000	160000	175000	174000	171000	(160-171)
5	170000	162000	158000	164000	176000	175000	174000	(164-174)
5	165000	160000	159000	161000	171000	169000	168000	(161-168)
4	172000	161000	159000	163000	171000	170000	167000	(163-167)
3	164000	157000	156000	161000	175000	172000	168000	(161-168)
1	164000	157000	155000	161000	168000	167000	166000	(161-166)
5	173000	162000	159000	167000	177000	176000	174000	(167-174)
4	170000	163000	161000	166000	175000	174500	171500	(166-171,5)
2	172000	164000	163000	166000	179000	177000	174000	(166-174)
Promedio	169400	162000	159700	164200	174600	173250	170950	(164,2-170,95)
Ponderación puntual	0,25	0,05	0,01	0,1	0,01	0,05	0,2	

Tabla 16. Resultados 2da. Ronda método Delphi (semana)

De 1 a 5, cómo califica su conocimiento sobre la demanda de energía eléctrica diaria en Colombia.	¿Cuál cree que es el valor de consumo de energía diaria muy alto del fin de semana? (Mw)	¿Cuál cree que es el valor de consumo de energía diaria muy bajo del fin de semana? (Mw)	¿Cuál cree que es el valor de la mediana de consumo de energía del fin de semana? (Mw)
4	158000	152000	154000
4	167000	135000	160000
5	167000	135000	150000
5	159000	133000	156000
4	160000	133000	154000
3	163000	131000	150000
1	160000	130000	155000
5	164000	134000	153000
4	163000	132500	154000
2	170000	139000	153000
Promedio	163100	135450	153900
Ponderación puntual	0,05	0,05	0,25

Cada pregunta, tenía la intención de conocer un percentil en especial de la demanda energética en Colombia según el experto, Por ello, se enfocó de tal manera que el asignara la importancia de cada dato acorde con el peso respectivo.

En la segunda ronda hay mayor homogeneidad en las respuestas, en relación a la primera, ya que los expertos son retroalimentados de forma que conozcan el criterio de sus compañeros, bajo el anonimato de todos.

A continuación, se mostrarán los histogramas de la estimación de demanda energética diaria acorde con la información proporcionada por los expertos para lunes a viernes y el fin de semana, donde el comportamiento de la energía es muy común.

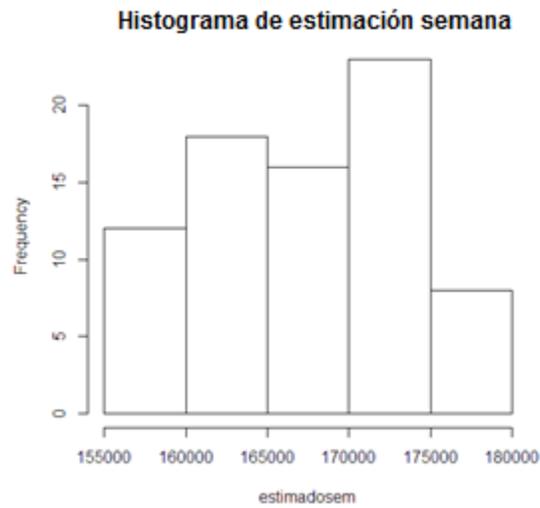


Figura 12. Gráfico de frecuencia de los datos de lunes a viernes

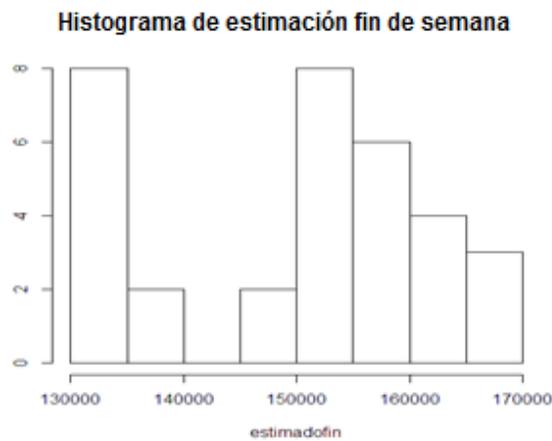


Figura 13. Gráfico de frecuencia de los datos del fin de semana

En la figura 12, se logra observar que de lunes a viernes, hay una alta frecuencia entre 165000 MWh y 175000 MWh, por lo que se puede intuir que el intervalo de alta densidad, que indica que ocurren datos con alta probabilidad, se encontrará en dicho rango de consumo. Lo anterior se corrobora con el valor de la mediana, media y tercer cuartil, observados en la tabla 17.

Tabla 17 Estadísticas básicas para el consumo en los días de lunes a viernes

Min.	1er Cuartil	Mediana	Media	3er Cuartil	Max.
155000	162000	168000	167700	173200	179000

Por otro lado para el fin de semana, acorde con la figura 13 y la tabla 18, puede verse que un intervalo de alta frecuencia de ocurrencia de consumo energético diario puede situarse entre 150000 MWh y 160000 MWh.

Tabla 18 Estadísticas básicas para el consumo en los días del fin de semana

Min.	1er Cuartil	Mediana	Media	3er Cuartil	Max.
130000	135400	154000	150800	160000	170000

Histograma del muestreo para pronóstico de demanda energética de una semana.

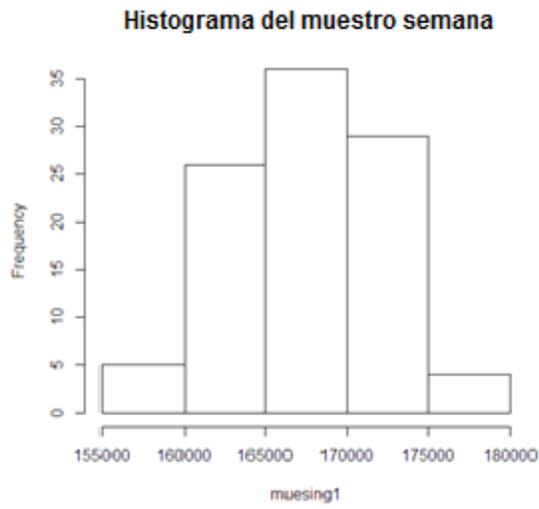


Figura 14. Gráfico de frecuencia para el pronóstico de lunes a viernes

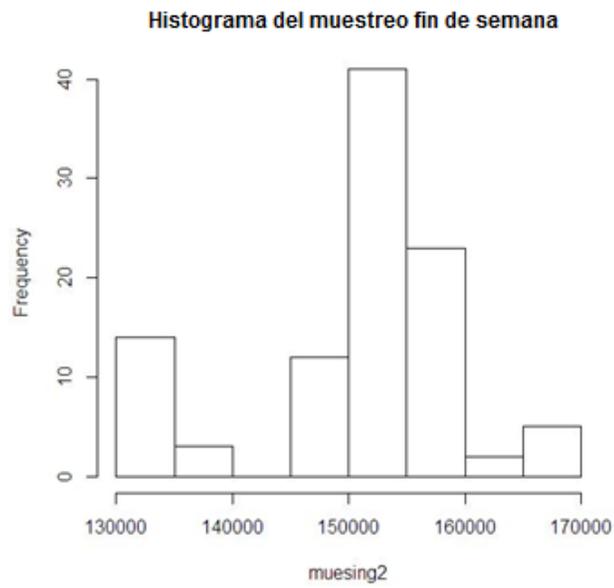


Figura 15. Gráfico de frecuencia para el pronóstico del fin de semana

Para la semana, se observa que el rango de valores de alta densidad estimado está entre 165000 MWh-170000 MWh, y para el fin de semana en la está entre 150000 MWh-155000 MWh (Figuras 14 y 15 y tablas19 y 20)

Tabla 19. Muestreo para pronóstico en semana

Min.	1er cuartil	Mediana	Media	3er. cuartil	Max.
157000	164800	168000	168200	172000	177000

Tabla 20. Muestreo para pronóstico del fin de semana.

Min.	1er cuartil.	Mediana	Media	3er cuartil.	Max.
130000	150000	153900	151600	156000	167000

A continuación se presenta el resultado del algoritmo descrito en la metodología de este trabajo (subsección 4.2), en el cual se tiene como información a priori los pesos por día, los datos elicitados, y la uniforme, como se describió previamente, para la media,

Intervalos de densidad de lunes a viernes después de la simulación

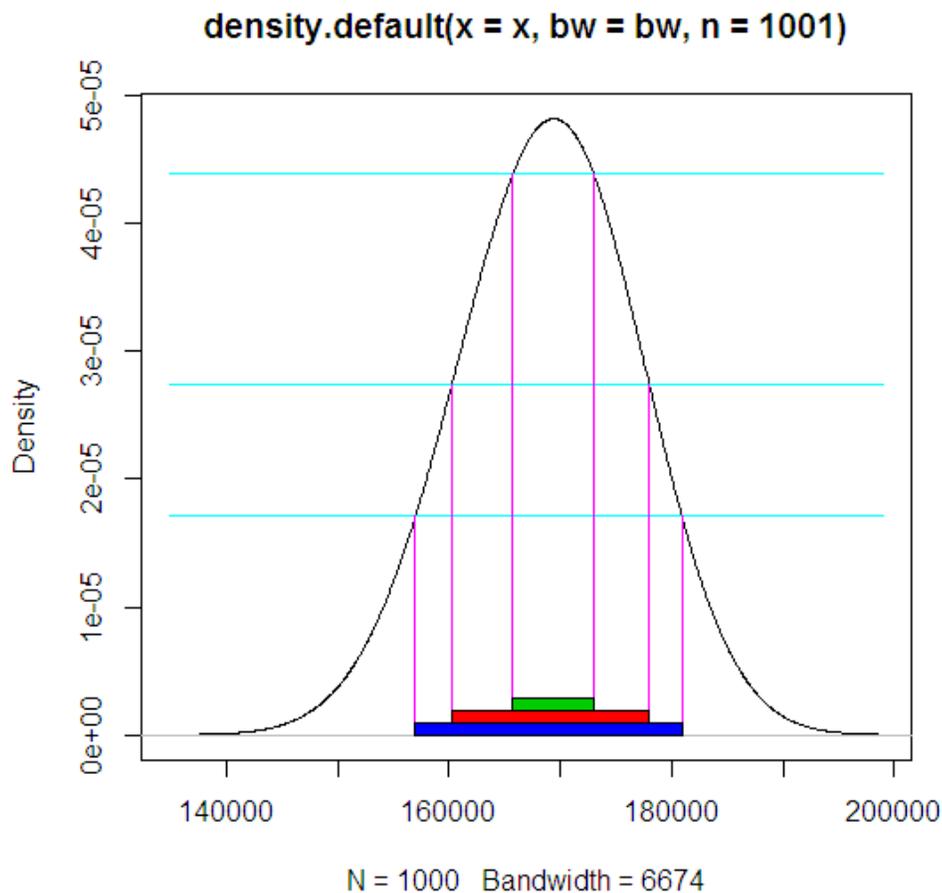


Figura 16. Gráfico de los intervalos de alta densidad del pronóstico de lunes a viernes

Ahora, se puede afirmar, a partir de la figura 16 que el intervalo de alta densidad, en donde estará ubicado el pronóstico será de 165000 Mwh a 172000 Mwh, fue acertada la primera estimación acerca de este, cualquier valor que esté dentro de este intervalo, se considera un pronóstico confiable, pero a partir de las 1000 iteraciones de la simulación podremos saber el valor medio de la simulación y ese será el pronóstico.

Intervalos de densidad después de la simulación para el fin de semana

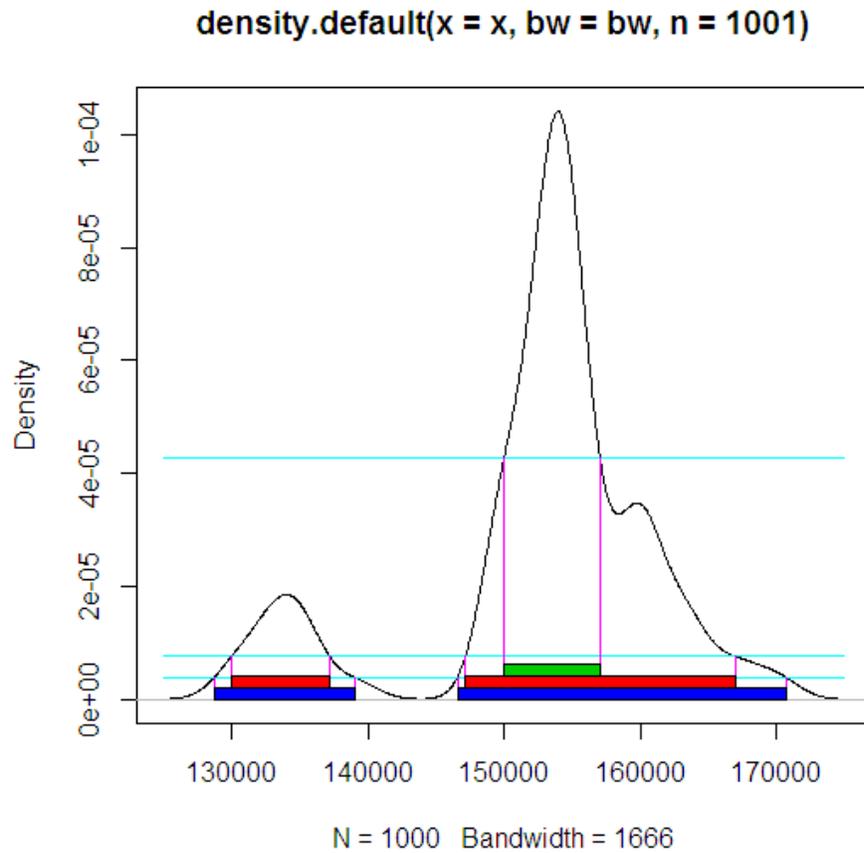


Figura 17. Gráfico de los intervalos de alta densidad del pronóstico del fin de semana

En la figura 17 observamos que el intervalo de alta densidad se encuentra alrededor de 150000 Mwh y 158000 Mwh, se corrobora la primera estimación que se realizó. La media de la simulación será la predicción de los datos futuros, este valor está situado en dicho intervalo de alta densidad

Tabla 21. Pronóstico para una semana no atípica

Fecha	Demanda Real (Mw)	Día	Pronóstico	Error relativo, semana común
17/12/2012	169756,92	Lunes	173572,2	2,25%
18/12/2012	170427,43	Martes	173305,1	1,69%
19/12/2012	173748,52	miércoles	167018,3	3,87%
20/12/2012	173399,14	Jueves	168861,6	2,62%
21/12/2012	170600,5	Viernes	169007,1	0,93%
22/12/2012	160288,86	Sábado	154644,4	3,52%
23/12/2012	145036,17	domingo	149644,1	3,18%
			MAPE	2,58%

Un MAPE bastante pequeño (tabla 21), teniendo en cuenta la cantidad de datos que se obtienen a partir de la elicitación a expertos Delphi. Es un modelo confiable para pronosticar a corto plazo, puesto que por esta técnica, solo obtenemos la media de toda la simulación para cada día, es decir, cada día (lunes, martes, miércoles...domingo) tiene un único pronóstico.

Ahora, en la tabla 23 se mostrará el pronóstico para días atípicos, para este proceso, fue necesario recurrir de nuevo a los expertos y preguntarles en que porcentaje, días como 24, 25 y 31 podrían disminuir o aumentar la demanda, con respecto a una semana en la que la demanda de energía se comporta de manera normal, tabla 22. Con base en esto y con el pronóstico obtenido de la simulación bayesiana, se calculó el pronóstico.

Tabla 22. Porcentaje disminución de energía

% de disminución promedio de energía	
24 de diciembre	15%
25 de diciembre	20%
31 de diciembre	15%

Tabla 23 Pronostico de demanda de energía para días atípicos

Fecha	Demanda Real (Mw)	Día	Pronóstico	% de disminución	Pronóstico real	Error relativo
24/12/2012	147566,05	Lunes	173572,2	15%	147536,37	0,020%
25/12/2012	131777,18	Martes	173305,1	20%	138644,08	5,211%
26/12/2012	160614,76	Miércoles	167018,3	0%	167018,3	3,987%
27/12/2012	162983,71	Jueves	168861,6	0%	168861,6	3,606%
28/12/2012	163891,65	Viernes	169007,1	0%	169007,1	3,121%
29/12/2012	155519,14	Sábado	154644,4	0%	154644,4	0,562%
30/12/2012	143156,82	Domingo	149644,1	0%	149644,1	4,532%
31/12/2012	143094,86	Lunes	173572,2	15%	147536,37	3,104%
					MAPE	3,018%

MAPE=2,58%

MAPE pronóstico=3,018%

El modelo es muy confiable para pronósticos a muy corto plazo, logrando un error relativo hasta del 0,020%. En resumen, el éxito de esta metodología reside en la confianza de los expertos. En este caso, las personas evaluadas dentro de XM compañía de expertos son profesionales que dedican la mayor

parte de su tiempo al despacho de energía en Colombia. La información a priori puede ser modificada con el mismo conocimiento de expertos, por ello, este método muestra flexibilidad y eficiencia a corto plazo y no requiere gran cantidad de datos o validar supuestos.

5.6 Modelo de regresión Gaussiana vía Monte Carlo por Cadenas de Markov.

Para esta técnica, se utilizó el lenguaje de programación para análisis estadístico R. Al igual que en las otras técnicas estadísticas, se realizaron varias estimaciones, a continuación expondremos las estimaciones con menor error.

En este modelo se incorporaron las siguientes variables:

- Tendencia t
- Día
- Rezago (Z_{t-1}).

Se realizó una simulación Montecarlo con 12000 iteraciones, y se hizo un quemado de las primeras 2000 variables aleatorias estimadas por el paquete estadístico MCMC regress, con el fin de garantizar estacionariedad en la serie, por lo que se espera que la prueba de Box Pierce se cumpla con una significancia del 5%.

Iteraciones = 2001:12000

Tamaño de muestra = 10000

1. Media y desviación estándar para cada variable, con el error estándar de la media

Tabla 24. Coeficientes Estimados

	Media	Desviación estándar	Error estándar previo (Naive)	Series de tiempo SE
Intercepto	-6,10E+06	7,75E+05	7,75E+03	7,75E+03
T	3,06E+03	3,87E+02	3,87E+00	3,87E+00
Día lunes	2,94E+04	8,74E+02	8,74E+00	8,74E+00
Día martes	2,41E+04	7,07E+02	7,07E+00	7,07E+00
Día miércoles	2,18E+04	7,57E+02	7,57E+00	7,57E+00
Día jueves	2,09E+04	7,68E+02	7,68E+00	7,46E+00
Día viernes	2,05E+04	7,68E+02	7,68E+00	7,42E+00
Día sábado	1,13E+04	7,66E+02	7,66E+00	7,66E+00
energía_1	5,56E-01	3,12E-02	3,12E-04	3,18E-04
sigma2	2,58E+07	1,37E+06	1,37E+04	1,40E+04

El modelo simula los parámetros de regresión y para cada uno genera 10000 datos, mostrando las estadísticas básicas en la tabla 24. La media es el estimador elegido como el coeficiente a incorporar en el modelo de regresión vía MCMC.

Ahora, para analizar la incorrelación y la estacionariedad de la serie, realizaremos la prueba de Box Pierce para cada variable simulada, los parámetros del modelo.

Tabla 25. Pruebas de estacionariedad e incorrelación

Prueba de Box-Pierce	
Variable	Valor P
Intercepto	0,6206
Tendencia T	0,6289
Día lunes	0,4435
Día martes	0,3584
Día miércoles	0,3424
Día jueves	0,1649
Día viernes	0,5753
Día sábado	0,5945
Rezago Zt-1	0,1287
Sigma2	0,0789

En la tabla 25 se observa que todas las variables cumplen la prueba los valores p están por encima del 5%, es decir los valores simulados son incorrelacionados,

Tabla 26 Pronostico de demanda de energía

Fecha	Día	Demanda Energía (Mw)	Zt-1	Tiempo	Pronóstico	MAPE
17/12/2012	Lunes	169756,92	147594,62	2012,95902	167044,72	1,60%
18/12/2012	Martes	170427,43	169756,92	2012,96175	174090,88	2,15%
19/12/2012	miércoles	173748,52	170427,43	2012,96448	172211,91	0,88%
20/12/2012	Jueves	173399,14	173748,52	2012,96721	173096,13	0,17%
21/12/2012	Viernes	170600,5	173399,14	2012,96995	172520,30	1,13%
22/12/2012	Sábado	160288,86	170600,5	2012,97268	161833,18	0,96%
23/12/2012	domingo	145036,17	160288,86	2012,97541	144780,33	0,18%
24/12/2012	Lunes	147566,05	145036,17	2012,97814	165681,24	12,28%
25/12/2012	Martes	131777,18	147566,05	2012,98087	161815,70	22,79%
26/12/2012	miércoles	160614,76	131777,18	2012,98361	150788,61	6,12%
27/12/2012	Jueves	162983,71	160614,76	2012,98634	165854,89	1,76%

28/12/2012	Viernes	163891,65	162983,71	2012,98907	166789,91	1,77%
29/12/2012	Sábado	155519,14	163891,65	2012,9918	158162,90	1,70%
30/12/2012	domingo	143156,82	155519,14	2012,99454	142187,82	0,68%
31/12/2012	Lunes	143094,86	143156,82	2012,99727	164695,20	15,10%

MAPE ajuste=1,90%

MAPE pronóstico=4,61%

Tabla 27. Comparación de resultados de los modelos Bayesianos para el pronóstico energético diario en Colombia.

Fecha	Demanda Real (Mw)	Día	Error relativo	
			MCMC	Delphi
17/12/2012	169756,9	Lunes	1,60%	2,25%
18/12/2012	170427,4	Martes	2,15%	1,69%
19/12/2012	173748,5	miércoles	0,88%	3,87%
20/12/2012	173399,1	Jueves	0,17%	2,62%
21/12/2012	170600,5	Viernes	1,13%	0,93%
22/12/2012	160288,9	Sábado	0,96%	3,52%
23/12/2012	145036,2	domingo	0,18%	3,18%
24/12/2012	147566,1	Lunes	12,28%	0,02%
25/12/2012	131777,2	Martes	22,79%	5,21%
26/12/2012	160614,8	miércoles	6,12%	3,99%
27/12/2012	162983,7	Jueves	1,76%	3,61%
28/12/2012	163891,7	Viernes	1,77%	3,12%
29/12/2012	155519,1	Sábado	1,70%	0,56%
30/12/2012	143156,8	domingo	0,68%	4,53%
31/12/2012	143094,9	Lunes	15,10%	3,10%
MAPE			4,62%	2,81%

En la tabla 28 se puede apreciar que el método Delphi se acerca mucho más a la realidad que la técnica MCMC, a la hora de pronosticar la demanda diaria energética de Colombia, generando un MAPE menor.

El método Delphi es el más eficiente de los dos bayesianos evaluados, para pronosticar demanda energética en días donde el consumo es atípico.

Tabla 28. Comparación del modelo clásico más eficiente con el modelo Bayesiano más eficiente.

Fecha	Demanda Real (Mw)	Error relativo	
		Dinámico	Delphi
17/12/2012	169756,92	0,11%	2,25%
18/12/2012	170427,43	1,22%	1,69%
19/12/2012	173748,52	0,37%	3,87%
20/12/2012	173399,14	0,16%	2,62%
21/12/2012	170600,5	0,94%	0,93%
22/12/2012	160288,86	1,25%	3,52%
23/12/2012	145036,17	2,07%	3,18%
24/12/2012	147566,05	2,06%	0,02%
25/12/2012	131777,18	0,25%	5,21%
26/12/2012	160614,76	1,05%	3,99%
27/12/2012	162983,71	0,76%	3,61%
28/12/2012	163891,65	0,15%	3,12%
29/12/2012	155519,14	0,99%	0,56%
30/12/2012	143156,82	2,91%	4,53%
31/12/2012	143094,86	0,54%	3,10%
	MAPE	0.99%	2,81%

Ahora, analizando los dos modelos más eficientes en la tabla 28, uno clásico y otro bayesiano, vemos que el Dinámico (modelo clásico) es el más asertivo para este tipo de demanda, esto se debe a que esta técnica permite realizar análisis sobre las variables exógenas, admitiendo variables de nivel que llevan a captar muy bien los picos atípicos. Pero no se puede dejar a un lado el método Delphi por simulación bayesiana, ya que este modelo alcanzó niveles muy buenos, con la ventaja de que no se necesitaron muchos datos, por lo tanto, si se presentara una variación inesperada en la demanda energética, por aspectos macro, los modelos clásicos no podrían captar dicha diferencia, ahí es donde esta metodología contemporánea podría jugar un papel muy importante, dependiendo netamente del criterio de expertos.

Un resumen de todos los resultados se presenta en la tabla 29 resaltando el modelo dinámico como el de mejor desempeño según el MAPE pronosticado y las técnicas bayesianas como una alternativa, cuando no se hace necesario el cumplimiento de premisas o no se tienen muchos datos históricos.

Tabla 29. Síntesis de resultados

Modelo	MAPE pronostico (%)	Premisas	Ventajas/Desventajas encontradas
Suavización exponencial simple	7,29	Residuales incorrelacionados	Incumple supuesto de correlación. Error de pronóstico inaceptable.
Suavización exponencial de Holt	6,91	Residuales incorrelacionados	Incumple supuesto de correlación. Error de pronóstico inaceptable.
SARIMA	6,2	Residuales incorrelacionados, y	Cumple incorrelación. Incumplimiento de normalidad, error de

Modelo	MAPE pronóstico (%)	Premisas	Ventajas/Desventajas encontradas
		normalidad	pronóstico poco aceptable.
Modelo Dinámico autorregresivo con errores SARIMA	0,99	Residuales incorrelacionados, y normalidad	Incumplimiento de normalidad. Error de pronóstico óptimo (el mejor de todos). Buen resultado para estimar atípicos.
MCMCregress	4,62	A priori de parámetros con distribución normal, verosimilitud Normal. Simulación de variable aleatoria para cada parámetro, incorporada en el paquete. Incorrelación en las variables aleatorias simuladas.	Error aceptable. Cumple incorrelación de las variables aleatorias simuladas (parámetros) con el paquete usado.
Modelo de elicitación Bayesiana	2,81	A priori uniforme para la media, datos con distribución empírica elicitada. Remuestreo y simulación de variable aleatoria uniforme en el intervalo de alta densidad, estimando la media de pronóstico.	Error muy aceptable. Los pesos de cada día son un importante insumo para esta técnica. Buen resultado para estimar atípicos.

CONCLUSIONES

1. De todos los modelos estimados y analizados, se encontró que el modelo Dinámico con errores SARIMA fue el óptimo para lograr el menor error de pronóstico (0,99%), acercándose más a la realidad, Esto se debe a que al incorporar variables significativas como los retardos, las variables ficticias, o de nivel, se explicaron mejor los cambios estructurales en la serie, logrando con ello proporcionar una herramienta que facilita la optimización del abastecimiento energético diario de Colombia, por la confianza del pronóstico.
2. El modelo SARIMA para el consumo energético, que es uno de los utilizados para estimar el pronóstico de demanda de energía por el Centro Nacional de Despacho (CND) mostró un problema con la validación de supuestos requerida, y además, fue superado por la eficiencia del modelo dinámico, lo cual sugiere replantearlo, con la herramienta proporcionada en este trabajo.
3. Una gran ventaja que tiene el método de regresión por MCMC es que no requiere un cumplimiento de los mismos supuestos de un modelo clásico. En este trabajo se muestra que el método simula los parámetros de un modelo de regresión, llegando a generarlos sin dependencia temporal. Se mostró además, una eficiencia en el pronóstico aceptable, brindando una alternativa cuando existen pocos datos y modelos difíciles de ajustar.

4. El método de predicción bayesiana por medio de la elicitación a expertos evidencia su favorabilidad para pronósticos a corto plazo, y sería una gran herramienta si no se cuenta con datos históricos.

5. Con este trabajo de investigación, el Centro Nacional de Despacho (CND) puede encontrar métodos para optimizar el abastecimiento energético diario para Colombia, obteniendo predicciones de demanda muy cerca de los reales, e incluso, usando información de expertos. De hecho, los expertos en pronóstico destacaron la importancia de los métodos estadísticos bayesianos para aproximarse a los valores de consumo energético de días atípicos, lo cual representa una ganancia alrededor de la planeación previa de abastecimiento que deben realizar.

RECOMENDACIONES

1. Cuando se tengan datos históricos confiables, el Centro Nacional de Despacho (CND) puede hacer uso del modelo dinámico para la predicción, y así poder garantizar confiabilidad y calidad en el servicio de despacho de energía eléctrica para el país.
2. Si no se cuenta con históricos confiables, o si hay un cambio a nivel macro en el país que modifique el comportamiento de consumo de energía eléctrica de manera inesperada, el CND puede utilizar la simulación Bayesiana por medio de la elicitación a expertos, obteniendo datos de gran importancia para la estimación de los modelos, como los valores de los pesos que caracterizan el tipo de demanda o el clima, y de esta obtener niveles de predicción a corto plazo totalmente confiables.
3. Otros trabajos alrededor del tema se pueden orientar incorporando de los precios de la energía a la modelación para el consumo energético, como también, evaluar con un modelo de optimización dinámico los costos del abastecimiento de energía en todo el país.

BIBLIOGRAFÍA

- Medina Flechas, S. C., & Mutis, H. (2004). Ilustración de la aplicación de la metodología arima para pronosticar la demanda de productos de consumo masivo para la limpieza del hogar en el mercado colombiano. *Universidad de los Andes*.
- Latin - Focus*. (2009). Recuperado el 3 de Octubre de 2012, de <http://www.latin-focus.com/spanish/>
- Aguilera Aguilera, J. R. (2009). Fundamentos de producción para licenciados en contaduría. *Universidad Michoacana de San Nicolas de Hidalgo* .
- Anderson, D., Sweeney, D., & Williams, T. (2009). *Métodos cuantitativos para los negocios*. Cengage Learning Editores.
- Andrade Sosa, H. H., Parra Valencia, J. A., & Pineda Ballesteros, E. (2004). Dinámica de sistemas y econometría: en la búsqueda de posibilidades o imposibilidades de reconocimiento a partir de la reflexión de una experiencia en economía. *Encuentro Colombiano de Dinámica de sistemas* .
- Arce, & Mahía. (2003). Modelos Arima. *Programa Citius*. Universidad Autónoma de Manizales

- Bellas Parias, L. D., & Martinez Calle, S. (2005). Una metodología de series de tiempo para el area de la salud. Caso práctico .
- Bijak, J., & Wiśniowski, A. (2010). Bayesian forecasting of immigration to selected European countries by using expert knowledge. *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 775-796.
- Botero, S., & Cano, J. (2008). Análisis de series de tiempo para la predicción de los precios de la energía en la bolsa de colombia. *Revista cuadernos de economía*.
- Bravo, M. d., & Arrieta, J. J. (2005). El método delphi. Su implementación en una estrategia. *Revista Iberoamericana de Educación*.
- Bruce , Bowerman, & Oconnell. (2007). *Pronósticos, series de tiempo y regresión: un enfoque aplicado*. Cengage Learning Editores.
- Cabanzo Niño , Y. (2011). aplicación de drp (plan de requerimiento de distribución) y planeación de la producción para la empresa plásticos monroy en la línea de producción de bolsas plásticas para basura. *Universidad Libre de Colombia*.
- Caridad, & Ocerin. (1998). *Econométrica: Modelos Econométricos y series temporales* . Reverté .
- Castro Zuluaga, C., & Botero Escobar , S. (2012). Metodología para la selección del parámetro alpha en el modelo de Suavización Exponencial: Un enfoque empírico. *Tenth LACCEI Latin American and Caribbean Conference*.
- Chavez, N. (1997). Modelos Arima. *Revista Ciencia y Cultura* .

- Conchi, A. (2012). Métodos de cadenas de Markov Monte Carlo. *Universidad Carlos III de Madrid*.
- Contreras , J., & Idrovo , B. (2009). Un Modelo SARIMA para Predecir la Tasa de Desempleo de Chile. *Munich Personal RePEc Archive*.
- De Arce, R., & Mahía , R. (2007). Modelos ARIMA. En *Modelos ARIMA* . U.D.I Econométrica e Informatica .
- Esparza Catalán, C. (2000). Series Temporales. *Laboratorio de Estadística* .
- Fornero, R. (2007). *Cronología fotográfica de las finanzas* . Universidad Nacional de Cuyo.
- Garcia Tabares , A. I. (2008). Propuesta de una metodología de estimación de demanda de accesorios de joyería, en una empresa de venta directa Colombiana. *Universidad Nacional de Colombia*.
- Gardner. (2006). Exponential smoothing: The state of the art Part II. *International Journal of Forecasting*, 637-666.
- Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian Data Analysis*. (C. & Hall, Ed.) (Second Edi., p. 668).
- Gill, J. (2002). *Bayesian Methods-A social and Behavioral Sciences Approach* (p. 459).
- Gómez, G. A., Carreño, J. C., & Zambrano Caviedes , J. (2011). Herramienta de software para el pronóstico de demanda horaria de potencia eléctrica en el sistema eléctrico de Codensa S.A. ESP. *Tecnura*.
- González , A., & Gómez, A. (2007). La relación entre los precios de los alimentos y la concentración de los supermercados en Chile: evidencia

de un modelo dinámico de panel y análisis de los impactos de las fusiones propuestas en la industria. *Economía y Negocios Universidad de Chile* .

Guerra, J., Sánchez, G., & Reyes, B. (1997). Series de tiempo para predecir la inflación en Venezuela. *Colección banca central y sociedad*.

Guisán, M. C. (2007). Modelos econométricos dinámicos y desarrollo económico: Análisis del salario real, la productividad y el empleo en los países de la OCDE, 1965-2005. *Universidad de Santiago España* .

Hanke, & Reitsch. (1996). *Pronósticos en los negocios*. Prentice Hall.

Hillier, F. S., & Lieberman, G. J. (2002). *Investigación de operaciones*. México: Mc Graw-Hill.

Hyndman, Koehler, Ord, & Snyder. (2008). *Forecasting with exponential smoothing: The state space approach*. Springer.

Jarque, & Bera. (1987). Test for Normality of Observations and Regression Residuals. *International Statistical*, 163-172.

Ji, X., & Chen, W. (2012). Evaluation of supply chain risk and its empirical analysis for apparel processing enterprises. *Journal of Convergence Information Technology*, 44-53.

Landeta, J., Barrutia, J., & Lertxundib, A. (2011). Hybrid Delphi: A methodology to facilitate contribution from experts in professional contexts. *Technological Forecasting and Social Change*, 1629–1641.

- Landeta, J., Mateya, J., & Ruíza, V. (2008). Results of a Delphi survey in drawing up the input–output tables for Catalonia. *Technological Forecasting and Social Change*, 32-56.
- Larget, & Simon, D. (2000). Análisis Bayesiano en biología y evolución molecular. Departamento de las matemáticas y de la informática. *Universidad Duquesne*.
- Ljung. (1978). On a Measure of a Lack of Fit in Time Series Models. *Biometrika*, 297-303.
- Mahia. (2001). Notas sobre estacionariedad de series temporales: Definición y contraste de Raíces Unitarias. *Seminario de utilización de los modelos econométricos para la simulación y predicción de la economía española*.
- Marshall , P. (1986). Predicción de inflación con modelos de series de tiempo multiples . *Cuadernos de Economía* .
- Maté, C. (2006). Modelos ARIMA. *Comillas Madrid Universidad Pontificia ICAI ICADE*.
- Mate, C. (s.f.). *Universidad Pontificia icai icade comillas*. Recuperado el 3 de Septiembre de 2012, de http://www.upcomillas.es/presim/documentos/pred_clasi_03.pdf
- Medina, & Garcia. (2005). Predicción de demanda de energía en Colombia mediante un sistema de inferencia difuso neurona. *ISSN 0120-9833*, 15-24.

- Montgomery, D., Peck, E. & Vining, G. (2012). *Introduction to Linear Regression Analysis*. John Wiley & Sons. 5ª Ed.
- Murillos, J., Trejos, A., & Carvajal, P. (2003). Estudio del pronóstico de la demanda de energía eléctrica, utilizando modelos de series de tiempo. *Scientia et Technica*.
- Pettyjohn, S. L. (2011). Global Defense Posture, Santa Monica, *RAND Corporation*, 2012. <http://www.rand.org/pubs/monographs/MG1244>.
- Pierce, B., & Pierce, D. (1970). Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models. *American Statistical Association*.
- Ramírez, Cardona, & González. (2012). Comparación de metodologías estadísticas en el pronóstico de la demanda. *Universidad Pontificia Bolivariana*.
- Robalino, A. (2005). *Análisis de series de tiempo*. Repositorio Escuela Superior Politécnica de Litoral.
- Roca Sagalés, O., & Pereira, A. (1998). Impacto de la inversión en infraestructuras sobre el producto, la ocupación y la inversión privada en España. *Revista Española de Economía*.
- Rueda, V., Velásquez Henao, J. D., & Franco Cardona, C. (2011). Avances recientes en la predicción de la demanda de Electricidad usando modelos no lineales. *Scielo*.
- Shapiro. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 591-611.

- Stevens, S. J. (1998). *Investigación de Operaciones, Un Enfoque Fundamental*. Mexico: Mc Graw Hill.
- Synder, R. (2002). Forecasting sales of slow and fast moving inventories. *European Journal of operational Research*, 648-699.
- Synder, R. (2002). Forecasting sales of slow and fast moving inventories. *European Journal of operational Research*, 648-699.
- Universidad de Washington. (20 de Mayo de 2013). *mcmcpack*. Recuperado el 2013, de <http://mcmcpack.wustl.edu/>
- Valencia, M., & Sierra, J. (2009). *Método de Elicitación Bayesiana*. Medellín.
- Vera, M. & Bustamante, J. (2007). Modelo dinámico para la generación de pronóstico usando redes neurales artificiales (RNA). *Visión Gerencial*, 130-142.
- Vermorel. (2011). Definición de Estacionalidad. *Lokad*. Disponible en: <http://www.lokad.com/es/definicion-estacionalidad>
- Villavicencio, J. (2010). Introducción a series de tiempo. Puerto Rico .
- Winston, W. L. (2004). *Investigación de operaciones: aplicaciones y algoritmos* (4ta ed.). México: Thomson.
- Zuluaga Mazo, A., Molina Parra, P. A., & Guisao Giraldo, É. Y. (2011). La pleneación de la demanda como requisito para la gestion de las cadenas de suministro en las empresas en Colombia. *Revista Politécnica* .