

**UNIVERSIDAD NACIONAL DE INGENIERÍA**

**FACULTAD DE INGENIERÍA MECÁNICA  
SECCIÓN DE POST GRADO Y SEGUNDA ESPECIALIZACIÓN**



**TESIS:**

**“ANÁLISIS ESTOCÁSTICO ARIMA PARA EL  
MODELAMIENTO Y PREDICCIÓN DE LA DEMANDA  
ELÉCTRICA EN EL SECTOR RESIDENCIAL DE LIMA SUR”**

**PARA OPTAR EL TÍTULO DE MAESTRO EN CIENCIAS CON  
MENCIÓN EN ENERGÉTICA**

**ING. HERNÁN ANTONIO ORBEZO URQUIZO**

**LIMA – PERÚ**

**2011**

## **RESUMEN**

En la presente investigación se aplica el análisis de series temporales bajo un enfoque estocástico ARIMA (Modelo Estocástico Lineal Integrado Autorregresivo- Medias Móviles), para la realización de proyecciones de la demanda de energía eléctrica residencial en la zona de Lima Sur (variable con tendencia, estacionalidad, ciclaje y aleatoriedad), habiéndose demostrado que permite mejorar la eficiencia y bondad predictiva al comparar y validar los resultados obtenidos con la técnica determinística de predicción Winter.

Una adecuada predicción de la energía demandada es determinante para decisiones de inversión en redes eléctricas, pues cuanto más precisas sean las predicciones menores serán los riesgos de incurrir en inversiones innecesarias; en ese sentido, ésta investigación representa la primera intención de formular un procedimiento técnico validado que respalde las decisiones de inversión a nivel de distribución eléctrica. Como resultado del análisis realizado se determinó que el consumo de energía residencial de Lima Sur tendrá un crecimiento conservador sostenido del 3,5% en promedio para el periodo 2010-2012. Incremento que también representa el crecimiento del consumo de energía en baja tensión, pues ésta serie energética presenta el mismo comportamiento que el consumo de energía residencial para la misma zona de análisis.

La metodología ARIMA proporciona mejores resultados predictivos, tanto gráficamente como en margen de error y en complejidad de la variable en estudio, frente cualquier otro método como el alisado exponencial compuesto Winter, demostrando la Hipótesis Nula formulada para la investigación, de ofrecer resultados con mejor eficiencia y bondad predictiva. Así con el método ARIMA se obtienen resultados con menor error absoluto porcentual promedio (1,59%) y menor suma de cuadrado de errores (1359).

### **Palabras Clave:**

Predicción, demanda eléctrica residencial, modelos ARIMA, serie temporal, procesos estocásticos, validación.

## **SUMMARY**

In the present research applied time series analysis in a stochastic approach ARIMA (Autoregressive Integrated Moving Average), for making forecasts of demand for residential electricity in Lima Sur (variable with trend, seasonality, cycling and randomness), having shown that improves efficiency and predictive goodness to compare and validate the results obtained with the deterministic prediction Winter technique.

An adequate prediction of the energy demand is critical to investment decisions in electricity networks, because the more accurate the predictions lower the risk of incurring unnecessary investments; in that sense, this research represents the first attempt to formulate a technical procedure validated to support investment decisions at the level of electrical distribution. As a result of the analysis determined that the residential energy consumption will grow Lima Sur sustained conservative 3.5% on average for the period 2010-2012. Increase also represents the growth of energy consumption at low voltage, as this energy range has the same behavior as the residential energy consumption for the same area of analysis.

The ARIMA methodology provides better predictive results, both graphically and in margin of error and complexity of the variable under study, compared with other methods such as exponential smoothing compound Winter, proving the null hypothesis formulated for the research, to provide results with better efficiency and predictive goodness. So with the ARIMA method results are obtained with lower mean absolute percentage error (1.59%) and lowest sum of squared errors (1359).

### **Keywords:**

Forecast, residential electricity demand, ARIMA models, time series, stochastic processes, validation.

## ÍNDICE

<b>AGRADECIMIENTOS</b>	<b>2</b>
<b>RESUMEN</b>	<b>3</b>
<b>SUMMARY</b>	<b>4</b>
<b>INTRODUCCIÓN</b>	<b>7</b>
<b>CAPÍTULO I – ASPECTOS GENÉRICOS DE LA INVESTIGACIÓN</b>	
1.1 Planteamiento del problema	9
1.2 Formulación del problema	10
1.3 Objetivos	10
1.3.1 Objetivo General	10
1.3.2 Objetivo Específico	10
1.4 Hipótesis	11
1.5 Metodología y nivel de investigación	11
1.6 Importancia de la investigación	11
1.7 Predicción de la demanda eléctrica. Experiencia nacional e internacional	12
<b>CAPÍTULO II - MODELOS PARA LA PREDICCIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA</b>	
2.1 Importancia de la predicción de la demanda de energía eléctrica	21
2.2 Factores que influyen en la demanda de energía eléctrica	23
2.3 Modelamiento de la predicción de energía eléctrica	24
2.4 Introducción a las series temporales	27
2.4.1 Series temporales	28
2.4.2 Análisis de la predicción con series temporales	31
2.5 Estudio comparativo de modelos univariantes	32
2.5.1 Métodos aplicables a series sin tendencia ni estacionalidad	33
2.5.2 Método de Holt	36
2.5.3 Método de Holt-Winter	36
2.5.4 Modelo Box-Jenkins	37
2.6 Formulación de la Metodología Estocástica ARIMA	39
2.6.1 Modelos ARIMA	40
2.6.2 Etapas para la elaboración del Modelo ARIMA	43
2.6.3 Estructura de cálculo con Modelos ARIMA	44
<b>CAPÍTULO III - MODELAMIENTO Y PREDICCIÓN</b>	
3.1 Consumo de energía eléctrica en el sector residencial	47
3.1.1 Características de las redes eléctricas	49
3.1.2 Importancia de la serie energética	50
3.1.3 Análisis preliminar	52
3.2 Análisis de la bases de datos	53

3.3	Proceso de cálculo predictivo. Método de WINTER	55
3.4	Resultados. Metodología ARIMA	58
3.4.1	Análisis de la estabilidad en varianza	59
3.4.2	Análisis de estabilidad en media regular y media estacional	60
3.4.3	Identificación de órdenes autoregresivos y medias móviles (p, P, q, Q)	60
3.4.4	Estimación de parámetros o coeficientes	61
3.4.5	Validación del modelo ARIMA óptimo por prueba de residuos	62
3.4.6	Validación de predicciones con modelo ARIMA óptimo por estadísticos de error	63
3.4.7	Predicción con modelo ARIMA representativo	64
<b>CAPÍTULO IV – DISCUSIÓN DE RESULTADOS Y CONTRATACIÓN DE LA HIPÓTESIS</b>		
4.1	Importancia de los resultados obtenidos	68
4.2	Contratación de la hipótesis	69
	<b>CONCLUSIONES</b>	<b>71</b>
	<b>RECOMENDACIONES</b>	<b>73</b>
	<b>BIBLIOGRAFÍA</b>	<b>74</b>
	<b>ANEXOS</b>	<b>77</b>

## INTRODUCCIÓN

El proyecto de investigación plantea el uso de modelos de series temporales bajo un enfoque estocástico ARIMA, para la realización de proyecciones de la carga vegetativa en una zona de concesión eléctrica, los cuales permitan mejorar la eficiencia y bondad predictiva, estableciendo procedimientos de fácil comprensión con la ayuda del programa informático SPSS<sup>1</sup>. Cabe señalar que una adecuada predicción de la energía demandada es una guía para eventuales decisiones de incrementar la capacidad generadora instalada, por cuanto a medida que las predicciones sean más precisas, menores serán los riesgos de incurrir en inversiones innecesarias y/o de causar insatisfacción de los usuarios. En ese sentido, en la zona de concesión de análisis se han realizado inversiones por un monto total de 47.1 millones US\$ para el año 2008, representando un crecimiento del 27% respecto del año 2007<sup>2</sup>.

El desarrollo de la investigación está dividido en cuatro capítulos, incluyendo, además conclusiones, recomendaciones y anexos que sustentan los procedimientos y cálculos realizados. En el primer capítulo se plantea la formulación del problema en la que se sustenta la investigación, también se hace mención de los objetivos e hipótesis de la investigación, así como la importancia de la predicción de demanda de energía eléctrica.

En el segundo capítulo se efectúa un estudio sobre la importancia de la predicción de la demanda de energía, y de los factores que influyen en su variación y los modelos para su análisis. También se realiza análisis comparativo de los modelos de previsión univariantes de la demanda de energía más utilizados actualmente, previa introducción al tema de series temporales.

En el tercer capítulo, se desarrolla la parte central del presente trabajo, es decir, el estudio de la formulación de la metodología estocástica ARIMA, así como el proceso de cálculo predictivo de la serie energética de la demanda de energía residencial, previa introducción al análisis

---

<sup>1</sup> Statistical Package for Social Sciences (Paquete Estadístico para Ciencias Sociales)

<sup>2</sup> Información obtenida de página web [www.luzdelsur.com.pe](http://www.luzdelsur.com.pe)

exploratorio de datos con el software SPSS. Asimismo, se desarrolla el proceso de cálculo predictivo bajo el método de Winter, efectuándose el análisis comparativo de los resultados obtenidos con la finalidad de demostrar la hipótesis nula planteada.

En el cuarto capítulo, se refuerza el sustento técnico de los resultados obtenidos, describiendo la importancia de los resultados obtenidos y principalmente efectuando la contrastación de los mismos respecto de la predicción realizada mediante alisado Winter.

Finalmente, se presentan las conclusiones obtenidas en la investigación, así como recomendaciones sobre posibles futuras líneas de investigación.

## **CAPÍTULO I**

### **ASPECTOS GENÉRICOS DE LA INVESTIGACIÓN**

#### **1.1 PLANTEAMIENTO DEL PROBLEMA**

Una característica importante de las sociedades industrializadas y en vías de desarrollo es el consumo de energía eléctrica, representando un factor básico para la producción en diversos sectores, entre los que se encuentran el industrial, comercial y residencial o doméstico. El creciente nivel de vida en los países en vías de desarrollo está asociado a una mayor demanda de energía eléctrica doméstica o residencial debido, entre otras causas, a la mayor utilización de aparatos electrodomésticos diversos.

La finalidad fundamental de una concesionaria eléctrica es satisfacer las necesidades energéticas de los habitantes bajo su ámbito de acción tan económicamente como sea posible, con un nivel aceptable de calidad, fiabilidad y seguridad. Lográndose a pesar de factores externos, como el traslado de numerosas plantas industriales, el cambio en el tipo de aparatos eléctricos empleados por sus clientes y la frecuencia con que son utilizados.

Según Lenton Ch.<sup>3</sup>, Latinoamérica enfrentará la insuficiencia de infraestructura en sistemas de generación y transmisión, considerando las expectativas sobre su crecimiento transversal que se mantendrá en torno al 4% anual. Además, señala que los inversionistas continuarán valorando los mercados eléctricos latinoamericanos como apuestas seguras, mientras EEUU y Europa luchan por recobrase de la crisis de la deuda. Lenton<sup>3</sup> estima que para el 2030 las nuevas inversiones en todo el sector eléctrico llegarán a los US\$ 600,000 millones en Latinoamérica, suma que en gran parte provendrá del sector privado.

Todo ello crea la necesidad que las empresas eléctricas cuenten o desarrollen procedimientos para el cálculo de las predicciones de sus principales variables energéticas. Si bien es cierto que actualmente en nuestro país, algunas empresas eléctricas vienen desarrollando la metodología ARIMA para toda su zona de concesión, sin embargo, no

---

<sup>3</sup> Publicación "Energía Eléctrica, Proyecciones 2011"



consideran las características particulares de cada cliente, ni realizar un procedimiento de verificación como se desarrolla en el presente trabajo de investigación. Es decir, sin una metodología adecuada las áreas de planificación de nuestras empresas eléctricas no tienen la capacidad de realizar una programación de sus proyectos de inversión para ejecutar anualmente como resultado de una planificación técnicamente sustentada, sino proyectan sus inversiones basados generalmente en criterios de prioridad sobre redes eléctricas que presentan problemas de sobrecarga, lo cual a futuro les representa problemas en cuanto a la calidad del servicio que prestan y a la limitada capacidad de vender más energía a sus clientes, en desmedro de las ganancias que se puedan obtener a nivel de empresa. Además, en las áreas de planificación de las empresas de distribución eléctrica se manejan variados criterios para la predicción de la demanda de energía, de tal forma que los profesionales de la misma área no se ponen de acuerdo sobre el tema; generándose las predicciones en base a opinión de varios expertos, sin ningún sustento técnico ni validado.

En ese sentido, la investigación formula una metodología que uniformice los criterios para la predicción de la demanda eléctrica a corto plazo, logrando que ésta sea más precisa y eficiente que otras usadas actualmente, buscando con ello proporcionar una herramienta técnica que respalde las decisiones de inversión en infraestructura eléctrica para satisfacer las necesidades energéticas de las empresas eléctricas de distribución logrando tener un adecuado sistema eléctrico que soporte los incrementos de carga que puedan ocurrir en el tiempo.

## **1.2 FORMULACIÓN DEL PROBLEMA**

Habiendo presentado la realidad problemática, la formulación del problema se traduce en las siguientes interrogantes: ¿De qué manera influye el análisis estocástico ARIMA en el modelamiento y predicción de la energía residencial?

## **1.3 OBJETIVOS**

### **1.3.1 OBJETIVO GENERAL**

Aplicando el análisis estocástico ARIMA se logrará modelar y predecir la demanda eléctrica del sector residencial en Lima Sur para un periodo de tres años, dando respaldo técnico sobre las decisiones de inversión en infraestructura eléctrica.

### **1.3.2 OBJETIVOS ESPECÍFICOS**

- Construir el modelo ARIMA univariante de la demanda mensual de energía eléctrica residencial en la zona sur de Lima, basado en los datos de consumo correspondientes a los meses del periodo entre Enero.1996 y Diciembre.2009.

- Realizar proyecciones de la carga vegetativa en la zona sur de Lima Metropolitana, que permita mejorar la eficiencia y bondad predictiva para la toma de decisiones en inversión de infraestructura eléctrica.
- Verificar el modelo desarrollado, comprobando si la predicción obtenida arroja valores con menor porcentaje de error respecto de otros modelos de estimación utilizados, como el modelo determinístico de alisados exponenciales.

#### **1.4 HIPÓTESIS**

Utilizando el análisis estocástico ARIMA para modelar y predecir la energía residencial, se obtendrá mejor eficiencia y bondad predictiva respecto de los resultados obtenidos por otras técnicas determinísticas como los alisados exponenciales.

#### **1.5 METODOLOGÍA Y NIVEL DE INVESTIGACIÓN**

- **Tipo de investigación:** Por el tipo de la investigación, el presente trabajo reúne las condiciones metodológicas de una investigación aplicando un método hipotético-deductivo y correlacional. Se utiliza conocimientos sobre el uso de modelos de series temporales bajo el enfoque estocástico ARIMA, a fin de aplicarlas en la realización de proyecciones de la energía eléctrica residencial.
- **Nivel de investigación:** De acuerdo a la naturaleza del trabajo de la investigación, reúne por su nivel las características de un estudio descriptivo y explicativo.

#### **1.6 IMPORTANCIA DE LA INVESTIGACIÓN**

La importancia radica en que una adecuada predicción de la energía eléctrica es determinante para decisiones de inversión en redes eléctricas, pues cuanto más precisas sean las predicciones menores serán los riesgos de incurrir en inversiones innecesarias. Logrando en consecuencia una adecuada priorización de los proyectos de inversión para la ampliación y/o reforzamiento de las redes eléctricas, tanto al nivel de transmisión en alto voltaje como de distribución en medio y bajo voltaje; lo cual también podrá servir para la formulación de un plan de expansión de nuevos servicios de electrificación residencial en cualquier zona de concesión.

Así, los problemas de exceso de capacidad de transformación de potencia o por el contrario de capacidad insuficiente pueden tener costos muy elevados, no sólo en términos económicos, sino también en la imagen que el cliente percibe en cuanto a la calidad del servicio que recibe. Además, permitirá una fijación tarifaria más eficiente bajo el concepto de desagregación de demandas, pues está demostrado que el cálculo de los pronósticos de

los componentes de una variable ofrece mejores resultados predictivos que pronosticar la variable global directamente.

Finalmente, se dará sustento técnico para la predicción de energía eléctrica a corto plazo mediante un nuevo procedimiento para la determinación del pronóstico de los consumos de energía residencial con característica de alta eficiencia predictiva basada en el rigor matemático-estadístico. Además, a la fecha no han realizado trabajos que aborden formalmente el tema de la estimación con metodología ARIMA de la demanda de energía eléctrica de la carga vegetativa en cualquier zona de concesión, siendo importante que cada empresa eléctrica privada pueda proyectar futuras inversiones en su sistema eléctrico como base para una adecuada toma de decisión.

## **1.7 PREDICCIÓN DE LA DEMANDA ELÉCTRICA. EXPERIENCIA NACIONAL E INTERNACIONAL**

Los métodos para la predicción de la demanda de energía eléctrica han sufrido una profunda evolución durante las últimas décadas. En el período anterior a 1970 los métodos de extrapolación de la tendencia bastaban para explicar el aumento de la energía demandada. A partir de 1970 las enormes incertidumbres que emergieron como consecuencia de la primera crisis del petróleo de 1973, hicieron necesario el desarrollo y aplicación de técnicas mucho más sofisticadas entre las que se encuentran los métodos econométricos y los métodos ARIMA<sup>4</sup>.

La predicción de la demanda de energía eléctrica ha recibido escasa atención hasta la década de 1970, ya que el entorno económico anterior a la primera crisis del petróleo de los años 1973-74 y en particular durante las décadas de 1950 y 1960, se podía considerar que el bajo precio de los combustibles utilizados en la generación de energía eléctrica era bastante predecible, los tipos de interés eran moderados y no existían demasiados problemas para que los mercados de capitales proporcionen los fondos necesarios para la expansión del sector eléctrico; sin embargo, la mencionada crisis introdujo numerosos factores de incertidumbre originando que los departamentos de planificación de las empresas eléctricas se viesen obligados a elaborar pronósticos de demanda para horizontes de cuatro o cinco años.

En las décadas posteriores se desarrollaron distintas metodologías para determinar el crecimiento futuro de la demanda de fuentes energéticas ante la incertidumbre de su

---

<sup>4</sup> Tesis doctoral "La demanda residencial de energía eléctrica en la Comunidad Autónoma de Andalucía: un análisis cuantitativo"

evolución en el tiempo. A continuación se presentan algunas experiencias en Latinoamérica, donde se aprecia que actualmente para la proyección de variables energéticas se aplica principalmente procedimientos econométricos, metodología ARIMA y de series temporales, como en Perú; aunque hay trabajos novedosos que proponen la utilización de la metodología de redes neuronales, como el caso argentino.

### **Perú**

Según el Plan Referencial de Electricidad 2008-2017<sup>5</sup>, en el numeral 3.2 “Estudio de Proyección de la Demanda del SEIN” menciona que la proyección de la demanda del SEIN<sup>6</sup> se realizó disgregando en dos grandes tipos de carga: Cargas Vegetativas y Cargas Mayores (Especiales/Incorporadas). Las cargas vegetativas corresponden principalmente a la demanda del mercado regulado considerando que este tipo de demanda es susceptible a la modelación mediante técnicas estadísticas, disponiéndose de datos para ello; mientras que las cargas especiales/incorporadas de crecimiento en forma discreta y con importante impacto sobre la demanda global han requerido un análisis individual cuyos resultados se agregan a los resultados del modelado estadístico.

Para la proyección de la demanda de cargas vegetativas se utilizan dos metodologías, una para el corto plazo (período 2009 – 2010), y la otra para el largo plazo en el período restante 2008 – 2017, que en conjunto dan una mayor validez a la proyección. La primera metodología utiliza las series temporales ARIMA<sup>7</sup> para proyectar la variable objetivo en forma mensual, y la otra metodología utiliza los modelos econométricos para el largo plazo en forma anual, como se muestra a continuación gráficamente.

---

<sup>5</sup> Dirección General de Electricidad, Ministerio de Energía y Minas.

<sup>6</sup> SEIN: Sistema Eléctrico Interconectado Nacional

<sup>7</sup> ARIMA: Auto Regressive Integrated Moving Average

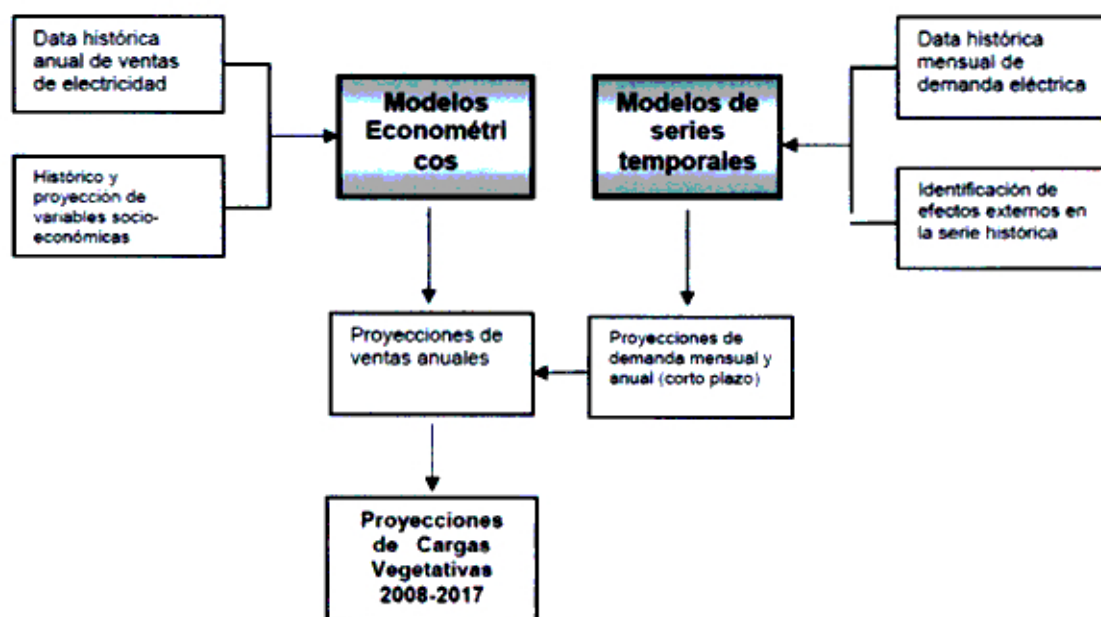


Fig. 1.3 Modelo Proyección Demanda Total

En ese sentido, los resultados indican que la demanda de energía crecerá con una tasa promedio anual de 8.71% para el escenario optimista, 7.55% para el escenario medio y 6.36% para el escenario conservador.

De otro lado, se indica respecto del Sistema Tarapoto – Moyobamba – Bellavista que, la demanda de energía y potencia de este sistema aislado y sus expectativas de crecimiento se consideran dentro de la proyección de las Cargas Vegetativas del SEIN; lo cual evidencia la perspectiva de informe macro del "Plan Referencial de Electricidad" sobre la proyección de energía y potencia a nivel nacional, sin considerar las características particulares de cada zona de concesión como es el caso del presente trabajo de tesis, donde se trata de explicar el comportamiento de la carga vegetativa en una zona de concesión determinada a través de la denominada carga residencial.

### **Argentina**

En la República Argentina actualmente la metodología de "Redes Neuronales" es utilizada para determinar las proyecciones de demanda doméstica, sin embargo, se desarrollan proyecciones de la demanda interna global que son desagregadas en función de la participación de cada región de acuerdo al registro facturado por las empresas distribuidoras de energía, apreciándose que los cálculos de crecimiento para el análisis neuronal se efectuó por medios de interpolación y extrapolación generalmente, y suavizado exponencial respecto de periodos anteriores, lo que no representa una buena base estadística para futuras proyecciones.

Una red neuronal es un conjunto de nodos interpuestos entre la/s variable/s de entrada (variable/s independiente/s) y la/s variable/s de salida (variable/s dependiente/s). Las redes neuronales se utilizan cuando resulta difícil encontrar una relación explícita entre variables. Tienen como ventaja adicional, no utilizada en el caso argentino, de permitir relacionar en conjunto de variables de salida y no sólo con una de ellas.

En ese sentido, la Secretaría de Energía<sup>8</sup> menciona que en ediciones anteriores se utilizaba un modelo econométrico para efectuar las proyecciones de demanda de energía eléctrica, donde se estimaba la evolución de la oferta de generación de energía eléctrica a partir de los valores de PBI y de la población. El problema fundamental radica en el comportamiento de ambas variables, así la evolución de la economía en su conjunto en los últimos años ha presentado dos periodos recesivos muy marcados, esto hace que el PBI presente brechas que no se reflejan en el consumo de energía eléctrica; con lo cual, el consumo de energía eléctrica no se correlaciona con índices agregados de la actividad económica, situación que se vuelve más crítica ante la crisis macroeconómica donde se observan bruscas caídas de actividad; por lo que, de los análisis que se realizaron concluyeron que los resultados obtenidos con una red neuronal son aceptables y pueden incluirse variables relacionadas con el consumo de energía eléctrica que afectan a todos los sectores.

Sobre el particular, en la página web de la Secretaría de Energía<sup>9</sup> se muestran las series históricas de energía eléctrica que maneja la citada entidad, las mismas que presentan sólo periodicidad anual, siendo agrupadas por cada región y distrito federal que forman parte de la República Argentina, teniendo las siguientes:

- Serie de cantidad de Usuarios total país 1991-2009
- Potencia instalada 1976-2009
- Serie Facturación en MWH 1970-2009
- Serie Generación asociada a redes 1970-2009
- Potencia Máxima del Sistema
- Consumo de Combustibles en Centrales de Generación de Energía Eléctrica 1970-2009

Como referencia en la Región de Jujuy, la Empresa Jujueña de Energía SA (EJESA)<sup>10</sup> en el marco de la Revisión quinquenal de tarifas para el periodo 2011-2016 para su zona de concesión establece que para el pronóstico de la demanda se determinaron tasas de crecimiento para sectores de consumo homogéneos, trabajándose con las series históricas de las ventas de energía de la distribuidora. Se aprecia que las series de datos de consumo

<sup>8</sup> Publicación "Prospectiva 2002"

<sup>9</sup> Página web <http://energia3.mecon.gov.ar/contenidos/verpagina.php?idpagina=3140>

<sup>10</sup> Informe "Pronóstico de Demanda, Revisión quinquenal de tarifas de EJESA, Periodo 2011-2016"

de energía utilizados para el pronóstico presentan sólo periodicidad anual, no así mensual. Para el caso particular de la proyección del sector residencial, este se establece a través de metodología econométrica como el producto de los consumos por usuario previstos para cada año por la cantidad de clientes sectoriales correspondientes, para lo cual, tanto el pronóstico del número de clientes como la estimación de los consumos se estiman en porcentajes que resultan de la aplicación de extrapolación por tendencia.

Del estudio en mención, se determinan las siguientes proyecciones:

**Tabla 1.1 Tasas de crecimiento – Región Jujuy**

Tipo de carga	2009-2021		2006-2011	
	Energía	Clientes	Energía	Clientes
Total	4.5%	2.3%	4.5%	2.4%
Residencial	3.7%	2.3%	3.7%	2.5%

Elaboración propia.

### **Colombia**

Para el caso colombiano, el procedimiento para realizar la proyección de demanda de energía eléctrica en su sistema eléctrico nacional integra la aplicación combinada de metodologías de series de tiempo y econométricas. Para la estimación a corto plazo se aplican las primeras que permiten tener en cuenta tendencias y coyunturas recientes del consumo y los efectos calendario. Para el largo plazo se aplican modelos econométricos que consideran la relación histórica del consumo de energía eléctrica con variables macroeconómicas y de población. A partir de la estimación del comportamiento futuro de éstas últimas realizado por las agencias del gobierno responsables se establecen escenarios de evolución del consumo y la demanda eléctrica. La Unidad de Planeamiento Minero Energético (UPME) publica periódicamente en su página web<sup>11</sup> las proyecciones de demanda de energía eléctrica y potencia para el sistema interconectado nacional.

El Ministerio de Minas y Energía de Colombia<sup>12</sup>, establece como máximo nivel de desagregación la proyección de la demanda regulada, no haciendo distinción entre los clientes en baja tensión respecto de los de media tensión, en donde se establece el uso de metodologías semejantes a los utilizados para la proyección de la demanda total nacional.

En ese sentido, se utilizan modelos econométricos para analizar el comportamiento anual de las series de ventas totales de energía, ventas sectoriales y demanda de energía con relación a diferentes variables como Producto Interno Bruto – PIB, valores agregados sectoriales nacionales, valor agregado total de la economía, consumo final de la economía,

<sup>11</sup> Página web <http://www.siel.gov.co/Inicio/Demanda/ProyeccionesdeDemanda/tabid/97/Default.aspx>

<sup>12</sup> Publicación "Proyección de Demanda de Energía Eléctrica y Potencia Máxima"

índices de precios, población, etc. A estos, se agrega posteriormente las pérdidas de energía a nivel de distribución, sub transmisión y transmisión, considerándose además las demandas de energía de cargas industriales especiales.

Para el caso de la proyección de la demanda doméstica de energía, se parte de la demanda de energía mensualizada a la cual se aplican los factores de carga correspondientes a cada mes teniendo en cuenta lo ocurrido en los tres últimos años, dando como resultado los valores de potencia máxima mensual doméstica, de los cuales se generan los valores de potencia máxima anual.

A continuación se muestra la metodología utilizada para la demanda total nacional<sup>13</sup>, donde se advierte que el método utilizado es básicamente econométrico:

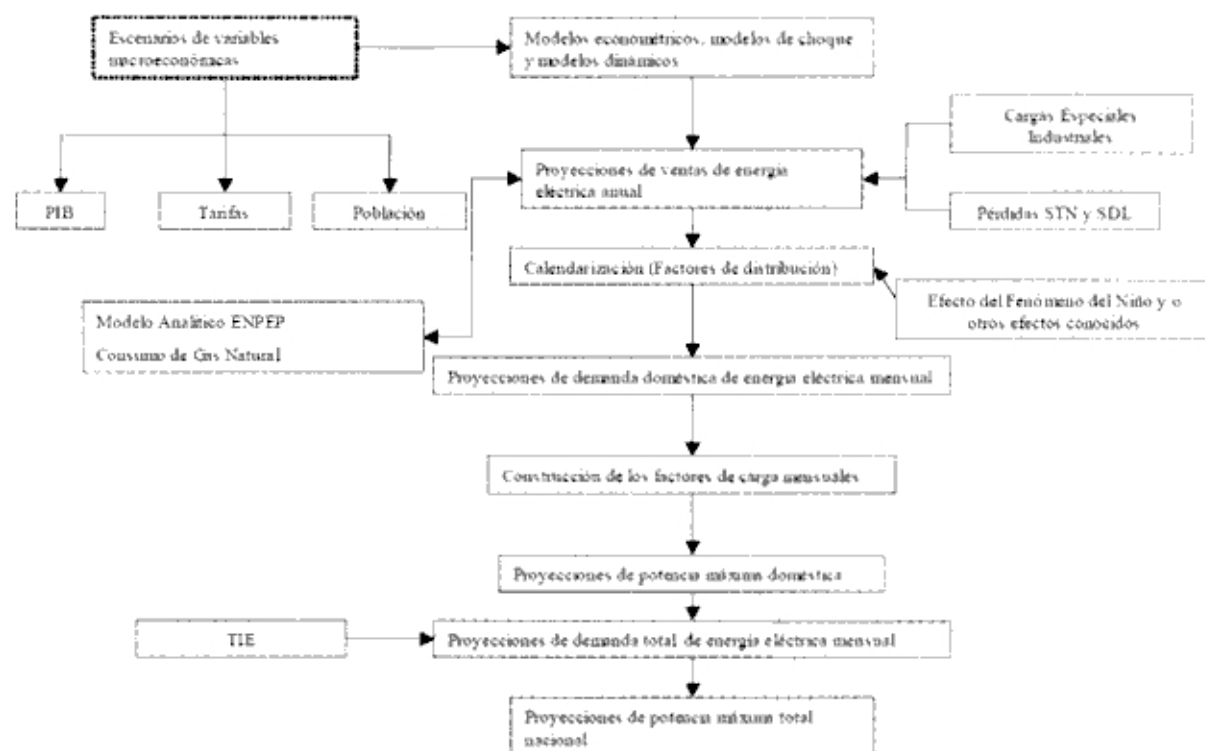


Fig. 1.1 Metodología Proyección Demanda Total Nacional

Es de precisar que, Murillo, Trejos, Carbajal<sup>14</sup> presentan un estudio que describe un pronóstico de demanda de energía eléctrica, utilizando la metodología ARIMA, y el paquete estadístico SPSS, donde los datos históricos han sido suministrados por la Empresa de Energía de Pereira.

<sup>13</sup> Publicación "Proyecciones de Demanda de Energía Eléctrica y Potencia Máxima, 2003-2011"

<sup>14</sup> Estudio del pronóstico de la demanda de Energía Eléctrica, utilizando modelos de series de tiempo



Se concluye que, es posible modelar los consumos de energía eléctrica en los municipios de Colombia utilizando la metodología propuesta por Box y Jenkins, dado que el comportamiento de estas series no es completamente aleatorio y se pueden describir como series de tiempo con una alta probabilidad de éxito en la modelación a través de dicha metodología; estableciendo que el modelo ARIMA que mejor ajusta describe el consumos de energía eléctrica a las 8 am es ARIMA (1,0,0)(2,1,0).

### Chile

La Comisión Nacional de Energía de Chile presenta un resumen<sup>15</sup> realizado sobre la base del estudio contratado durante el año 2008 al Programa de Gestión y Economía Ambiental del Departamento de Ingeniería Industrial de la Universidad de Chile, denominado "Diseño de un Modelo de Proyección de Demanda Energética Global Nacional de Largo Plazo", cuyos contenidos y conclusiones del estudio han servido de insumo para su trabajo.

En el citado resumen, se establece que el modelo de proyección seleccionado será de carácter híbrido, basado en un enfoque sectorial que combina un análisis econométrico cuando las tendencias parecen robustas para el sector, y opinión experta y análisis de uso final cuando se esperan cambios en estas tendencias, como se muestra a continuación.

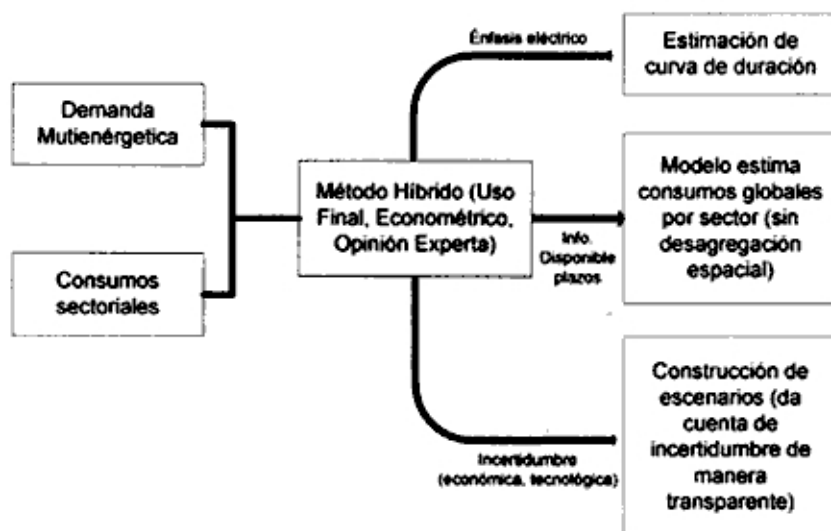


Fig. 1.2 Modelo Proyección Demanda Total

En ese sentido, para realizar las estimaciones y proyecciones de los distintos consumos de energía, se emplea una metodología multisectorial y subsectorial que se fundamenta en funciones de producción para cada sector, lo cual entrega una medida de tendencia de los consumos de largo plazo, bajo la siguiente fórmula.

<sup>15</sup> Publicación "Modelo de Proyección Demanda Energética Nacional de Largo Plazo"

$$Y_t = \prod_i C \cdot P_{it}^{a_i} \dots\dots\dots(1.1)$$

Donde:

$Y_t$ , consumo de energía del sector o subsector

C, constante

$P_{it}$ , producción tipo i del sector o subsector en el tiempo

t y  $a_i$ , elasticidad del consumo energético respecto de la producción tipo i en el sector o subsector tratado.

De otro lado, el máximo nivel de desagregación es el denominado "Sector Comercial, Público y Residencial", sobre el cual se menciona que los datos históricos muestran que el consumo total se ha incrementado un 3,9% promedio anual entre 1982-2006, mientras que para el período 2006-2030 se estima que el consumo crecerá con una tasa promedio anual de 3,8%.

**México**

En el gobierno mexicano, la Secretaría de Energía (SENER)<sup>16</sup> señala que el estudio y proyección del mercado eléctrico requiere del análisis de la información sobre el consumo de electricidad de los diversos tipos de usuarios, así como de la evolución esperada de la economía, entre otras variables. Señalando además que mediante diversas metodologías econométricas se especifican las variables económicas y sociales relevantes para definir el nivel, estructura del consumo y expectativas de desarrollo del mercado eléctrico, estableciendo supuestos macroeconómicos, poblacionales, precios de combustibles, precios promedio de la electricidad, así como los relacionados con el uso más eficiente de la energía eléctrica por la aplicación del horario de verano y otros programas específicos.

Para el caso de la demanda de energía residencial para el periodo 2009-2024, aplicando metodología econométrica, estiman que crecerá 3.9% durante el periodo, el comercial 2.7% y el de servicios 2.1%, habiendo utilizado los siguientes modelos:

Modelo de saturación de usuarios:

$$RSAT_t = f (RSAT_{t-1}, RSAT_{t-2}, T, VISAT) \dots\dots\dots(1.2)$$

Donde:

- UR = Número de usuarios del sector residencial
- V = Número de viviendas

---

<sup>16</sup> Publicación "Prospectiva del Sector Eléctrico 2009-2024"

- SAT = Coeficiente de saturación de usuarios = UR/V
- RSAT = (1 - SAT)/SAT
- T = Tiempo.
- VISAT = 1 en 1996, 0 en el resto

Modelo de ventas por usuario:

$$VU_t = f(VU_{t-1}, SAT_t, CP/V_t, P_t, PRED_t, T) \quad \dots\dots\dots(1.3)$$

Donde:

- VU = Ventas por usuario del sector residencial
- SAT = Coeficiente de saturación de usuarios = UR/V
- CP/V = Consumo privado por vivienda
- P = Precio medio real de electricidad en el sector residencial
- PRED = Precio relativo de los electrodomésticos
- T = Tiempo

**Conclusión:**

De lo expuesto, se puede afirmar que a la fecha no se conocen trabajos que aborden formalmente el tema de la estimación con metodología ARIMA de la demanda de energía eléctrica en la zona de concesión de Lima Sur y más concretamente la del sector residencial. De hecho, únicamente se tiene experiencia en el país sobre la utilización de esta metodología a nivel nacional, no haciéndose la desagregación por zonas de concesión que es importante para que cada empresa eléctrica privada pueda proyectar futuras inversiones en su sistema eléctrico con características muy particulares que deben ser analizadas como base para la toma de decisión. En consecuencia, puede considerarse a esta investigación como la primera en Perú que se propone tratar formalmente el modelamiento de la demanda de energía residencial.

## **CAPÍTULO II**

### **MODELOS PARA LA PREDICCIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA**

El objetivo de todo sistema eléctrico es dar cobertura a la demanda de electricidad de forma eficiente. En este capítulo se estudian modelos de previsión para el horizonte temporal a corto plazo. El ámbito del estudio es el consumo residencial de energía en la zona sur de Lima Metropolitana utilizando datos históricos mensuales desde 1996 a 2009. La demanda de electricidad ha sido objeto de diversos estudios, sin embargo, para el caso particular de Lima Metropolitana los estudios han estado reservados básicamente a organismos públicos, así como a las propias compañías eléctricas pero a nivel macro, y que nunca han llegado al nivel desagregado de energía residencial que se plantea en el presente trabajo.

Este capítulo se encuentra organizado como sigue: En la primera sección se comenta la importancia de la predicción de la demanda de energía eléctrica. En la segunda sección se contemplan algunos de los principales factores que influyen en la demanda de dicha energía en los sectores industrial, comercial y residencial o doméstico. En la tercera sección, se ocupa de los principales métodos que se suelen utilizar en el pronóstico de la demanda de electricidad, comentándose las principales ventajas e inconvenientes de cada uno de ellos. En la cuarta sección se realiza una introducción a las series temporales. En la quinta sección, se efectúa un estudio detallado y comparativo de los modelos univariantes, es decir, de aquéllos que tratan de explicar la variación de la demanda a través de sí misma. Finalmente, en la sexta sección se desarrolla teóricamente la formulación de la metodología estocástica ARIMA.

#### **2.1 IMPORTANCIA DE LA PREDICCIÓN DE LA DEMANDA DE ENERGÍA ELÉCTRICA**

La importancia de la predicción de la demanda de energía eléctrica radica en que ayuda a determinar posible una carencia de capacidad de suministro de energía al usuario final y, en consecuencia, pudiera ser conveniente considerar la construcción de nuevas plantas de generación, subestaciones y de líneas de transmisión o por el contrario aumentar la

potencia de los transformadores de potencia instalados en las citadas subestaciones o, en el futuro existirá un exceso de capacidad que pudiera aconsejar la no utilización de parte del parque ya existente.

En ese sentido, el sector energético es intensivo en capital con inversiones a largo plazo, puesto que se necesita al menos una década para planificar y construir una nueva planta generadora, y al menos de tres años para la construcción de una subestación de transformación y de la línea de transmisión, por lo que, una previsión correcta de la demanda de energía eléctrica es un requisito imprescindible para lograr las metas previstas de calidad y fiabilidad del servicio, ya que la creciente dependencia a la electricidad aumenta los inconvenientes causados a los consumidores si se producen deficiencias en el suministro del servicio. Es de precisar que, la electricidad es uno de los sectores que mantendrá en el futuro una de las tasas de crecimiento más altas.

En consecuencia, sin una adecuada representación de las necesidades futuras de consumo eléctrico, los problemas de exceso de capacidad, o por el contrario de capacidad insuficiente, pueden tener costos altos. La correcta previsión de la demanda también desempeña un importante papel en las decisiones de una compañía eléctrica respecto a qué cantidad será conveniente comprar a las generadoras del sector.

Un argumento importante a tener en cuenta es el económico, ya que si las predicciones resultan ser demasiado bajas pueden tener lugar carencias de energía cuyos costos habitualmente son mucho mayores que el valor de la energía no suministrada. Por el contrario si las previsiones resultan demasiado altas, los costos de oportunidad pueden ser muy elevados al tener comprometidos, de forma improductiva, cuantiosos fondos económicos durante largos períodos de tiempo. Ambas desviaciones, por exceso o por defecto, acaban repercutiendo sobre el usuario final, es decir, si se produce una carencia de electricidad, el precio de ésta se incrementará y el abonado pagará más por la energía consumida; si por el contrario las erróneas predicciones se traducen en un exceso de capacidad de generación eléctrica, los costos asociados con la capacidad ociosa, serán trasladados también al consumidor, salvo que se considere como reserva de capacidad remunerable.

Existen al menos tres motivos para modelar la demanda de energía. En primer lugar, el suministro razonablemente fiable de energía es vital para el funcionamiento de cualquier economía moderna. En segundo lugar, la ampliación de los sistemas de suministro de

energía requiere muchos años. En tercer lugar, las inversiones necesarias en tales sistemas son altamente intensivas en capital.

Finalmente, los programas de planificación del sector eléctrico se basan fundamentalmente, en las predicciones que se realicen sobre las puntas de potencia y energía demandada y, de acuerdo con ellas, se pueden adoptar las eventuales decisiones de incrementar la capacidad eléctrica instalada. Por lo tanto, cuanto más acertadas sean las predicciones, menores serán los riesgos de incurrir en inversiones innecesarias y/o de causar insatisfacción a los usuarios.

## **2.2 FACTORES QUE INFLUYEN EN LA DEMANDA DE ENERGÍA ELÉCTRICA**

Frecuentemente, en la demanda de energía eléctrica conviene segmentar los kilovatios-hora consumidos entre aquellos usuarios que se observa un comportamiento similar. En ese sentido, está muy difundida la división en tres grupos de demanda: residencial, comercial e industrial aunque, a veces, resulta útil segmentar aún más las clases de clientes servidos a fin de realizar una predicción más detallada de las tendencias de los distintos tipos de usuarios.

El precio relativo de la electricidad es un factor importante en la cantidad de energía consumida por los clientes. La influencia de otros factores puede ya no ser tan clara, por ejemplo, el precio de los productos sustitutos (petróleo, gas natural, etc.). Es de precisar que, la magnitud del ingreso familiar es un factor importante en el consumo del sector residencial que en el industrial.

Otros factores de considerable importancia en la demanda de energía eléctrica también pueden depender del sector considerado, así en el sector residencial se pueden considerar factores relevantes el número de consumidores, el precio de los electrodomésticos y el número de personas por vivienda, etc. También varía la demanda doméstica o residencial de electricidad según la hora del día, ya que se suele consumir más electricidad durante las primeras horas del día antes de producirse la incorporación a la jornada laboral declinando posteriormente el consumo eléctrico doméstico durante el desarrollo de la misma y volviendo a aumentar dicho consumo cuando tiene lugar el regreso al hogar luego de finalizar la jornada laboral. El sector residencial principalmente contribuye a las puntas de demanda del sistema de generación y transmisión eléctrica.

En el sector comercial son factores relevantes la magnitud del espacio ocupado por las oficinas, el nivel de empleo en el sector servicios, el número de potenciales clientes en la proximidad del establecimiento comercial, el tamaño de la población escolar, los niveles de consumo de los ciudadanos, etc. En el sector industrial se suelen considerar factores influyentes en la demanda de energía eléctrica, el nivel de producción industrial, el nivel de empleo en este sector, la productividad de los distintos establecimientos industriales, la normativa medio ambiental, etc.

La demanda de energía eléctrica está sujeta a alteraciones debido a un gran número de factores. El consumo de electricidad varía de acuerdo con la estación del año y la hora del día. Con la estación del año se consume más electricidad durante el verano que en la primavera, el invierno o el otoño, estaciones estas últimas en las que el clima es más moderado. En ese sentido, la época del año en la que los diversos alimentadores que comprenden la zona de concesión de Lima sur alcanzan el máximo de potencia demandada o la máxima energía consumida corresponde a los meses de verano, especialmente entre febrero y marzo.

### **2.3 MODELAMIENTO DE LA PREDICCIÓN DE ENERGÍA ELÉCTRICA**

Un sistema de predicción de la demanda eléctrica debe ser principalmente óptimo, robusto y calculable. En este sentido, el horizonte temporal de la predicción está relacionado con los problemas a los que se enfrenta una empresa eléctrica, básicamente se pueden considerar tres horizontes temporales teóricos (corto, medio y largo plazo) para la predicción de la demanda de energía.

Los métodos de previsión de la demanda de energía eléctrica han evolucionado considerablemente desde la década de 1970 hasta la actualidad. En general el horizonte temporal ha sido uno de los factores primordiales que han tenido en cuenta las metodologías utilizadas, pudiendo citarse las siguientes: de tendencia temporal, de series cronológicas o temporales y econométricas.

- **Métodos de tendencia temporal**

Estos métodos eran casi los únicos utilizados antes de la década de 1970, debido a las tendencias estables de la demanda en dicha época, pero actualmente su utilización prácticamente ha quedado reducida al muy corto plazo. En esta técnica se obtiene el

ajuste de la tendencia general histórica en kWh facturados o en kW de las puntas de potencia.

Algunas de las técnicas utilizadas en estos métodos recurren a una transformación previa de los datos de carga a fin que el ajuste de la tendencia explique con mayor facilidad la tendencia del consumo. Otras técnicas de esta metodología suponen una tasa de crecimiento constante de la demanda, mientras que otras se basan en una curva de saturación que tiene un período inicial de crecimiento lento, seguido de un período intermedio de crecimiento rápido, al que sigue un período final en el que declina la tasa de crecimiento alcanzando la energía demandada su nivel de saturación.

Entre las principales ventajas de estos métodos se pueden citar su economía y rapidez, el no necesitar un gran apoyo técnico y su mínima exigencia en cuanto a los datos requeridos. En cuanto a sus desventajas puede decirse que estos métodos no son sensibles a los cambios estructurales y que no disponen de un trazado histórico para interpretar las causas de los errores pasados de predicción a fin de corregir el modelo.

- **Métodos de series cronológicas**

Los métodos de series cronológicas efectúan la predicción de la demanda de energía eléctrica analizando la trayectoria de los datos históricos y proyectándola en el futuro. En particular los métodos de Box-Jenkins suelen resultar adecuados cuando existe una fuerte autocorrelación entre los datos de la serie analizada. El modelo inicial propuesto por Box y Jenkins se basa en el estudio sistemático de los procesos ARIMA<sup>17</sup> en tres fases (identificación, estimación y diagnóstico).

Si el valor de la demanda actual puede expresarse sólo como combinación lineal de un número determinado de datos pasados, estamos ante un proceso autorregresivo (AR). En cambio, si expresamos la demanda presente sólo como combinación lineal de los errores pasados de predicción, el proceso se denomina de medias móviles (MA). Cuando para expresar la demanda actual necesitamos las dos combinaciones lineales mencionadas, el proceso se conoce como ARMA y en todos los casos anteriores es necesario identificar el orden de los polinomios de retardo, estimar los parámetros del modelo identificado y proceder a la validación de este último.

Cuando los datos de la serie evidencian un comportamiento no-estacionario es necesario diferenciar los datos hasta que se logre la estacionariedad del proceso,

---

<sup>17</sup> Auto Regresivo Integrado de Medias Móviles



debiendo identificar también el orden de las diferencias aplicadas, ello da lugar a los denominados modelos ARIMA. Cuando los datos indican un comportamiento estacional de la serie, el método para construir el modelo es similar al que se describe antes.

Entre las principales ventajas del análisis de series temporales se encuentran su bajo costo, su moderado error en el corto plazo y que sólo utiliza datos históricos de la serie que se trata de predecir, pudiendo citarse entre sus inconvenientes que, al suponer permanente la caracterización de la trayectoria de los datos históricos, se está postulando que las condiciones socioeconómicas del pasado serán invariables en el futuro.

- **Métodos econométricos**

Los métodos econométricos se basan en ecuaciones que relacionan la demanda de electricidad con factores externos. Generalmente se considera que algunos de los modelos de este tipo pueden ser los más complejos de todos los utilizados para la predicción de la demanda de la energía eléctrica.

En estos modelos la demanda futura se determina considerando la influencia que en esta última tienen algunas variables explicativas, cuyos valores históricos son un requisito imprescindible para la utilización de estos modelos. Entre las variables explicativas para estos modelos se encuentran el número de clientes, la renta de los mismos, el precio de la electricidad, el nivel de actividad económica, las variables climáticas, el nivel de equipamiento de los hogares, etc.

Los modelos econométricos son herramientas predictivas muy potentes, sin embargo, algunas de sus limitaciones es que se pronostica el futuro basándose en relaciones entre las variables que tuvieron lugar en el pasado. Otra limitación radica en la consistencia de la relación entre las variables relevantes, ya que dicha relación puede ser causal o simplemente casual. Tampoco se puede despreciar la influencia en el consumo de energía eléctrica de variables de nueva aparición o emergentes que en el pasado, o no existían, o no habían sido suficientemente identificadas quedando englobadas en la perturbación aleatoria.

Con frecuencia este método se aplica en modalidad uniecuacional, y en particular en la forma de un modelo de regresión lineal. El análisis de regresión proporciona una medida de la significación de los coeficientes de las variables y de la fiabilidad de la predicción. Los límites de confianza de los valores de la predicción proporcionan una indicación del

riesgo asociado a la misma. El análisis de regresión consiste en formular y contrastar un modelo que estudie el impacto de las variables explicativas (independientes) sobre la variable explicada (dependiente).

Algunos investigadores que se han ocupado de la demanda de energía eléctrica en el sector residencial, incluyen variables demográficas en los modelos que proponen. El modelo de Crow, Robinson y Squiteri<sup>18</sup>, se incluyen como variables explicativas además de determinadas variables económicas, el número medio de miembros de la unidad familiar y el número de habitaciones del hogar, como se muestra a continuación:

$$KWH = a_1 - a_2PELEC + a_3INC + a_4CONS + a_5R + a_6S \dots\dots\dots(2.1)$$

Donde:

- KWH: la energía consumida.
- PELEC: el precio de la electricidad.
- INC: el nivel de ingresos.
- CONS: el número de consumidores.
- R: el número de habitaciones por hogar.
- S: el tamaño familiar medio (número de personas).

Entre las ventajas de los modelos econométricos se encuentran que pueden medir el efecto de las causas subyacentes de la tendencia y de la incertidumbre de la predicción y que los modelos pueden ser re estimados. Entre sus desventajas pueden citarse que requieren considerable destreza en la utilización de los métodos econométricos y que su costo puede ser relativamente elevado.

## 2.4 INTRODUCCIÓN A LAS SERIES TEMPORALES

Un proceso estocástico es una familia de variables aleatorias que corresponden a momentos sucesivos del tiempo. En cada período o momento temporal se dispone de una variable que tendrá su correspondiente distribución de probabilidad, como se muestra a continuación.

---

<sup>18</sup> Tesis doctoral "La demanda residencial de energía eléctrica en la Comunidad Autónoma de Andalucía: un análisis cuantitativo"

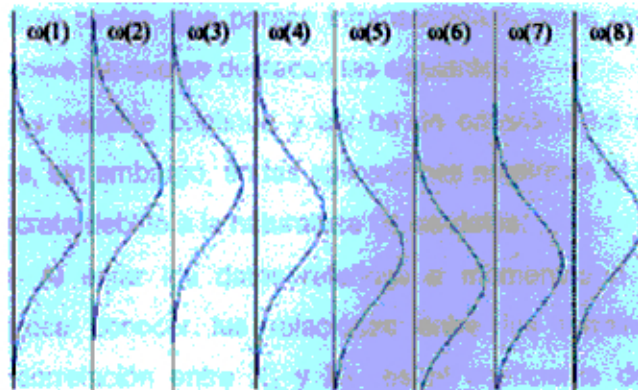


Fig. 2.1 Ejemplo de procesos estocásticos

Un proceso estocástico no tiene ni principio ni fin, esto es, el tiempo transcurre desde menos infinito hasta más infinito, por lo que, un proceso estocástico será un conjunto de infinitas variables aleatorias.

$$t \in (-\infty, +\infty) \quad \dots\dots\dots(2.2)$$

$$\omega(-\infty), \dots, \omega(-1), \omega(0), \omega(1), \omega(2), \omega(3), \dots, \omega(T), \omega(T+1), \dots, \omega(+\infty) \quad \dots\dots\dots(2.3)$$

Se dice que un proceso estocástico es estacionario si su media es constante e independiente del tiempo y la covarianza entre dos variables aleatorias del proceso no depende del tiempo cronológico en que se midan las mismas, sino sólo del número de periodos de separación entre dichas variables aleatorias, en consecuencia la varianza del proceso es finita y constante.

Se considera que una serie cronológica o temporal es parte de la realización de un proceso estocástico o que el mecanismo de generación de una serie es un proceso estocástico. Se suele decir que una serie temporal es una sucesión de valores generalmente registrados a intervalos equidistantes del tiempo.

#### 2.4.1 SERIES TEMPORALES<sup>19</sup>

**Definición:** Una serie temporal, llamada también serie histórica o serie cronológica, es una sucesión de valores observados de una variable referida a momentos o a periodos de tiempo diferentes, generalmente regulares. Por lo tanto, se concluye que es una secuencia ordenada de observaciones sobre una variable en particular.

**Características:** La característica distintiva de una serie temporal, en contraposición a las observaciones de corte transversal, es que los datos aparecen ordenados

<sup>19</sup> "Trabajo sobre modelamiento y predicción de la Producción de Electricidad (GWH), utilizando la metodología estocástica ARIMA"

cronológicamente. Este hecho, que parece intrascendente tiene diversas repercusiones teóricas y prácticas, entre las que se destacan las siguientes:

- El tiempo es una variable continua y así ha de considerarse en la formulación de modelos teóricos, sin embargo, en las aplicaciones empíricas el tiempo se trata como una variable discreta debido a la naturaleza de los datos.
- Autocorrelación. Al estar los datos referidos a momentos o períodos de tiempo sucesivos, interesa conocer las relaciones entre los términos consecutivos. El coeficiente de correlación entre  $X_1$  y  $X_2$  es el coeficiente de autocorrelación de segundo orden, y así sucesivamente.
- Para medir la amplitud, longitud o tamaño de la serie temporal se puede utilizar el tiempo transcurrido durante el período muestral o el número de observaciones, ambas variables están relacionadas por la unidad de tiempo utilizada en las observaciones.

**Finalidad de las Series Temporales:** En términos generales, el análisis de series temporales persigue la descripción de una serie temporal dada, y la predicción de su evolución futura. Para ello es posible adoptar los enfoques alternativos univariante y multivariante. Para nuestro análisis usaremos el enfoque univariante que hace uso de los datos históricos de la variable (serie temporal) para elaborar un modelo que describa aceptablemente el comportamiento de la misma en el pasado. Tal modelo descriptivo puede usarse para realizar predicciones de los valores futuros de la variable. Este tipo de predicción también se denomina auto proyectivo.

**Componentes de una serie de tiempo:** La descomposición clásica es un método que se basa en el supuesto de que los valores observados de una serie son el resultado de agregar todos o alguno de los componentes siguientes:

- **Tendencia:** Una serie de tiempo con tendencia es aquella que contiene un componente de largo plazo que representa el crecimiento o declinación de la serie a través de un período amplio.
- **Estacionalidad:** Se define como estacional una serie de tiempo con un patrón de cambio a si mismo año tras año. Por lo regular, el desarrollo de una técnica de pronóstico estacional comprende la selección de un método multiplicativo o uno de adición y estimar después índices estacionales a partir de la historia de la serie.
- **Ciclo:** El efecto cíclico se define como la fluctuación en forma de onda alrededor de la tendencia. Los patrones cíclicos tienden a repetirse en los datos cada dos, tres o más años. Es difícil establecer un modelo para estos patrones cíclicos porque no son estables.

- **Irregular:** El componente irregular de la serie de tiempo es el factor residual, es decir, "todo lo que sobra" y toma en consideración las desviaciones de los valores reales de la serie de tiempo en comparación con los esperados, es el elemento aleatorio.

La tendencia y las perturbaciones aleatorias están presentes en toda serie. La ciclicidad y la estacionalidad pueden estar presentes o ausentes. La identificación de un modelo consiste en determinar cuáles de los componentes anteriores deben estar incluidos en el mismo de forma que se obtenga una representación satisfactoria de los datos.

Estos componentes no pueden ser observados directamente por separado, pero pueden estimarse con diferentes técnicas de descomposición. El método de descomposición que se deberá aplicar en cada caso dependerá del esquema de agregación de los componentes que se suponga presente la serie analizada. Los esquemas más frecuentes son el aditivo que postula que la suma de los componentes constituye la serie observada; y el multiplicativo que supone que la serie observada es el resultado de multiplicar los componentes. La razón de utilizar una de estas dos hipótesis, en lugar de otras más complejas, radica fundamentalmente en la sencillez y operatividad de las mismas.

El esquema aditivo postula una relación del tipo:

$$Y_t = T_t + C_t + E_t + I_t \quad \dots\dots\dots(2.4)$$

Donde:

$T_t$ : valor de la tendencia.

$E_t$ : valor del componente estacional.

$C_t$ : valor del factor cíclico.

$I_t$ : valor del componente irregular.

Es decir, el valor de la serie temporal para cualquier instante es igual a la suma de los valores correspondientes a sus cuatro componentes.

El esquema multiplicativo se expresa mediante la relación:

$$Y_t = T_t \times C_t \times E_t \times I_t \quad \dots\dots\dots(2.5)$$

Lo que significa que el valor de la serie temporal en cualquier instante es igual al producto de los valores correspondientes a sus cuatro componentes.

Si las observaciones de la variable se efectúan en sub períodos (meses, trimestres, días, horas) se puede determinar el esquema de agregación de las componentes a partir de la

representación gráfica. Si las fluctuaciones de los valores de la variable por encima y por debajo de la tendencia presentan un patrón regular en el transcurso del tiempo, el modelo de agregación será el aditivo. Si por el contrario la amplitud de dichas fluctuaciones crece o decrece con la tendencia el esquema o modelo es el multiplicativo.

## **2.4.2 ANÁLISIS DE LA PREDICCIÓN CON SERIES TEMPORALES**

Una serie de tiempo puede combinar los patrones de tendencia, estacionalidad y ciclaje. Sin embargo, alguno de estos patrones puede dominar la serie, existen técnicas que permiten identificar el elemento dominante, como lo es el método de descomposición. A estos patrones hay que añadir un elemento indeseable en toda serie de tiempo, que siempre existe y que es la aleatoriedad de la observación.

En consecuencia, el análisis cuantitativo de predicción de series temporales energéticas se puede realizar mediante análisis determinístico (medias móviles, alisados exponenciales, ajustes de tendencias) o análisis estocástico (ARIMA). En principio, ambas técnicas pueden ser aplicables, la elección de una u otra depende de las características particulares de cada serie (tendencia y estacionalidad), del tamaño muestral disponible de la serie, del horizonte temporal de predicción, de la complejidad del método y del grado de precisión con que se desea llevar a cabo dicho análisis. Bajo el criterio estrictamente matemático, interesan los métodos que comporten los errores de predicción más bajos.

Una serie temporal energética se caracteriza por poseer tanto tendencia como estacionalidad. Desde el punto de vista de la información estadística, una variable energética manifiesta sólo tendencia cuando la serie histórica se contabiliza anualmente, mientras que, muestra tendencia y estacionalidad cuando dicha serie es contabilizada mensualmente. Para determinar el tipo de metodología predictiva a utilizar, fundamentalmente, ha de tenerse en cuenta el tamaño de la serie histórica.

**Selección del Modelo de Series Temporales:** La elección del modelo de series temporales que se utiliza en el cálculo de las proyecciones de una determinada variable histórica, se realiza en base a lo siguiente:

- El tamaño muestral de la serie histórica y su nivel de desagregación: datos anuales o mensuales.
- La tendencia evolutiva de la serie.
- La ciclicidad de la serie.
- La estacionalidad de la serie.
- La irregularidad de la serie.

- El rigor metodológico.
- Calidad de los resultados.

El modelo determinístico de alisados exponenciales Holt se utiliza cuando se trata de series históricas que poseen solamente un patrón de tendencia, es decir, que carecen de ciclicidad y estacionalidad. El modelo determinístico de series temporales Winters se utiliza cuando la serie histórica posee tendencia y estacionalidad. Para nuestro caso se utilizará como elemento comparativo metodológico de pronóstico frente al modelo desarrollado.

El modelo estocástico ARIMA es de mayor complejidad metodológica y se utiliza en el cálculo de proyecciones de series que poseen patrones de tendencia, estacionalidad e irregularidad. El refinamiento predictivo es alto y se puede utilizar para proyectar series temporales mensuales típicas de la actividad energética, como es el caso de la energía eléctrica.

Cada una de las series históricas o temporales originadas por las diversas actividades energéticas tiene un comportamiento particular, dado que poseen un determinado tipo de tendencia y componente estacional. Para alcanzar buenos resultados, las técnicas de predicción deben estar en estrecha relación con las características propias de cada variable temporal. Así, se evaluará las predicciones de variables energéticas mediante el análisis cuantitativo univariante de series temporales, utilizando el enfoque estocástico bajo el que se hace la formulación de la modelización ARIMA y el enfoque de alisado exponencial de Winter.

## **2.5 ESTUDIO COMPARATIVO DE MODELOS UNIVARIANTES.**

El objetivo del análisis de series temporales es estudiar el comportamiento de una o varias variables a lo largo del tiempo, o investigar la estructura de la relación entre las variables de un proceso estocástico.

La sistematización de las diferentes metodologías empleadas en el tratamiento de una serie temporal puede hacerse en base a criterios diversos. El enfoque que se trata a continuación es el enfoque determinista y univariante el cual trata de explicar la trayectoria de una variable a través de la información contenida en los datos históricos de su correspondiente serie, es decir, intenta capturar el comportamiento sistemático que muestra el pasado de la misma y en base a ello realizar predicciones respecto al futuro.

Una vez definidos los posibles esquemas que pueden seguir las series, se estudian los distintos modelos univariantes estudiados para la demanda de electricidad, en función de su grado de complejidad.

El objetivo de las técnicas de predicción no causal es obtener estimaciones o pronósticos de valores futuros de una serie temporal a partir de la información histórica contenida en la serie observada hasta el momento actual. Estas técnicas no requieren la especificación de los factores que determinan el comportamiento de la variable, sino que se basan únicamente en la modelización del comportamiento sistemático de la serie. Se consideran tres modelos posibles del comportamiento sistemático de una serie temporal: modelo estacionario (sin tendencia), modelo con tendencia lineal y modelo con estacionalidad. La técnica de predicción adecuada dependerá del modelo de comportamiento de la serie.

## 2.5.1 MÉTODOS APLICABLES A SERIES SIN TENDENCIA NI ESTACIONALIDAD

### Modelos Ingenuos

El modelo más simple de serie temporal es el puramente aleatorio,  $Y_t = \varepsilon_t$ , habitualmente denominado ruido blanco. Si una serie tiene estructura de ruido blanco, su gráfica sigue una trayectoria que no obedece a ningún patrón. Una serie ruido blanco es impredecible, y por tanto, no admite ningún tipo de pronóstico ya que no contiene ninguna información sobre relaciones temporales. Formalmente, si tal serie se designa por  $\varepsilon_t$ , decimos que  $\varepsilon_t$  es un ruido blanco si cumple las condiciones siguientes:

$$E[\varepsilon_t] = 0, \forall t, \dots\dots\dots(2.6)$$

$$E[\varepsilon_t^2] = \sigma^2, \forall t; \dots\dots\dots(2.7)$$

$$E[\varepsilon_t \varepsilon_{t'}] = 0, t \neq t' \dots\dots\dots(2.8)$$

Se dice que, para una determinada serie un modelo es adecuado cuando éste ha extraído la máxima información sobre la serie y todo lo que queda los residuos es aproximadamente ruido blanco. Si los residuos aún contuvieran alguna información útil, ésta podría utilizarse para mejorar el modelo hasta entonces contemplado.

Otro tipo de modelo ingenuo es el denominado paseo aleatorio, el cual es un modelo de serie temporal en el que la diferencia entre dos observaciones consecutivas en el tiempo es un ruido blanco,  $Y_t = Y_{t-1} + \varepsilon_t$ . Este tipo de modelos se caracteriza por su simplicidad y se suelen emplear como inicio en la modelización. La predicción realizada por el modelo se hace coincidir, con el último valor observado, de tal forma que se expresa  $Y_{t+1}^* = Y_t$



Donde:  $Y_t$  es el valor observado de la variable aleatoria en el instante  $t$ , e  $Y_{t+1}^*$  la predicción que efectúa en el tiempo  $t$  sobre el valor que la serie tendrá en el periodo  $t+1$ .

### Modelos de Medias Móviles

La media móvil es una media aritmética que se caracteriza porque toma un valor diferente en cada momento del tiempo y porque en su cálculo no entran todas las observaciones de la muestra disponible. Para la aplicación de este método hay que definir su longitud ( $p$ ), es decir, el número de observaciones que intervienen en el cálculo de cada media móvil.

El término general de una media móvil de orden  $p$  (siendo  $p$  un número impar) centrada en el periodo  $t$  será:

$$MM(p)_t = \frac{\left( \frac{Y_{t-(p-1)/2} + Y_{t-1} + \dots + Y_{t+(p-1)/2} \right)}{p} \dots\dots\dots(2.9)$$

En el caso que  $p$  sea par, la expresión correspondiente vendrá dada por:

$$MM(p)_{t+0.5} = \frac{\left( \frac{Y_{t-(p-1)/2+0.5} + Y_{t-1} + \dots + Y_{t+(p-1)/2+0.5} \right)}{p} \dots\dots\dots(2.10)$$

Donde ahora la media móvil quedará descentrada o centrada entre el periodo  $t$  y  $t+1$ , es decir, en  $t+0.5$ . Los valores calculados a través de las medias móviles se utilizarán para predecir el valor de la variable en el periodo  $t+1$ .

### Modelo de Alisado Exponencial

Las técnicas de alisado o suavización exponencial se encuentran entre las más utilizadas en la predicción, por su facilidad de utilización y comprensión conceptual. Los estudios comparativos empíricos le asignan una considerable exactitud, robustez y flexibilidad de adaptación ante la posibilidad de un cambio en la estructura de los datos. Además, estos métodos suelen requerir menor número de datos que los métodos correlacionales.

Las series más propicias para la utilización de los métodos de alisado exponencial suelen ser aquellas cuyos datos no muestran una marcada estructura correlacional y, en particular, estos métodos suelen ser convenientes en situaciones en las que se presenta al menos una de estas circunstancias:

- El escaso número de datos históricos no permite calcular un número adecuado de los coeficientes de correlación.

- Las correlaciones calculadas no son estables. Cuando las autocorrelaciones de los datos permanecen aproximadamente constantes, el método de Box-Jenkins o el de regresión dinámica suelen presentar mejores resultados que el alisado exponencial.

Varios de los métodos de alisado exponencial resultan ser casos particulares de la familia de modelos de Box-Jenkins, aunque la filosofía de aplicación de ambas metodologías sea completamente diferente.

Los métodos de suavización exponencial se basan en un modelo estructural de datos de la serie cronológica en el que se supone que ésta presenta algunos o todos de los componentes ya comentados anteriormente, como son el nivel, la tendencia, componentes ciclo y la perturbación aleatoria.

A continuación se comentan brevemente algunos métodos de este tipo:

### **Alisado Exponencial Simple**

Cuando la serie presenta un comportamiento estacionario, es decir, no tiene tendencia y puede ser modelada como  $X_t = a + u_t$  (donde  $u_t$  es un término de perturbación aleatorio, con valor esperado cero y varianza constante para todo  $t$ , e independiente de  $X_t$  para todo  $t$ ), el método de predicción adecuado es el alisado exponencial simple (AES). Este método estima para cada período "T" el parámetro "a" como suma ponderada de todas las observaciones anteriores, dando mayor importancia a las observaciones más recientes que a las más antiguas. La expresión de cálculo es:

$$a_T = S_T = \alpha X_T + (1 - \alpha)S_{T-1} \dots\dots\dots(2.11)$$

Donde  $S_{T-1}$  es la estimación de  $a$  obtenida en el período T-1 y  $\alpha$  es la constante de alisado que toma valores entre 0 y 1. Como se observa, el AES actualiza período "a" período las estimaciones de "a" incorporando la nueva información.

La elección de la constante de alisado determina las características operativas del AES, ya que la rapidez con que se adaptan las predicciones a los posibles cambios experimentados por el valor de "a" depende de  $\alpha$ . Si  $\alpha$  es grande (próximo a 1) el AES se adapta rápidamente a los cambios experimentados en el valor de "a" y, en consecuencia, deberá escogerse un valor grande de  $\alpha$  cuando  $a$  es poco estable. Por el contrario, si la serie es muy estable, el valor de  $\alpha$  deberá ser pequeño para conseguir eliminar al máximo las fluctuaciones aleatorias debidas al término de perturbación y conseguir un mejor alisado.

La predicción del valor de  $X_t$  para los periodos  $T+1, T+2, \dots$  realizada al finalizar el periodo actual  $T$  es:

$$\hat{X}_{T+1|T} = \hat{X}_{T+2|T} = S_T \quad \dots\dots\dots(2.12)$$

En cuanto al valor de  $\alpha$ , la mayoría de los programas de ordenador proporcionan su valor óptimo por el simple procedimiento de elegir aquel valor de  $\alpha$  que minimiza la suma de los errores cuadráticos medios de predicción de las primeras 20 a 30 observaciones.

### 2.5.2 MODELO DE HOLT<sup>20</sup>

**Aplicable a series con tendencia y sin estacionalidad.** En el método de alisado exponencial de dos parámetros de Holt se supone que los datos fluctúan alrededor de un nivel variable  $b_t$  que sigue aproximadamente la tendencia de una línea recta que puede ser modelada como  $x_t = a + b_t + u_t$ . Las ecuaciones de predicción y actualización en este método son las siguientes:

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1}) \quad \dots\dots\dots(2.13)$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1} \quad \dots\dots\dots(2.14)$$

$$F_{t+m} = L_t + b_t m \quad \dots\dots\dots(2.15)$$

Donde  $\alpha$  y  $\beta$  son parámetros con valores entre cero y uno (cuanto menores sean estas constantes más alisada será la serie de predicciones), siendo  $L_t$  y  $b_t$  estimaciones del nivel y de la pendiente de la serie en el tiempo  $t$ , respectivamente. La ecuación (2.13) pone de manifiesto que la actualización del nivel se obtiene como una ponderación de los datos nuevos (primer sumando) y la mejor estimación del nuevo nivel basado en datos anteriores (segundo sumando) lo cual persigue que  $L_t$  tenga, aproximadamente el nivel del último dato. La ecuación (2.14) actualiza la tendencia. Finalmente la ecuación (2.15) es la que se utiliza para predecir los datos futuros.

### 2.5.3 MODELO DE HOLT-WINTER<sup>21</sup>

**Aplicable a series con tendencia y estacionalidad.** El método de Holt utiliza un modelo que no admite la presencia de variaciones estacionales. Si le añadimos un tercer término para recoger este elemento, tenemos el modelo de Winter que, considera la serie como formada por tres factores (además del aleatorio): nivel, tendencia y ciclos.

Como el método de Holt, este de Holt-Winter está pensado para series que presenten una tendencia aproximadamente lineal es su evolución. Ahora la expresión de obtención de las previsiones es:

$$x_{t+m} = T_{t+m} E_{t+m-L} \quad \dots\dots\dots(2.16)$$

<sup>20</sup> Publicación "Análisis de series temporales y Técnicas de Previsión"

En el que se supone que:  $T_{t+m} = S_t + m.b_t$  .....(2.17)

Donde:

$x_{t+m}$  es la previsión para el periodo  $t+m$  realizada desde  $t$

$T_{t+m}$  es la previsión sin estacionalidad (previsión de la tendencia)

$E_{t+m-L}$  es el factor estacional correspondiente al mismo periodo que se desea prever ( $t+m$ ), pero en el ciclo anterior

$S_t$  es el factor de nivel

$b_t$  es factor de tendencia (pendiente)

Sustituyendo en la expresión anterior, resultará:

$$x_{t+m} = (S_t + m b_t) E_{t+m-L} \dots\dots\dots(2.18)$$

Las expresiones correspondientes para el cálculo de los distintos elementos de la previsión, similares a las del método de Holt, son:

$$S_t = \alpha \frac{Z_t}{E_{t-L}} + (1-\alpha)(S_{t-1} + b_{t-1}) \quad 0 < \alpha < 1 \quad \dots\dots\dots(2.19)$$

$$b_t = \beta(S_t - S_{t-1}) + (1-\beta)b_{t-1} \quad 0 < \beta < 1 \quad \dots\dots\dots(2.20)$$

$$E_t = \gamma \frac{Z_t}{S_t} + (1-\gamma)E_{t-L} \quad 0 < \gamma < 1 \quad \dots\dots\dots(2.21)$$

Donde  $\alpha$ ,  $\beta$  y  $\gamma$  son las constantes de suavizado a estimar en cada caso, y  $L$  es la longitud del ciclo estacional. El primero está relacionado con el factor aleatorio de la serie, el segundo con la tendencia (representa la pendiente) y el tercero con la componente estacional. La mayor capacidad de representación de este modelo frente al anterior se consigue a costa de una mayor complejidad en la estimación, que requiere la exploración de los valores de los tres parámetros usando como criterio de selección el valor de la varianza residual o la desviación absoluta media, junto con un buen ajuste gráfico entre la serie original y la suavizada. Si los valores de  $\alpha$  y  $\beta$  superan el 0.5, es posible que el uso del método de Holt-Winters no sea adecuado.

#### 2.5.4 MODELO BOX-JENKINS<sup>22</sup>

**Método Dinámico.** En la metodología de Box-Jenkins el modelo se construye exclusivamente a partir de los datos sin necesidad de sugerir ningún modelo a priori que pudiera resultar arbitrario. Cuando los datos presentan una estructura correlacional considerablemente estable y no estacional el método Box-Jenkins aventaja con frecuencia a los modelos exponenciales de alisado simple y de Holt, debido a que estos modelos son casos particulares de la familia de modelos de Box-Jenkins (que también contemplan las

<sup>22</sup> Libro "Time Series Analysis: Forecasting & Control"

series estacionales), ventaja que no es empíricamente tan frecuente frente al modelo multiplicativo de Holt-Winters, posiblemente debido a que este método no se corresponde con ningún miembro de la familia de modelos Box-Jenkins, lo que sí ocurre, en cambio, para el modelo *aditivo* de Winters.

Entre las causas que posiblemente hayan dificultado una mayor aceptación de la metodología de Box-Jenkins pueden encontrarse su mayor complejidad, costo y requerimiento de número de datos históricos (al menos 50) en relación con otros métodos.

El objetivo principal de la metodología Box-Jenkins consiste en construir un modelo que reproduzca adecuadamente la estructura correlacional de los datos. Un modelo Box-Jenkins se construye a partir de las funciones muestrales de autocorrelación y de autocorrelación parcial de los datos correspondientes a la variable investigada. La optimización del método depende de las propiedades de los datos.

Según expusieron sus autores al presentar su método Box, G.E.P. y Jenkins, G.M. (1976), la primera etapa del mismo consiste en la identificación de uno o varios modelos candidatos de entre una extensa y compleja familia de modelos como una aproximación al verdadero proceso generador de los datos observados. Seguidamente se procede a la estimación de los parámetros del modelo mediante el ajuste del correspondiente modelo candidato a los datos históricos. A continuación tiene lugar la etapa de contrastes-diagnósticos, generalmente mediante el análisis de la autocorrelación de los residuos (diferencias entre los datos observados y los pronósticos del modelo correspondientes al período histórico). Si el modelo candidato no supera dichos contrastes-diagnóstico es rechazado, procediendo de nuevo a repetir el ciclo identificación-estimación-diagnóstico hasta que finalmente se acepta un modelo, momento a partir del cual se trabaja bajo el supuesto de que tal modelo representa la verdadera distribución estadística de los datos observados y, por extensión, también de los datos futuros. Se obtiene así la distribución de la serie en el futuro condicionada a las observaciones del pasado y del presente. Si la hipótesis anterior es correcta, la media condicional de tal distribución futura es la predicción óptima mínimo-cuadrática, siendo la varianza de dicha distribución un estimador válido de la precisión del pronóstico. En el caso de que más de un modelo candidato supere los *contrastes-diagnósticos*, Box y Jenkins recomiendan elegir el de estructura más simple.

El enfoque univariante Box-Jenkins incluye una familia de modelos que utilizan una notación estándar, teniendo como la representación más general un modelo ARIMA, la misma que se desarrollará en el presente trabajo.

## 2.6 FORMULACIÓN DE LA METODOLOGÍA ESTOCÁSTICA ARIMA<sup>23</sup>

La estructura típica del modelo que sirve de base para obtener la predicción de la variable energética  $X_t$  es:

$$X_t = f(X_{t-1}, X_{t-2}, \dots) + u_t \quad \dots\dots\dots(2.22)$$

La variable de error  $u_t$  se denomina ruido blanco, pues trata de una variable sin ninguna pauta reconocible en el tiempo que pueda ser utilizada para predecir el comportamiento de la variable dependiente. Desde un punto de vista estocástico, una serie temporal se define como el conjunto de valores observados de distintas variables aleatorias correspondientes a períodos de tiempo consecutivos, dichos períodos tienen la misma amplitud y la serie tiene un carácter discreto. Bajo este enfoque, una serie temporal por muy larga que sea, es considerada como una sola realización de un proceso estocástico. Para poder, a partir de una sola realización, efectuar inferencias sobre un proceso estocástico, este proceso debe cumplir las condiciones de estacionariedad y ergodicidad.

La estacionariedad se caracteriza mediante las siguientes propiedades: las esperanzas matemáticas de las variables aleatorias no dependen del tiempo (son constantes), las varianzas tampoco dependen del tiempo y son finitas, y las covarianzas entre dos períodos de tiempo distintos solamente dependen del lapso de tiempo transcurrido entre estos dos períodos. La función de autocorrelación (fac), que es la relación entre la covarianza y la varianza, y la función de autocorrelación parcial (facp), son los elementos estadísticos fundamentales en el proceso de identificación de un modelo predictivo ARIMA. En condiciones de estacionariedad la fac  $r_k$ , para todo  $k$ , se expresa como:

$$\rho_k = \frac{\text{cov}(X_t, X_{t+k})}{\text{var}(X_t)} \quad \dots\dots\dots(2.23)$$

La condición de ergodicidad permite que el proceso de inferencia se pueda realizar de una forma adecuada. De manera intuitiva podemos señalar que la ergodicidad posibilita obtener estimadores consistentes de los parámetros que configuran las funciones de auto covarianza y auto correlación. Por ejemplo, un proceso estocástico estacionario es ergódico en la media que es posible estimar consistentemente este parámetro haciendo uso de la media muestral temporal.

---

<sup>23</sup> Publicación "Predicciones de variables energéticas mediante análisis de series temporales"

En principio, la mayoría de series que representan sistemas económicos y energéticos no satisfacen las condiciones de estacionariedad, ello es debido a que en la serie original existe tendencia, la varianza no es constante y presenta variaciones estacionales (variabilidad de la media). Sin embargo, es posible transformar estas series no estacionarias en otras aproximadamente estacionarias mediante adecuadas operaciones algebraicas. Una vez realizada la transformación se aplica el análisis predictivo diseñado para series estacionarias de origen.

### 2.6.1 MODELOS ARIMA

El condicionante fundamental para la aplicación de un modelo estocástico hacia la descripción y pronóstico de una variable temporal discreta, es que ésta obedezca a un proceso estocástico lineal estacionario. En nuestro caso las variables históricas contabilizadas en forma mensual poseen patrones de tendencia y estacionalidad, consecuentemente no son estacionarias de origen. Por tanto, para la aplicación de la metodología estocástica, previamente se tienen que hacer transformaciones matemáticas a la variable original.

Para salvar la no estacionariedad ocasionada por el patrón de tendencia, se procede a diferenciar regularmente a la variable original. El número de veces que debe diferenciarse hasta alcanzar la estacionariedad constituye el grado u orden de homogeneidad o integración.

Si la serie original,  $X_t$ , es homogénea de orden  $d$ , entonces:

$$(1-L)^d X_t = Z_t \quad \dots\dots\dots(2.24)$$

Donde:  $t = 1, 2, \dots, T$

Con lo que la nueva serie  $Z_t$  es estacionaria, siendo "d" el orden de integración regular, y "L" el operador de retardos.

En general, diremos que un proceso es integrado de orden "h" cuando al diferenciarlo "h" veces se obtiene un proceso estacionario<sup>24</sup>. Así, a un proceso integrado  $X_t$  se le denomina proceso autorregresivo-medias móviles integrado ARIMA (p,d,q), si tomando diferencias de orden d se obtiene un proceso estacionario  $Z_t$  del tipo ARMA (p, q), el cual se expresa como sigue:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + u_t - \theta_1 u_{t-1} - \dots - \theta_q u_{t-q} \quad \dots\dots\dots(2.25)$$

ó

$$\phi(L)(1-L)^d X_t = \theta(L)u_t \quad \dots\dots\dots(2.26)$$

<sup>24</sup> Libro "Estadística Modelos y métodos"

Donde los  $\phi_i$  son coeficientes de la parte autorregresiva, los  $q_i$  son coeficientes de la parte de medias móviles,  $f(L)$  es el operador autorregresivo,  $q(L)$  es el operador de medias móviles, y  $(u_t, u_{t-1}, u_{t-2}, \dots)$  es una secuencia de perturbaciones aleatorias distribuidas idéntica e independientemente con media cero y varianza constante, lo que se conoce como ruidos blancos.

Donde  $\phi(L)$  es el operador polinomial de retardos de un proceso estocástico lineal estacionario del tipo autorregresivo puro de orden  $p$ , definido por:

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad \dots\dots\dots(2.27)$$

Donde  $\theta(L)$  es el operador polinomial de retardos de un proceso estocástico lineal estacionario del tipo medias móviles puro de orden  $q$ , definido como:

$$\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \dots - \theta_q L^q \quad \dots\dots\dots(2.28)$$

La eliminación de la tendencia de la serie original suele conseguirse mediante las diferenciaciones implícitas. Sin embargo, en ocasiones se observa también que existe una tendencia en la varianza, esto es, que la dispersión de las observaciones no es constante a lo largo del tiempo, la cual no se elimina mediante estas diferenciaciones. Cuando se presenta esta situación la transformación adecuada puede consistir en tomar logaritmos neperianos. Esta posibilidad de transformar la serie se puede concretar de forma más general mediante la transformación Box-Cox, quedando el modelo ARIMA expresado por:

$$\phi(L)(1-L)^d (X_t^{(\lambda)} - \mu) = \theta(L)u_t \quad \dots\dots\dots(2.29)$$

Donde  $\mu$  es la media de  $X_t^{(\lambda)}$  con potencia de transformación  $\lambda$  ( $\lambda$  es un número real), siendo:

$$X_t^{(\lambda)} = \frac{X_t^{(\lambda)} - 1}{\lambda}, \text{ si } \lambda \neq 0 \quad \dots\dots\dots(2.30)$$

$$X_t^{(\lambda)} = \ln X_t, \text{ si } \lambda = 0 \quad \dots\dots\dots(2.31)$$

Para determinar experimentalmente el valor adecuado de  $\lambda$  se consideran diversos valores de este parámetro y se elige aquél que consiga que el gráfico de media-desviación típica se aproxime a una línea horizontal. También se puede determinar aquel valor de  $\lambda$  para el que sea mínima la correlación entre la desviación típica y la media de la serie. En los casos frecuentes en que la variabilidad de los datos aumenta con el nivel de los mismos se suele conseguir la estacionariedad de la varianza mediante la transformación logarítmica ( $\lambda=0$ ) o calculando la raíz cuadrada ( $\lambda=0,5$ ) de los datos originales.



En cuanto a la estacionariedad en media, el procedimiento habitual para conseguirla suele consistir en diferenciar regular y/o estacionalmente los datos (los cuales ya se suponen estacionarios en varianza)  $Y_t$ , construyendo una nueva serie  $Z_t = Y_t - Y_{t-1}$ . Caso de que después de realizar la operación mencionada todavía no se haya conseguido la estacionariedad de la media, puede ser necesario diferenciar la serie  $Z_t$  pero, en la práctica, no suele ser necesario diferenciar los datos más de una o dos veces para lograr dicha estacionariedad, teniendo en cuenta que no se debe incurrir en una sobre diferenciación superflua de los datos que distorsionaría la naturaleza de la serie que se pretende analizar.

La ecuación general de un modelo ARIMA está referida a la variable diferenciada  $w_t$  correspondiendo a otra ecuación en la variable original  $Y_t$  a la que se denomina la variable integrada (I) de  $w_t$ , lo que añadido al proceso autorregresivo (AR) de medias móviles (MA) al que se refiere la ecuación explica el término ARIMA con que se suele denominar a estos procesos.

Otra fuente de no estacionariedad en muchas de las series reales, como es este caso de proyección de demanda donde se contabiliza la variable en forma mensual, lo constituye la estacionalidad. Para desestacionalizar la serie se procede con la diferenciación estacional.

A un proceso integrado estacional  $X_t$  se le denomina proceso autorregresivo-medias móviles integrado, ARIMA (P, D, Q), si tomando diferencias estacionales de orden D, con periodo "s", se obtiene un proceso estacionario  $Z_t$  del tipo ARMA (P, Q) y se expresa mediante:

$$Z_t = \Phi_1 Z_{t-s} + \Phi_2 Z_{t-2s} + \dots + \Phi_p Z_{t-ps} + \delta + u_t - \Theta_1 u_{t-s} - \dots - \Theta_q u_{t-qs} \quad \dots\dots\dots(2.32)$$

Donde:

$$Z_t = (1 - L^s)^D X_t$$

De forma resumida también se puede expresar como:

$$\Phi_p(L^s)(1 - L^s)^D X_t = \delta + \Theta_q(L^s)u_t \quad \dots\dots\dots(2.33)$$

Donde:

$$\Phi_p(L^s) = 1 - \Phi_1 L^{1s} - \Phi_2 L^{2s} - \dots - \Phi_p L^{ps} \quad \dots\dots\dots(2.34)$$

$$\Theta_q(L^s) = 1 - \Theta_1 L^{1s} - \Theta_2 L^{2s} - \dots - \Theta_q L^{qs} \quad \dots\dots\dots(2.35)$$

Así, los  $\Phi_i$  son coeficientes de la parte estacional autorregresiva, los  $\Theta_i$  son coeficientes de la parte estacional de medias móviles,  $\Phi(L^s)$  es el operador autorregresivo estacional,  $\Theta(L^s)$  es el operador de medias móviles estacional, D es el orden de diferenciación estacional y s es el número de observaciones anuales.

En forma general, los modelos que tienen en cuenta tanto el efecto de tendencia como la estacionalidad de una serie temporal se denominan Modelos Estacionales Multiplicativos ARIMA (p, d, q)x(P, D, Q)s, y se expresan como:

$$\Phi_p(L^s)\phi_p(L)(1-L^s)^D(1-L)^d X_t = \Theta_q(L^s)\theta_q(L)u_t \quad \dots\dots\dots(2.36)$$

En forma general, esta expresión se puede dar como:

$$\Phi_p(L^s)\phi_p(L)[(1-L^s)^D(1-L)^d X_t - \mu] = \Theta_q(L^s)\theta_q(L)u_t \quad \dots\dots\dots(2.37)$$

Donde:

$\mu$  es la media de  $Z_t = (1-L^s)^D(1-L)^d X_t$

### 2.6.2 ETAPAS PARA LA ELABORACIÓN DEL MODELO ARIMA

Partiendo de la serie temporal se trata de averiguar qué modelo ARIMA(p,d,q)xARIMA(P,D,Q) es susceptible de haber generado dicha serie, es decir, qué modelo representa adecuadamente el comportamiento de la misma, con el fin de utilizarlo para obtener predicciones de valores futuros de la serie en cuestión. Para ello se siguen cuatro etapas: identificación, estimación, validación, y predicción.

#### Identificación

Identificar una serie temporal consiste en inducir a partir de los datos, la función de autocorrelación muestral y la función de autocorrelación parcial muestral, indicando qué modelos ARIMA se adaptarían mejor a las características de la serie. Cuando se trata de una serie no estacionaria, primeramente se procede a estacionar la serie, tanto en media, es decir, identificación de los valores d y D (estacionalidad), como en varianza, esto es identificar el valor de  $\lambda$ . Una vez que esta serie transformada es estacionaria (en media y en varianza) se deben de averiguar los posibles valores tanto de la parte regular del modelo (autorregresiva "p", y medias móviles "q") como de la parte estacional (autorregresiva "P", y medias móviles "Q").

#### Estimación

Identificados los posibles modelos que han podido generar la serie temporal, se trata de cuantificar sus parámetros. Los dos problemas fundamentales a los que se enfrenta la estimación de los modelos ARIMA son el de los valores iniciales (de los parámetros, de la serie y de los ruidos) y el de no linealidad.

Se trata de estimar los parámetros  $B_i, i = 0, \dots, p + P + q + Q$ , donde:

$$B_i = \phi_i; i = 1, \dots, p$$

$$B_i = \Phi_{i-p}; i = p + 1, \dots, p + P$$

$$B_i = \theta_{i-(p+P)}; i = p + P, \dots, p + P + q$$

$$B_i = \Theta_{i-(p+P+q)}; i = p + P + q + 1, \dots, p + P + q + Q$$

$$B_i = C; i = 0$$

Si  $B_i$  es la estimación del parámetro  $B_i$ , la primera etapa en la validación del modelo consistirá en comprobar si los coeficientes  $B_i$  son significativamente distintos de cero. Para ello, sobre cada parámetro se planteará la hipótesis nula, esto es  $H_0: B_i = 0$ . Dicha hipótesis puede ser interpretada como que la variable asociada al parámetro  $B_i$  no mejora el ajuste con respecto al obtenido con las restantes variables incluidas en el modelo. Si el p-valor asociado al valor del estadístico de contraste  $t$  es menor que  $\alpha$ , se rechazará la hipótesis nula al nivel de significación  $\alpha$ .

### Validación

La etapa de validación se centra fundamentalmente en analizar si los residuos del modelo ( $u_t$ ) tienen un comportamiento similar a las perturbaciones del mismo ( $u_t$ ), esto es, si puede afirmarse que son semejantes a un ruido blanco. Adicionalmente, se tratará de comprobar la calidad de las estimaciones, así como el cumplimiento de las condiciones de estacionariedad e invertibilidad que deben satisfacer los parámetros de estos modelos.

### Predicción

Tras la validación viene el fin básico de esta metodología, esto es, la obtención de predicciones de valores futuros de la serie temporal. Una vez obtenidas las predicciones del modelo se trata de volver a chequear la adecuación del mismo, pudiendo utilizar para ello, tanto métodos no paramétricos (como el error cuadrático medio) como paramétricos (estadísticos de contenido informativo, exactitud y corroboración).

## 2.6.3 ESTRUCTURA DE CÁLCULO CON MODELOS ARIMA<sup>25</sup>

Una vez cuantificada la serie energética histórica se procede a determinar qué modelo óptimo ARIMA es susceptible de haber generado dicha serie, el cual servirá para determinar las predicciones. Este proceso de cálculo se resume en el diagrama de flujo mostrado a continuación.

<sup>25</sup> Publicación "Modelos estocásticos ARIMA para predicción de variables energéticas"

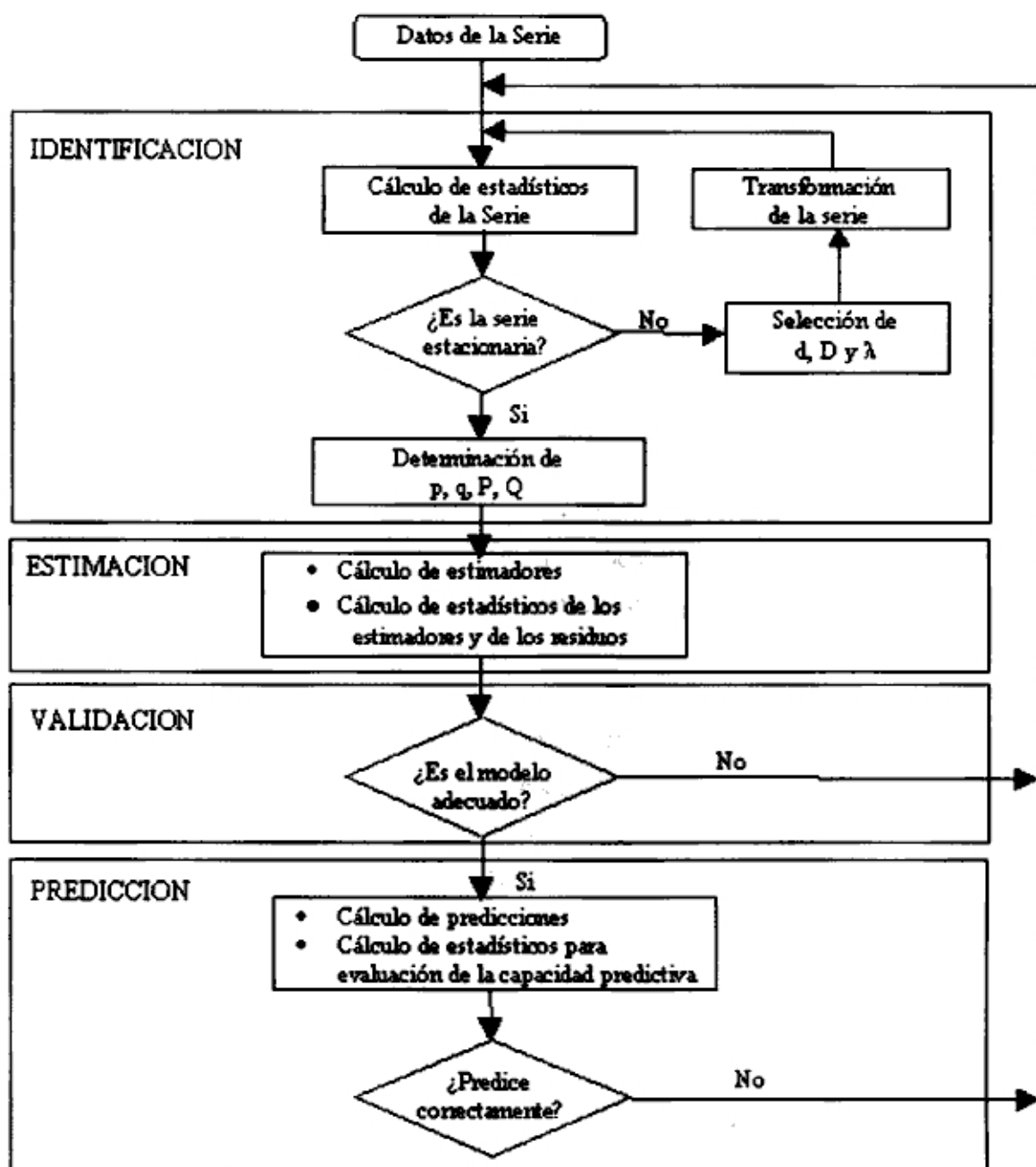


Fig. 2.2 Diagrama de flujo para la elaboración del modelo ARIMA

### Conformación de la serie temporal

Para elegir el tamaño de la serie temporal Consumo\_Energía\_GWH a efectos de su modelamiento predictivo, se debe tener en consideración las siguientes características:

- Tamaño de la muestra: Desde el punto de vista cuantitativo, cuantas más observaciones posea la serie habrá más posibilidades de ensayar diversos modelos de predicción.
- Patrones de comportamiento histórico: Cuando la serie muestra patrones de tendencia y estacionalidad bien definidos a lo largo del horizonte histórico, permitirá mejores posibilidades de alcanzar una buena calidad de predicción estadística.

- Origen de los datos: Dependiendo del grado de manipulación de cada una de las observaciones de la serie, estas pueden proporcionar mayor o menor representatividad estadística.

### **Elección del horizonte predictivo**

No existe un patrón específico que indique una relación entre el tamaño del horizonte histórico y el tamaño del horizonte predictivo, todo dependerá de las características intrínsecas de la serie temporal de estudio. De acuerdo a las características evolutivas de la muestra, así como de la capacidad predictiva que se ha de demostrar en la elección del modelo predictivo, es posible hacer predicciones en una proporción de 1 a 4 y 5 respecto al número de observaciones mensuales. En nuestro caso se trata de una serie de 168 observaciones de la serie temporal Consumo\_Energía\_GWH, por lo que, determinamos que nuestro horizonte de predicción será de tres años o 36 meses. Horizonte sugerido que se sustenta tomando en consideración la limitación de contar con una base de datos adecuada, y regular capacidad evolutiva.

De otro lado, se estima que las predicciones tienen las características de un proceso dinámico que debe ser validado año a año una vez obtenido el modelo predictivo, en base a lo cual, una vez comprobada las predicciones de esta tesis se podrá establecer un periodo de predicción de cinco años como mínimo, para un trabajo similar futuro.

## **CAPÍTULO III**

### **MODELAMIENTO Y PREDICCIÓN**

En este capítulo se busca encontrar un modelo a través del análisis de series temporales, que ayude a la descripción y previsión de la evolución de la demanda mensual de energía eléctrica residencial en la zona sur de Lima Metropolitana. El conocimiento de lo sucedido en el pasado a través de la observación de sus regularidades constituye un soporte para la elaboración de predicciones con respecto al futuro. En este apartado se analizan los datos de la demanda mensual de energía eléctrica en el sector residencial de la citada zona durante el periodo comprendido entre enero de 1996 y diciembre de 2009. La base de datos utilizada procede de la data que maneja la empresa Luz del Sur S.A.A, sin embargo, se tuvo que trabajar la citada base de datos para poder seleccionar aquellos consumos que corresponden sólo a clientes residenciales en baja tensión, en ese sentido, se designa por Consumo\_Energía\_GWH a la serie cronológica definida por tales datos.

En la primera sección, se efectúa análisis macro del consumo de energía eléctrica en el sector residencial bajo la zona de concesión. En la segunda sección, se detallan las fuentes datos utilizadas en este trabajo y las limitaciones que se han encontrado en la obtención de la base de datos a utilizar. En la tercera sección, se desarrolla el proceso de cálculo predictivo bajo el método de Winter. En la cuarta sección se muestran los resultados de la predicción realizada con la metodología ARIMA y con el software SPSS. Finalmente, en la quinta sección se efectúa el análisis comparativo de los resultados obtenidos.

#### **3.1 CONSUMO DE ENERGÍA ELÉCTRICA EN EL SECTOR RESIDENCIAL**

El presente estudio sobre el consumo de energía residencial de la zona de Lima Sur se desarrolló tomando como referencia la zona de concesión de la empresa Luz del Sur SAA. Sobre el particular, Luz del Sur (antes Edelsur) fue creada el 1° de enero de 1994. El 18 de agosto de 1994 Ontario Quinta A.V.V. adquirió el 60% de las acciones de la Empresa del

Estado. El Estado Peruano vendió al sector privado la totalidad de las acciones que mantenía en Luz del Sur. En agosto de 1996 la empresa se separó creándose la actual Luz del Sur SAA a la que se le transfirió la concesión de distribución de energía eléctrica.

La zona de concesión comprende un área aproximada de 3.000km<sup>2</sup>, que incluye 30 de los más importantes distritos de Lima, los que en conjunto superan los 3 millones de habitantes. En esta zona, que se extiende a lo largo de 120km de costa se concentra la más importante actividad comercial, de servicios, turística y una significativa parte de las empresas productivas del país. Esto, sumado al creciente estándar de vida de la población y a la gran disponibilidad de recursos constituye un excelente potencial de desarrollo para la industria, el comercio, y negocios de diverso tipo.

Es de precisar que, la zona de concesión se compone por 35 distritos con un total de 877,297 clientes totales a Dic.2006, entre los que se tiene principalmente a los siguientes: Asia, Ate-Vitarte, Barranco, Cieneguilla, Chaclacayo, Chilca, Chorrillos, Chosica, El Agustino, Jesús María, La Molina, La Victoria, Lima Cercado, Lince, Lurín, Mala, Matucana, Miraflores, Pachacamac, Pucusana, Punta Hermosa, Punta Negra, Ricardo Palma, San Bartolo, San Borja, San Isidro, San Juan Miraflores, San Luis, San Mateo, Santa Anita, Santa Eulalia, Santiago de Surco, Surquillo, Villa El Salvador, Villa María del Triunfo, tal como se grafica a continuación.



Fig. 3.1 Zona de Concesión Luz del Sur SAA

Finalmente, como anexo se tiene la Grafica A.0 que muestra la tipología de la red a evaluar, evidenciada en el esquema unifilar de la red de alta tensión de toda la zona de concesión de Luz del Sur SAA, actualizado a la fecha del presente análisis<sup>26</sup>.

### 3.1.1 CARACTERÍSTICAS DE LAS REDES ELÉCTRICAS

Las redes de las concesionarias eléctricas de acuerdo al alcance y cantidad de suministros que incluye se dividen o segmentan como se detalla a continuación:

- a. Subestaciones de distribución (SED): Área geográfica determinada por la frontera eléctrica de las redes y suministros en baja tensión de cada SED.

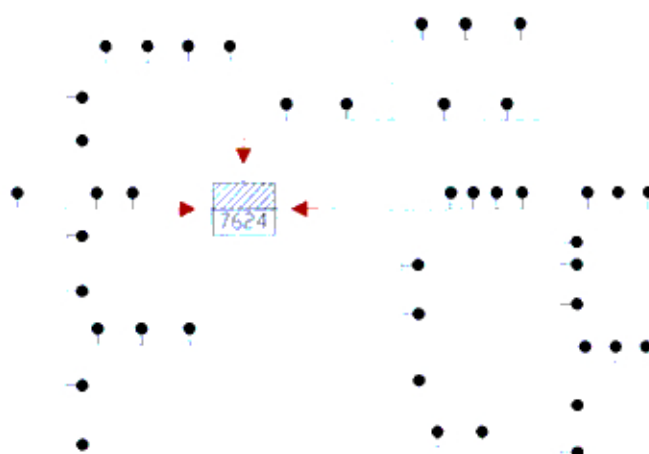


Fig. 3.2 Subestación de distribución N° 7624

- b. Alimentadores de MT: Área geográfica determinada por la frontera eléctrica de cada alimentador que es la agrupación de varios SED y clientes regulados en media tensión.



Fig. 3.3 Alimentador U24

- c. Subestaciones de transformación (SET): Área geográfica determinada por la frontera eléctrica de cada SET que es la agrupación de varios alimentadores en MT y clientes libres en media tensión.

<sup>26</sup> Obtenido de página web <http://www.luzdelsur.com.pe>



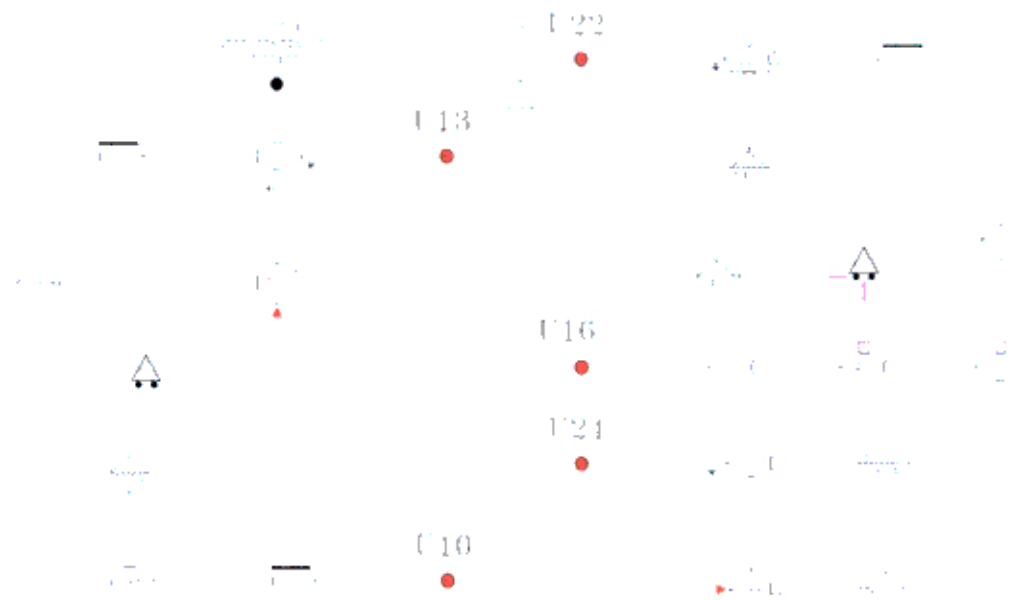


Fig. 3.4 Subestación de transformación Neyra (U)

**Dinámica de la topología de las redes:** Los sistemas de distribución o redes eléctricas son dinámicos, es decir, las configuraciones de las redes varían en el tiempo, ya sea por ampliaciones y/o traslados de cargas de las redes. De igual manera, las redes de las propias SED se amplían o reemplazan. Por lo que, para el presente estudio el sistema eléctrico a evaluarse corresponde a una topología fija en el tiempo con la estructura presentada en Diciembre 2006.

### 3.1.2 IMPORTANCIA DE LA SERIE ENERGÉTICA

El consumo de energía eléctrica en el sector residencial como serie energética representa dentro de la zona de concesión una variable representativa y de importancia por la información que contiene, como el número de clientes que representa el 90,1% del total de clientes. Asimismo, el consumo o ventas de energía residencial (GWH) representan el 39,7% del consumo de energía en toda la zona de concesión<sup>27</sup>, es decir, más de la tercera parte de la energía que vende Luz del Sur es energía residencial, como se observa de los gráficos mostrados a continuación.



Fig. 3.5 Distribución de Clientes (2008)

<sup>27</sup> Información obtenida de página web <http://www.luzdelsur.com.pe/>

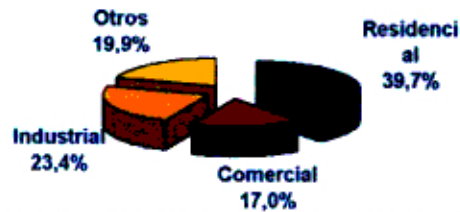


Fig. 3.6 Distribución Ventas Totales (GWh) (2008)

De otro lado, las inversiones en infraestructura eléctrica para toda la zona de concesión en el año 2008 fue por un monto total de 47,1 millones US\$, representando un crecimiento del 27% respecto del año 2007, con una inversión anual por distrito de 1,35 millones US\$ aproximadamente<sup>28</sup>; por ello, es necesario una correcta proyección del crecimiento del consumo de energía (GWh) pues la importancia de la predicción de esta variable radica en que puede determinar una posible carencia de capacidad de suministro de energía o, por el contrario, un exceso de capacidad que pudiera aconsejar la no utilización de parte del parque ya existente, como instrumento previo para la toma de decisiones en las inversiones.

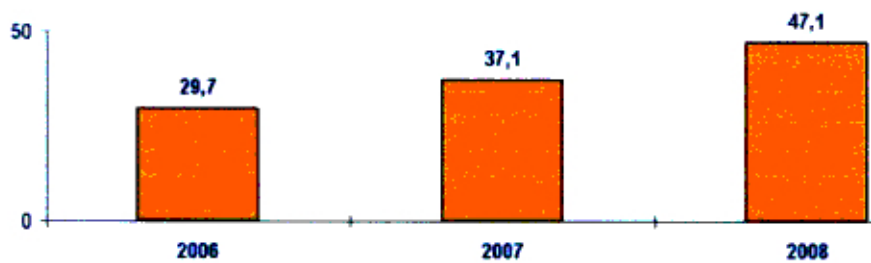


Fig. 3.7 Inversión anual (millones US\$)

En ese sentido, la predicción del consumo eléctrico de la zona de concesión se basa fundamentalmente en las predicciones que se realicen sobre las puntas de potencia y energía demandada. Así, los consumos totales de energía se desagregan de acuerdo al tipo de carga, siguiendo los criterios aplicados por el Ministerio de Energía y Minas para la proyección de la demanda global del SEIN<sup>29</sup>:

- a) Cargas Vegetativas: Definidas como las cargas de las que se dispone de amplia información estadística.
- b) Cargas Especiales: Definidas como cargas industriales o mineras que por su magnitud, su alto factor de carga y reciente incorporación al sistema eléctrico.

Como se demostrará más adelante en algunas SET de la zona de concesión el crecimiento porcentual de las cargas vegetativas es igual al crecimiento del consumo de energía del sector residencial.

<sup>28</sup> Información obtenida de página web <http://www.luzdelsur.com.pe/>

<sup>29</sup> Plan Referencial de Electricidad 2008-2017

### 3.1.3 ANÁLISIS PRELIMINAR

De otro lado, se tuvo acceso a una base de datos donde se incluye información respecto de los consumos de energía anual por alimentador (incluidos suministros en baja y media tensión), del cual se obtuvo los consumos históricos de energía anual por SET desde el año 1996 hasta el 2006, cuyos valores se muestran en el Tabla A.2 del Anexo, de donde se obtiene el gráfico mostrado a continuación, apreciándose que los consumos de energía de todos los SET presentan tendencia creciente, pudiendo identificarse que los diez de mayor consumo de energía son: SET Gálvez, SET Luis Neyra, SET Puente, SET Balnearios, SET Limatambo, SET San Isidro, SET Salamanca, SET Barranco, SET Villa María y SET Chorrillos.

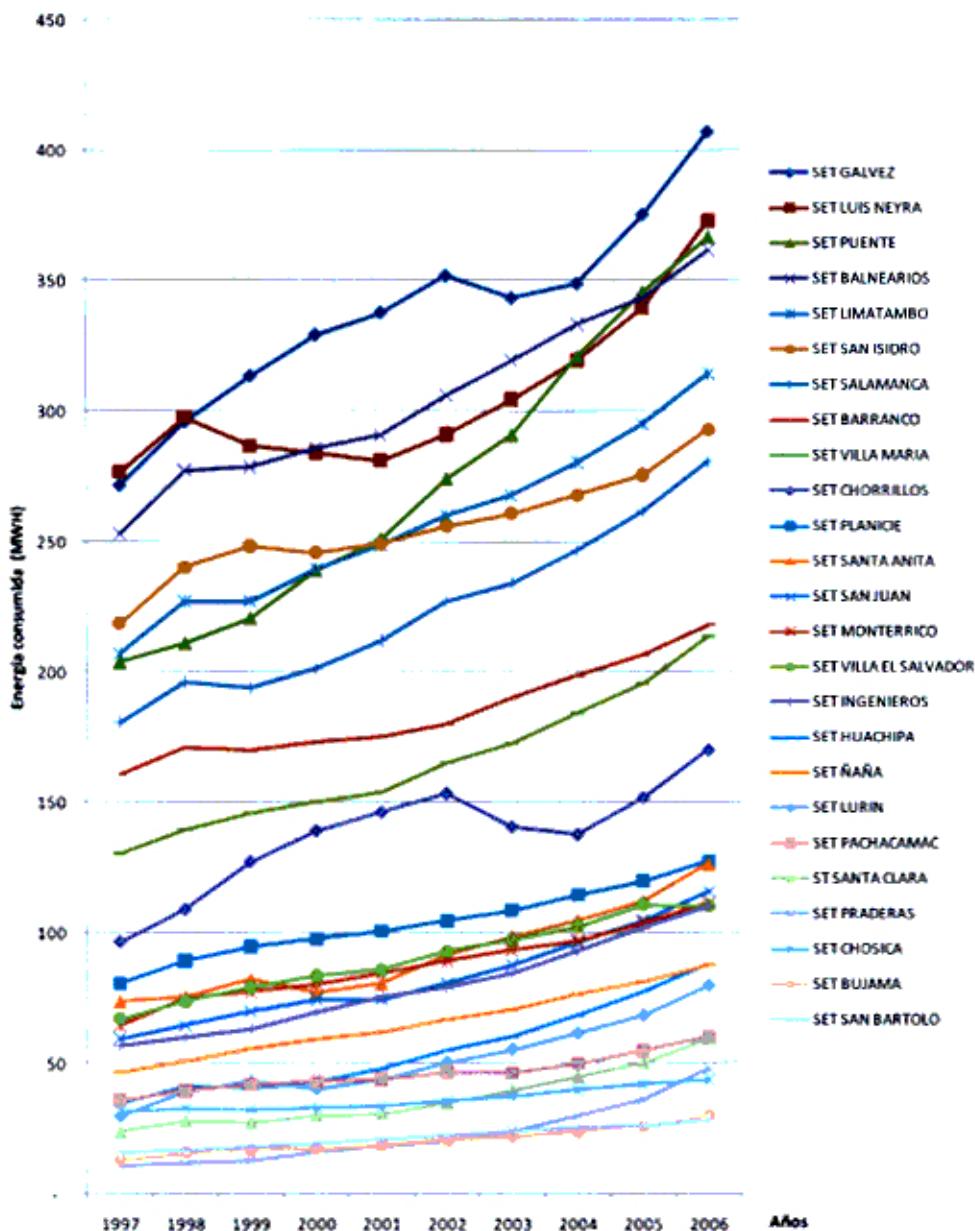


Fig. 3.8 Consumo total de energía por SET (Sub Estación de Transformación)

Asimismo, se analizó la relación entre los consumos de energía en baja y media tensión respecto del consumo total por SET en los últimos diez años, tomando como referencia los diez SET de mayor consumo, cuyo sustento están en el cuadro y gráficos mostrados en la Tabla A.3 del Anexo. De cuyo análisis se advierte que, la carga total del SET Villa María tiene un consumo de energía en baja tensión que representa el 89% de la carga total; asimismo, el SET Gálvez tiene un consumo de energía BT representa el 86% del total, por lo que, podría afirmarse que ambos SET presentan comportamiento mayoritariamente residencial; entendiéndose por comportamiento residencial, aquel sistema que presenta en su diagrama de carga diaria, el mayor consumo de energía en el periodo de Hora Punta.

De otro lado, por ejemplo el SET Puente tiene un consumo de energía en MT representa el 57% de la carga total, por lo que podría afirmarse que tiene un comportamiento industrial; entendiéndose por comportamiento industrial aquel sistema que presenta un diagrama de carga diaria con mayor consumo de energía constante en el periodo de Horas Fuera de Punta.

Por lo expuesto, se puede afirmar que el consumo de energía total de cualquier SET tendrá comportamiento residencial o industrial, dependiendo si éste tiene mayor porcentaje de carga en baja o media tensión, respectivamente.

### **3.2 ANÁLISIS DE BASE DE DATOS**

Como es de conocimiento general una correcta previsión requiere datos lo más exactos posibles, ya que sin ellos los resultados no son confiables. Sin embargo, a la fecha no existen datos disponibles en ningún estudio preliminar a nivel de energía eléctrica residencial en Lima Sur, data requerida para la previsión planteada en el presente trabajo. En algunos casos los datos deseados sólo están disponibles a nivel nacional, resultando inadecuados para la previsión de la demanda de energía eléctrica residencial en la citada zona de concesión. En otros casos, sólo ha sido posible obtener los valores de algunas variables de frecuencia anual, pero no de frecuencia mensual.

En el presente trabajo se han obtenido los datos necesarios para explicar la demanda de energía eléctrica mensual en el sector residencial de la zona sur de Lima. Los datos analizados en el trabajo se utilizan como series cronológicas del consumo de energía para el periodo comprendido entre enero de 1996 y diciembre de 2009.

En ese sentido, el primer paso llevado a cabo fue la recolección y evaluación de la información disponible. Del análisis efectuado, se ha llegado a la conclusión que para el caso de los concesionarios de electricidad no se cuenta con un sistema uniforme de información y, que los datos disponibles tienen que ser procesados en cada oportunidad que se requiera efectuar un análisis en particular. Esto hace ver la necesidad de realizar esfuerzos a la implementación de bases de datos de buena calidad, con el propósito de evaluar diferentes aspectos de la demanda eléctrica en determinada zona de concesión.

Los datos del consumo de energía proceden de la base de datos de la empresa concesionaria Luz del Sur S.A.A. De coordinaciones efectuadas, la Gerencia de Sistemas se encargó de recopilar información respecto de la energía consumida mes a mes por cada suministro o cliente bajo el ámbito de la concesión. Además se solicitó otra base de datos en donde se incluya información de cada cliente en cuanto, por ejemplo, al tipo de tarifa de cada suministro y distrito de ubicación, todo ello con la finalidad de tener mayor facilidad al momento de trabajar las series.

**Depuración:** Luego de obtener la base de datos inicial se procedió a realizar un tratamiento preliminar de depuración al universo de suministros eléctricos o clientes, así por ejemplo:

- Se eliminaron los consumos de todos los totalizadores, pues en ellos se incluye el consumo de dos o más suministros eléctricos para evitar duplicidad en consumos de energía.
- Se eliminaron lecturas atípicas, debido a errores de facturación que posteriormente se regularizan por medio de refacturaciones. Las más notorias encontradas son los consumos atípicos por lectura en un medidor nuevo, por ejemplo.
- Para aquellos suministros antiguos que no poseían ubicación geográfica en el sistema, se procedió a ubicarlos uno a uno por medio de su sector, carga, y dirección.
- Se eliminaron los suministros que no registran consumo de energía en los últimos doce meses, tomando como referencia diciembre del 2006.
- Se eliminaron todos los suministros en baja tensión que pertenecían a las tarifas BT2, BT3 y BT4 por no corresponder a clientes residenciales, pues muchos de estos corresponden a oficinas administrativas, pequeñas fábricas, etc, y su carga no tiene característica netamente residencial.

**Nivel de tensión:** El nivel de alimentación de algunos suministros no permanece constante en el tiempo. Así existen suministros que se iniciaron en BT y migraron a MT o viceversa. Para el caso de suministros que pasaron de BT a MT, se considerará que siempre han sido de MT, ya que se trata de un cliente con energía creciente significativa. Para el caso de

suministros que pasaron de MT a BT, también se considerará que siempre han sido de MT, pues se redujo la demanda del cliente de MT y su consumo pasó a ser pequeño, el ignorar un cliente en BT no afecta la evaluación grupal de la demanda de BT.

**Base de datos final:** La base de datos fue trabajada con el programa Microsoft Office Access para generar las consultas que permitan obtener información importante, así por ejemplo, se obtuvo los consumos de energía mensual por alimentador, por SET y por nivel de tensión. Se obtuvo finalmente la base de datos con el consumo de energía total residencial en la zona de concesión de Lima Sur, como se muestra en el Tabla A.1 desde Ene.1996 hasta Dic.2009, denominada Consumo\_Energía\_GWH.

De otro lado, en el trabajo elaborado por la empresa Consultora APEI<sup>30</sup> se realizaron corridas sobre la base de 132 meses de energía, a nivel de SET, alimentador, SED y sector, concluyéndose que, dada la poca carga existente SED y alimentadores, y su no regularidad, se obtienen inconsistencias y pobres ajustes, siendo por esto no recomendado para el uso de proyección, pues poseen bajos valores (energía) y presentan altas variaciones. A pesar de las diversas alternativas de transformaciones para estacionalizar la serie original, estas series no pueden pasar a ser estacionaria (en media o varianza), por lo que los resultados del cálculo resultan dudosos.

### 3.3 PROCESO DE CÁLCULO PREDICTIVO. MÉTODO DE WINTER

En esta sección se procede a estimar el modelo de series de tiempo con datos mensuales del consumo de energía residencial aplicando el Método de Winter, ya que la serie presenta tendencia creciente y estacionalidad. Este procedimiento suaviza componentes irregulares de datos de series temporales, para lo cual hace uso de una variedad de modelos que incorporan diferentes supuestos acerca de la tendencia y la estacionalidad. A continuación se presentan los resultados obtenidos de aplicar el software SPSS a la serie temporal Consumo\_Energía\_GWH.

Para determinar las constantes de alisado que minimizan los errores de estimación el software SPSS ofrece dos opciones para ser determinados. Una opción es con la denominada "Búsqueda en rejilla" donde se puede especificar un intervalo de valores para cada una de las constantes de alisado Alfa, Gamma y Delta, entre los cuales el sistema determinará aquellos que optimicen la predicción. Otra opción, es tomar los valores de las

---

<sup>30</sup> "Evaluación del Programa ARIMA"

constantes de alisado que el software proporciona por defecto. Para nuestro análisis se han desarrollado ambos casos a fin de determinar cual ofrece mejor ajuste predictivo.

Los coeficientes estacionales que el método Winter utiliza para inicializar el alisado son los que se obtienen aplicando el método de descomposición estacional establecido por defecto en el sistema.

### CASO A: "Búsqueda en rejilla"

Con lo cual el sistema seleccionará el valor de  $\alpha$  que proporcione menor suma de cuadrados de los errores de estimación (SSE), obteniéndose los siguientes resultados:

#### Descripción del modelo

Nombre del modelo	MOD_5	
Serie	1	Consumo_Energia_GWH
Modelo multiplicativo de Winters	Tendencia	Lineal
	Estacionalidad	Multiplicativo
Longitud del periodo estacional		12

Aplicando las especificaciones del modelo de MOD\_5

#### Parámetros del suavizado

Serie	Alpha (Nivel)	Gamma (Tendencia)	Delta (Estación)	Sumas de los errores cuadráticos	gl error
Consumo_Energia_GWH	.10000	.10000	.00000	2970.71823	155

A continuación, se muestran los parámetros con las sumas menores de errores cuadráticos. Estos parámetros se utilizan para pronosticar.

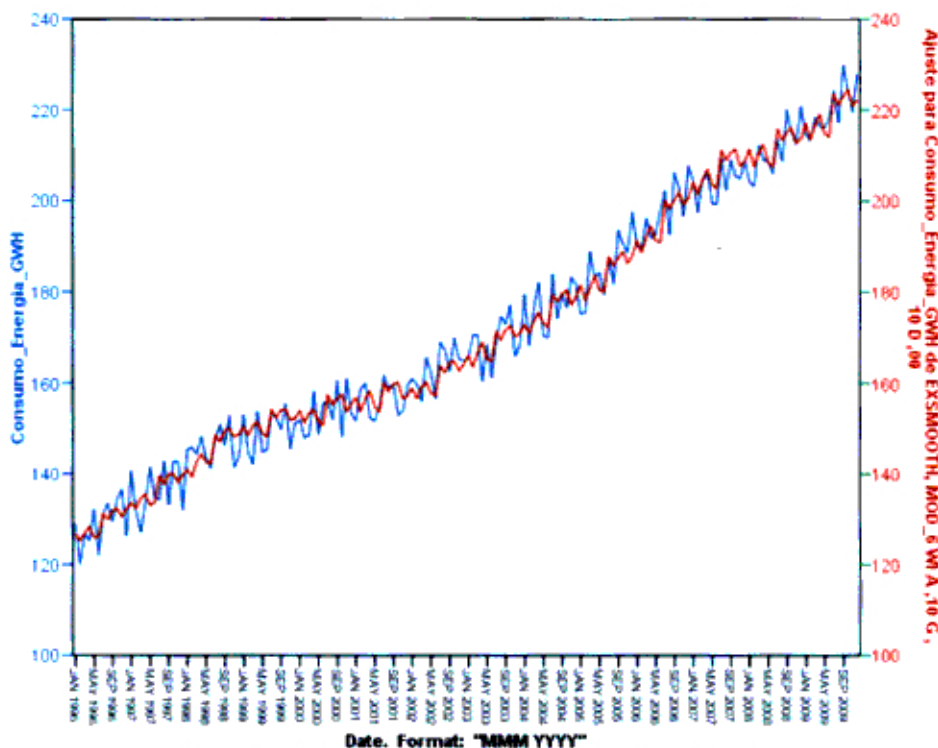


Fig. 3.9 Resultado suavizado exponencial vs. Data histórica

## CASO B: "Valores por defecto"

Se obtuvieron los siguientes resultados:

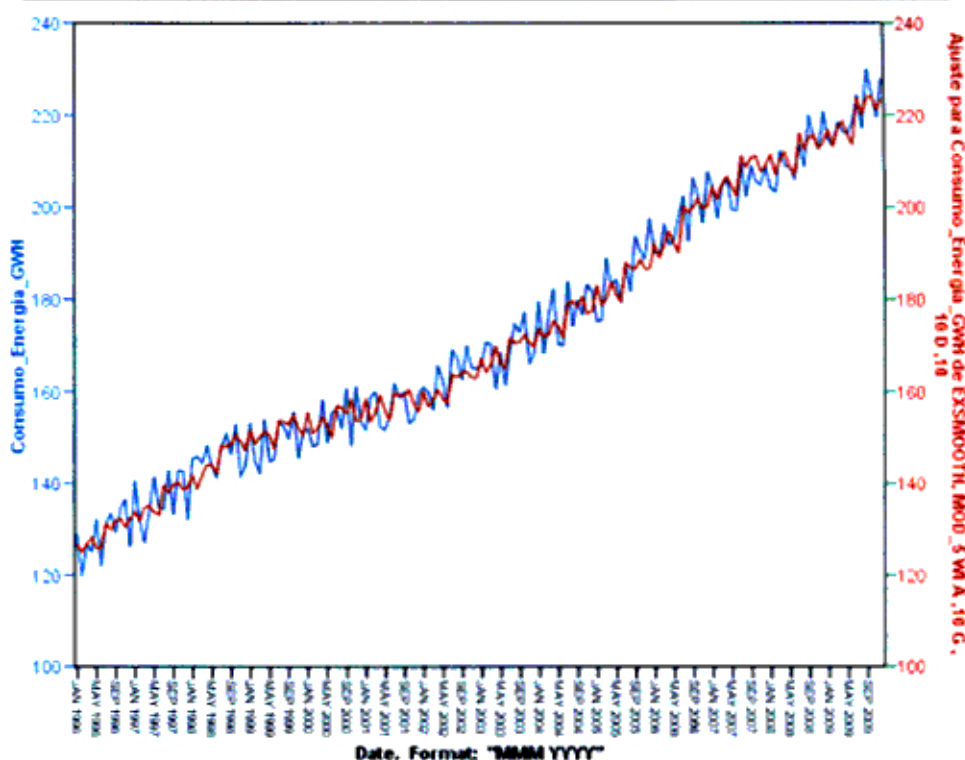
### Descripción del modelo

Nombre del modelo	MOD_6	
Serie	1	Consumo_Energia_GWH
Modelo multiplicativo de Winters	Tendencia	Lineal
	Estacionalidad	Multiplicativo
Longitud del periodo estacional		12

Aplicando las especificaciones del modelo de MOD\_6

### Parámetros del suavizado

Serie	Alpha (Nivel)	Gamma (Tendencia)	Delta (Estación)	Sumas de los errores cuadráticos	gl error
Consumo_Energia_GWH	,10000	,10000	,10000	3116,43420	155



De lo expuesto, se tiene que para la serie de consumo de energía residencial analizada, el método que representa mejor la variable analizada es el suavizado WINTER "Búsqueda por rejilla" con la menor suma de errores cuadráticos, datos que servirán para efectuar la comparación respecto del modelamiento realizado con la metodología ARIMA y demostración de la Hipótesis Nula.

A continuación, en cuadro resumen se muestran los indicadores estadísticos como resultado de aplicar todas las técnicas de suavizado existentes, de donde se advierte que el suavizado Winter "Búsqueda por rejilla" muestra los menores errores estadísticos, por lo



que se concluye que esta será el suavizado que será comparado con la metodología ARIMA.

**Tabla 3.1 Resumen resultados suavizado (Variable: Consumo\_Energia\_GWH)**

Indicador	SIMPLE Valores por defecto	SIMPLE Búsqueda por rejilla	HOLT Valores por defecto	HOLT Búsqueda por rejilla	WINTER Valores por defecto	WINTER Búsqueda por rejilla
MAE	7,28	4,96	4,02	4,02	3,56	3,53
MAPE	4,51%	3,07%	2,43%	2,43%	2,16%	2,14%
MSE	102,54	45,00	22,02	22,02	18,55	17,68
SSE	17227	7560	3700	3700	3116	2971
RMS	10,13	6,71	4,69	4,69	4,31	4,21

### 3.4 RESULTADOS. METODOLOGÍA ARIMA

El análisis univariante de series temporales mediante la metodología ARIMA con ayuda del software SPSS, se basa en que una serie temporal obedece a un proceso estocástico, sea utilizado para describir y predecir el comportamiento futuro de las series energéticas. De acuerdo a las características de la serie se muestran los resultados de calcular el modelo ARIMA univariante que describe mejor su comportamiento. Es de precisar que, el procedimiento y fundamentación del cálculo predictivo se desarrolla en el Anexo N° A.7.

Como aplicación, se realiza el cálculo predictivo de una serie energética representativa típica, mediante los métodos ARIMA. Así, tenemos la serie Consumo Mensual de Energía en Lima Sur expresada en Giga Watio Hora entre los años 1996 y 2009, con predicciones para los años del 2010 al 2012, con un total de ciento sesenta y ocho observaciones. Para realizar la estimación del modelo, se seguirá el procedimiento de Box y Jenkins; en primer lugar, se identificará la variable a proyectar para luego realizar la estimación y evaluar los residuos; determinándose el orden de integración de la serie, para luego proceder a identificar su proceso generador de datos.

Como análisis previo se muestra la evolución de la serie consumo de energía residencial en Lima Sur:

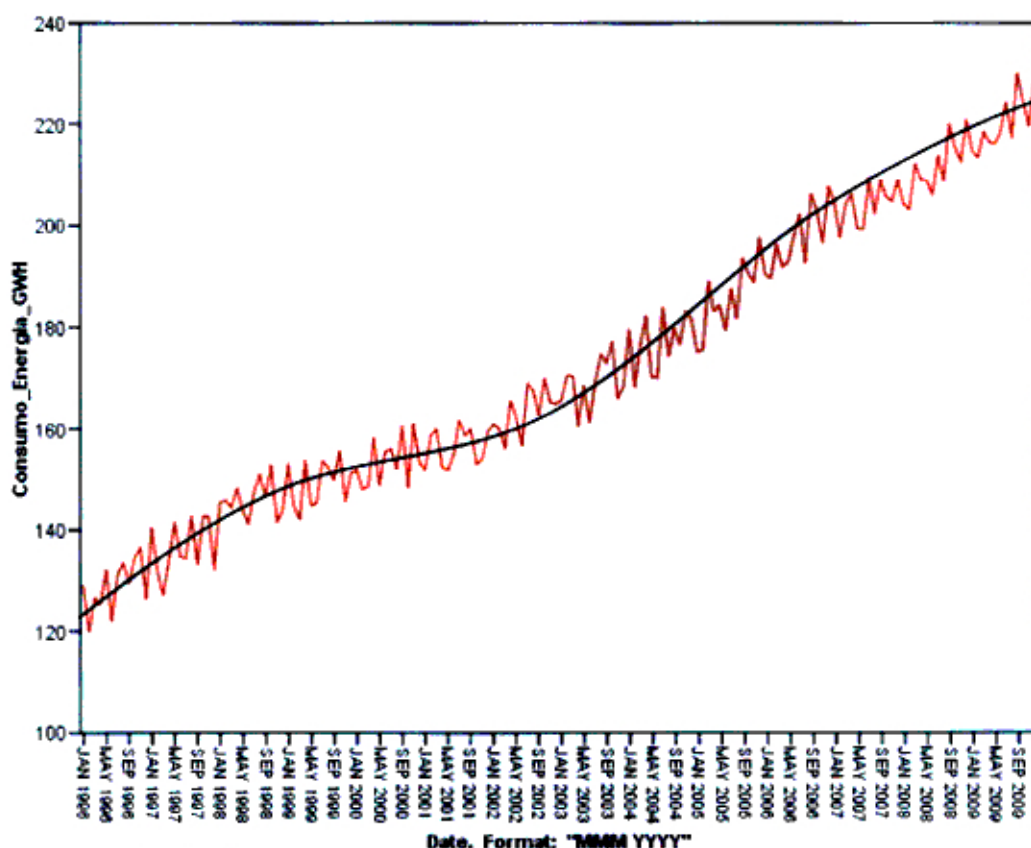


Fig. 3.11 Gráfico de la serie histórica Consumo Mensual de Energía en Lima Sur

Del gráfico anterior se puede señalar que la serie presenta tendencia creciente, asimismo se aprecia una componente estacional con amplitudes pronunciadas y un patrón de irregularidades significativo.

### 3.4.1 ANÁLISIS DE LA ESTABILIDAD EN VARIANZA

En general, cualquier serie natural a lo largo de su evolución histórica puede presentar una variabilidad no constante, esto es que la varianza sea dependiente del tiempo. En muchos casos, la variabilidad no aumenta con el tiempo sino con el nivel de la serie, por lo que, se tendrá que efectuar una transformación de la variable para estabilizar la serie en varianza.

Para identificar si la varianza es constante en una serie natural, se procede a evaluar el comportamiento variabilidad-nivel, agrupándose las observaciones de la serie por periodos de tiempo con el mismo número de observaciones. Luego dependiendo del grado de complejidad de la serie, se puede realizar una visualización gráfica o en el caso más amplio efectuar una prueba de comprobación de la hipótesis (Prueba de Levene).

Del análisis efectuado, se elaboró el cuadro resumen basándose en la media de cada caso mostrado a continuación, del cual se advierte que la serie original presenta el menor estadístico Levene con un valor de significancia o P\_valor mayor a 0.05, con una potencia

estimada de 0.596 es decir que la potencia a modelar estaría en el orden de 1/2, por lo tanto, la serie original necesita ser transformada con la función Raíz Cuadrada para tener estabilidad en varianza.

**Tabla 3.2 Cuadro resumen: Estabilidad en Varianza**

Estimación	Estadístico Levene	Inclinación	Significancia o P Valor	Transformados Potencia	Conclusión
0.596	1,108	0.404	0.356	Sin transformación (1)	
	3,372	0.084	0.000	Recíproco (-1)	
	2,203	0.058	0.012	1/raíz cuadrada (-½)	
	1,402	-0.033	0.164	Logaritmo natural (0)	
	1,032	-0.008	0.424	Raíz cuadrada (1/2)	Ok, la serie necesita esta transformación

### 3.4.2 ANÁLISIS DE ESTABILIDAD EN MEDIA REGULAR Y MEDIA ESTACIONAL

Para estabilizar la media regular de la serie en estudio puede ser necesario aplicar diferencias regulares (de orden  $d$ ) y estacionales (de orden  $D$ ). Los órdenes de diferenciación se determinan luego del análisis de estabilidad en varianza.

A continuación se efectúa la comparación por estadísticos descriptivos de las tres series creadas por diferenciación regular, estacional y combinada, obteniéndose la siguiente tabla:

**Tabla 3.3 Estadísticos descriptivos**

	N	Media	Desv. típ.
r2_Con_Ene_GWH	168	13,0437	1,07228
DIFF(r2_Con_Ene_GWH,1)	167	,0224	,26438
SDIFF(r2_Con_Ene_GWH,1,12)	156	,2679	,21190
SDIFF(dif_reg,1,12)	155	-,0016	,29854
N válido (según lista)	155		

De la tabla 3.3, se observa que la desviación típica menor corresponde a aplicar una diferencia estacional de orden 1 a la serie transformada raíz cuadrada del Consumo\_Energia\_GWH ( $Y_t$ ), por lo tanto, será tomada para el modelamiento en ARIMA con  $d=0$  y  $D=1$ . Sin embargo, las desviaciones típicas mostradas son cercanas entre sí, por lo que, a modo de comprobación se desarrollará el procedimiento de autocorrelaciones para todos los posibles casos.

### 3.4.3 IDENTIFICACIÓN DE ÓRDENES AUTOREGRESIVAS Y MEDIAS MÓVILES (p,P,q,Q)

De lo hallado en el punto anterior, los órdenes  $d$  y  $D$  son el número de veces que se tiene que diferenciar la serie previamente estabilizada en varianza hasta obtener un proceso estable en varianza y media, es decir, la serie original ya se encuentra estacionaria. El valor de los órdenes autoregresivos regular ( $p$ ) y estacional ( $P$ ), y los órdenes medias móviles

regular (q) y estacional (Q), se obtiene cuando la transformación realizada antes alcanza a comportarse como un proceso ARMA(p,q)x(P,Q). En ese sentido, un AR(p) presenta efectos directos de observaciones separadas por "1,2,...,p" retardos y los efectos directos de las "p+1,..." son nulos<sup>31</sup>.

Del análisis efectuado se puede concluir que, para seleccionar la diferencia adecuada se debe encontrar aquellos resultados en donde se note que tanto en la función de autocorrelación normal, como parcial tienen coeficientes altos o que muestren significancia en los mismos retardos. En consecuencia tenemos el modelo (La gráfica de FAC se muestra en Anexo A.8):

$$\text{ARIMA } (p,d,q)x(P,D,Q) = \text{ARIMA } (3,0,0)x(0,1,1).$$

Sin embargo, también analizaremos los modelos identificados y mostrados en cuadro a continuación, los cuales serán analizados para observar si sus raíces son significativas y si pasan las pruebas de residuos, escogiendo el modelo que consiga el mejor ajuste:

**Tabla 3.4 Modelos ARIMA identificados**

Modelo	Tipo de transformación	Modelo ARIMA (p,d,q)x(P,D,Q)	Contraste de significancia de parámetros
M1	Consumo_Energía_GWH	(1 0 0) (0 1 1) cte	Con constante. Coeficientes estables y significativos
M2	Consumo_Energía_GWH	(1 1 0) (0 1 1) cte	Con constante. Coeficientes estables y significativos
M3	Consumo_Energía_GWH	(3 0 0) (0 1 1) cte	Con constante. Coeficientes estables y significativos
M4	Consumo_Energía_GWH	(3 1 0) (0 1 1) cte	Con constante. Coeficientes estables y significativos

### 3.4.4 ESTIMACIÓN DE PARÁMETROS O COEFICIENTES

Las estimaciones y ajustes calculados, de los coeficientes autoregresivos y medias móviles regulares y estacionales, se muestran a continuación para cada modelo pre seleccionado:

#### MODELO M1: ARIMA (1 0 0) (0 1 1) con cte

$$(1-0,129xL^1) \times (1-L^{12}) \times (\text{Consumo\_Energía\_GWH}) = (1-0,505xL^{12}) \times u_t + 0,263 \dots\dots(3.1)$$

Donde:  $\Theta_1 = SMA(1) = 0.505$ ,  $\phi_1 = AR(1) = 0.129$

#### MODELO M2: ARIMA (1 1 0) (0 1 1) con cte

$$(1+0,542xL^1) \times (1-L^{12}) \times (1-L) \times (\text{Consumo\_Energía\_GWH}) = (1-0,676xL^{12}) \times u_t \dots\dots(3.2)$$

Donde:  $\Theta_1 = SMA(1) = 0.676$ ,  $\phi_1 = AR(1) = -0.542$

#### MODELO M3: ARIMA (3 0 0) (0 1 1) con cte

$$(1-0,484xL^3) \times (1-L^{12}) \times (\text{Consumo\_Energía\_GWH}) = (1-0,745xL^{12}) \times u_t + 0,263 \dots\dots(3.3)$$

<sup>31</sup> Libro "Estadística Modelos y métodos"

Donde:  $\Theta_1 = SMA(1) = 0.745$ ,  $\phi_3 = AR(3) = 0.484$

**MODELO M4: ARIMA (3 1 0) (0 1 1) con cte**

$$(1+0,220xL^3) \times (1-L^{12}) \times (1-L) \times (\text{Consumo\_Energía\_GWH}) = (1-0,758xL^{12}) \times u_t \dots (3.4)$$

Donde:  $\Theta_1 = SMA(1) = 0.758$ ,  $\phi_3 = AR(3) = -0.220$ ,  $\Phi_1 = SAR(1) = 0$

Es de precisar que, en todos los casos la significancia de la Constante es mayor al 5% establecido como parámetro máximo, por lo que se concluye que la Constante es significativa debiendo ser tomada en cuenta para el cálculo correspondiente.

**3.4.5 VALIDACIÓN DEL MODELO ARIMA ÓPTIMO POR PRUEBA DE RESIDUOS**

La validación de los mejores modelos ARIMA calculados y seleccionados consiste en comprobar si satisface la hipótesis relativa de que los residuos del modelo ( $a_t$ ) obedecen al comportamiento de ruido blanco. Para saber si las observaciones están incorreladas y procede de una distribución normal se efectua la prueba de incorrelación de residuos.

**Prueba de Incorrelación de residuos**

De los resultados dados por el software SPSS respecto de las autocorrelaciones ACF, analizadas para cada modelo se tiene que, los residuos del modelo M3 están incorrelados, los residuos caen dentro de la banda de confianza, asimismo el estadístico de Box-Ljung de la tabla correspondiente muestra que todos los valores pasan la prueba de significancia, por lo que este modelo es el óptimo.

**Hallando la ecuación de la serie de acuerdo al modelo ARIMA**

Partimos el análisis de la Ecuación General:

$$\Phi_p(L) * \phi_p(L) * (1-L)^D * (1-L)^d * X_t = \Theta_q(L) * \theta_q(L) * \mu_t + C \dots (3.5)$$

De los procedimientos antes desarrollados se tienen los siguientes datos:

- Auto regresivo:  $p = 3$ ,  $P = 0$
- Diferenciación:  $d = 0$ ,  $D = 1$
- Media Móvil:  $q = 0$ ,  $Q = 1$
- Periodo de Ciclo:  $S = 12$
- Constante  $C = 0,263$

Despejando finalmente, tendremos la ecuación de predicción de la serie:

$$X_{T+1} = 0,484 X_{T-2} + X_{T-11} - 0,484 X_{T-14} + \mu_{T+1} - 0,745 \mu_{T-11} + 0,263 \dots (3.6)$$

Donde:

$$X_T = \sqrt{\text{Consumo\_Energía\_GWH}}$$

Se comprueba que, la ecuación de predicción es la suma de dos componentes: un componente permanente o de tendencia, que es un polinomio de orden (d-1) con coeficientes que dependen del origen de la predicción, y un componente transitorio, que es una mezcla de términos cíclicos y exponenciales<sup>32</sup>.

### 3.4.6 VALIDACIÓN DE PREDICCIONES CON MODELO ARIMA ÓPTIMO POR ESTADÍSTICOS DE ERROR

La mayoría de estudios ponderan una evaluación cualitativa de los resultados, y en aquellos donde se cuantifica el análisis optan por el uso de estadísticos simples como el error cuadrático medio, el error absoluto medio, entre los principales. Para la validación presentada se ha optado por utilizar la suma del cuadrado de los errores (SSE), el error absoluto medio (MAE) y el error porcentual normalizado (MAPE), estos estadísticos permiten medir la precisión de la predicción, entendiéndose por precisión, al promedio del grado de correspondencia entre pares individuales de valores pronosticados y valores observados.

Para el presente trabajo, como parámetros de evaluación del modelo de predicción se utilizan los indicadores SSE, MAE y MAPE por cada observación, teniendo así una evolución temporal del error y poder analizar con detalle el comportamiento del modelo durante todo el periodo de predicción. La comparación de estos estadísticos entre los modelos identificados nos posibilita la evaluación de la calidad predictiva, eligiéndose aquel que minimice estos estadísticos. A continuación se muestra la comparación de los estadísticos de error de los modelos seleccionados, procedimiento de cálculo que se detallan en la Tabla A.5 y son mostrados a continuación:

**Tabla 3.5 Resumen - Error Estadístico**

Ind.	ERR_1 100_011	ERR_2 110_011	ERR_3 300_011	ERR_4 310_011
MAE	0.153	0.166	0.133	0.137
MAPE	1.19%	1.29%	1.04%	1.07%
SSE	5.93	7.24	4.63	4.75

De los cuatro modelos ARIMA seleccionados se tiene que, todos responden a un modelo ARIMA de la serie original, de los cuales el modelo M3 muestra menores valores en el

<sup>32</sup> Libro "Estadística Modelos y métodos"

cuadro de estadísticos de error, así mismo el modelo M3 es también el modelo cuyos residuos caen dentro de la banda de confianza al verificar la correlación de residuos.

Finalmente, se concluye que el modelo  $ARIMA(p,d,q)x(P,D,Q) = ARIMA(3,0,0)x(0,1,1)$  del Consumo\_Energía\_GWH incluido Constante, es el mejor modelo ARIMA para el presente análisis.

#### **3.4.7 PREDICCIÓN CON MODELO ARIMA REPRESENTATIVO**

El cálculo de las predicciones de la serie temporal Consumo\_Energía\_GWH, utilizando modelos ARIMA determinados en el punto anterior, se realiza a partir de las expresiones matemáticas obtenidas en el proceso de estimación y ajuste.

Para nuestro caso hemos seleccionado cuatro modelos, cuyos coeficientes son significativos y pueden ser buenos modelos predictivos, sin embargo, como se demostró el modelo M3 ( $ARIMA(3,0,0)x(0,1,1)$ ) cumple con las condiciones necesarias en las pruebas de residuos, por lo que, sólo desarrollaremos este modelo pues ofreció mejores resultados predictivos.

Teniendo en cuenta las características desarrolladas anteriormente, de acuerdo a las características evolutivas de la muestra, así como de la capacidad predictiva que se ha demostrado en la elección del modelo predictivo, es posible hacer predicciones en una proporción de 1 a 4 respecto al número de observaciones mensuales. Para nuestro caso se trata de una serie temporal de 168 observaciones, que tiene comportamiento ligado al clima y al desarrollo económico y social, y siendo que las predicciones calculadas corresponden a un proceso dinámico es que el modelo debe ser validado cada año, por lo que, definimos que nuestro horizonte de predicción será de tres años o 36 meses.

La capacidad predictiva del modelo ARIMA seleccionado se comprueba con el cálculo de predicciones en base a un horizonte histórico muestral con 36 observaciones menos que el horizonte histórico real de la serie Consumo\_Energía\_GWH, quiere decir que las estimaciones de los coeficientes del modelo se obtuvieron inicialmente considerando las observaciones hasta diciembre del 2009, comprobando posteriormente la capacidad predictiva sobre todo el intervalo de análisis 1996 al 2009 (llamado periodo de validación), comparando las predicciones de este modelo con las observaciones reales.

De otro lado, los valores resultados del cálculo efectuado con la modelación ARIMA se muestran en el Tabla A.5, los mismos que son mostrados gráficamente a continuación en donde se puede apreciar la aproximación calculada con el software SPSS:

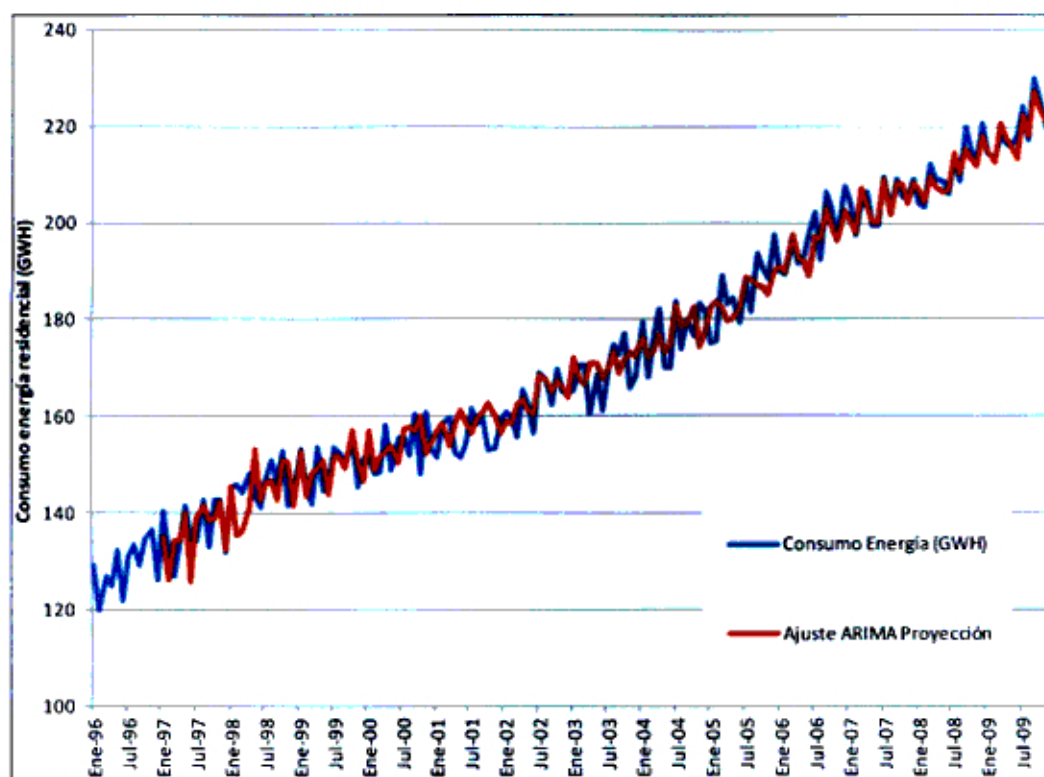


Fig. 3.12 Gráfico Consumo histórico vs. Consumo Metodología ARIMA

Del gráfico anterior, se observa que el procedimiento seguido y los valores obtenidos son los adecuados, pues la aproximación a la serie original Consumo\_Energía\_GWH es muy buena, de excelente calidad predictiva, con bajos porcentajes de error calculados.

Luego, se procede a calcular la proyección para los años 2010, 2011 y 2012, para lo cual nuevamente hay que aplicar la metodología ARIMA. Con los valores obtenidos se procede a graficar la predicción calculada, como se muestra a continuación:



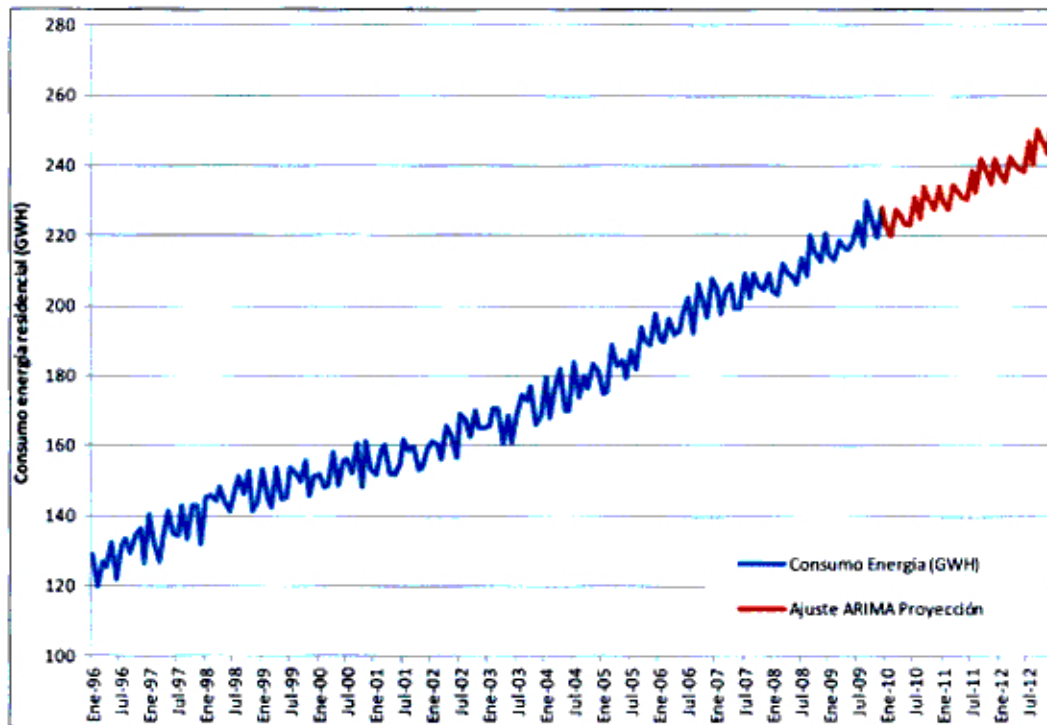


Fig. 3.13 Gráfico Consumo histórico vs. Proyección consumo ARIMA

Observamos que para el periodo predictivo calculado, éste sigue la tendencia y estacionalidad de los años anteriores, lo cual evidencia que el procedimiento seguido y los valores obtenidos son los adecuados pues la aproximación es muy buena. Finalmente, se muestra gráficamente los límites superiores e inferiores dados por el software SPSS para el periodo predictivo calculado:

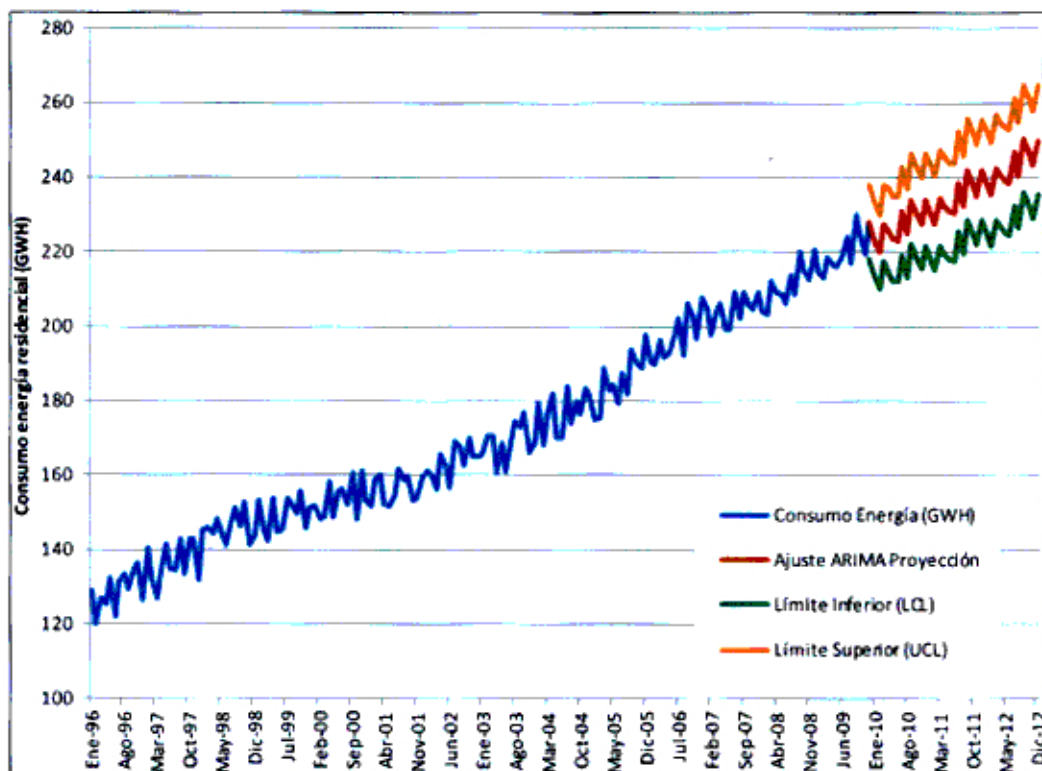


Fig. 3.14 Gráfico Consumo histórico vs. Proyección consumo ARIMA (max-min)

Del gráfico anterior se concluye que para este tipo de serie energética, los rangos entre mínimos y máximos para la proyección en periodos a corto plazo crece de tal forma, que da valores confiables para ser tomados en cuenta para decisiones de planeamiento energético en la zona de concesión, por lo que, se recomienda dar periodos predictivos de tres años bajo el procedimiento de cálculo desarrollado. Las predicciones salieron dentro de lo esperado matemática y gráficamente con valores aceptables dentro de los tres años de predicción.

De la proyección efectuada para el periodo 2010-2012 del consumo de energía residencial en la zona de Lima Sur, como resultado de la aplicación de la metodología ARIMA, se tiene que para el año 2010 se tendrá un crecimiento del 3,35%, para el año 2011 del 3,57% y para el 2012 del 3,48%; hechos que constituyen sustento técnico para concluir que el consumo de energía del sector residencial tendrá un crecimiento conservador en un escenario optimista con crecimiento anual promedio del 3,50%. A continuación se muestra los porcentajes de variación respecto del escenario elegido:

**Tabla 3.6 Resumen - Proyección crecimiento porcentual**

ESCENARIO	AÑO		
	2010	2011	2012
Pesimista	1,66%	2,93%	3,08%
Medio	2,53%	3,26%	3,29%
Optimista	3,35%	3,57%	3,48%

Es de precisar que, una vez que se valide la tasa de crecimiento estimada del 3.50% respecto el crecimiento real para el periodo 2010-2012 se podrá usar este mismo porcentaje de variación para los años 2013 y 2014; con ello se dará sustento técnico para que en un futuro trabajo similar se podrá establecer un periodo de predicción mínimo de cinco años.

## **CAPÍTULO IV**

### **DISCUSIÓN DE RESULTADOS Y CONTRASTACIÓN DE LA HIPÓTESIS**

En este capítulo, habiéndose determinado que el modelo  $ARIMA(3,0,0) \times (0,1,1)$  ofrece mejores resultados predictivos que los otros tres modelos  $ARIMA$  pre seleccionados, se establece la importancia de los resultados obtenidos y principalmente se efectúa la contrastación de los resultados obtenidos respecto de la predicción mediante alisado Winter, ya que con ello queda totalmente sustentado técnicamente que para el periodo 2010-2012 en un escenario optimista el consumo de energía tendrá un crecimiento anual promedio del 3,50%.

#### **5.1 IMPORTANCIA DE LOS RESULTADOS OBTENIDOS.**

Dada que la participación del consumo eléctrico residencial es del orden del 40% del total en la zona de concesión, para que los resultados obtenidos sean de mayor utilidad y tengan mayor sustento para orientar las inversiones en la expansión del sistema eléctrico a nivel de distribución, es aconsejable incluir también la predicción de la eléctrica industrial en baja tensión para lo cual se desarrolla a continuación un procedimiento empírico alternativo tomando como referencia los consumos de energía totales en baja tensión de toda la zona de concesión y comparándolos gráficamente con los consumos residenciales de energía. Es de precisar que, el consumo de energía total en BT corresponde a las cargas del sector residencial principalmente, incluyendo también las cargas comerciales e industriales que por su pequeña envergadura consumen energía en BT.

Como resultado se muestra el gráfico Fig. 3.14, de donde se puede concluir que ambas curvas muestran el mismo comportamiento en cuanto a tendencia y estacionalidad, con un promedio de las desviaciones porcentuales de la media del 1,86%. Por lo que, se puede concluir por extensión que la proyección de la carga total en baja tensión tendrá el mismo comportamiento que la proyección de carga residencial en cuanto a variación porcentual, es decir, que tendrá crecimiento anual promedio del 3,5% en el escenario optimista.

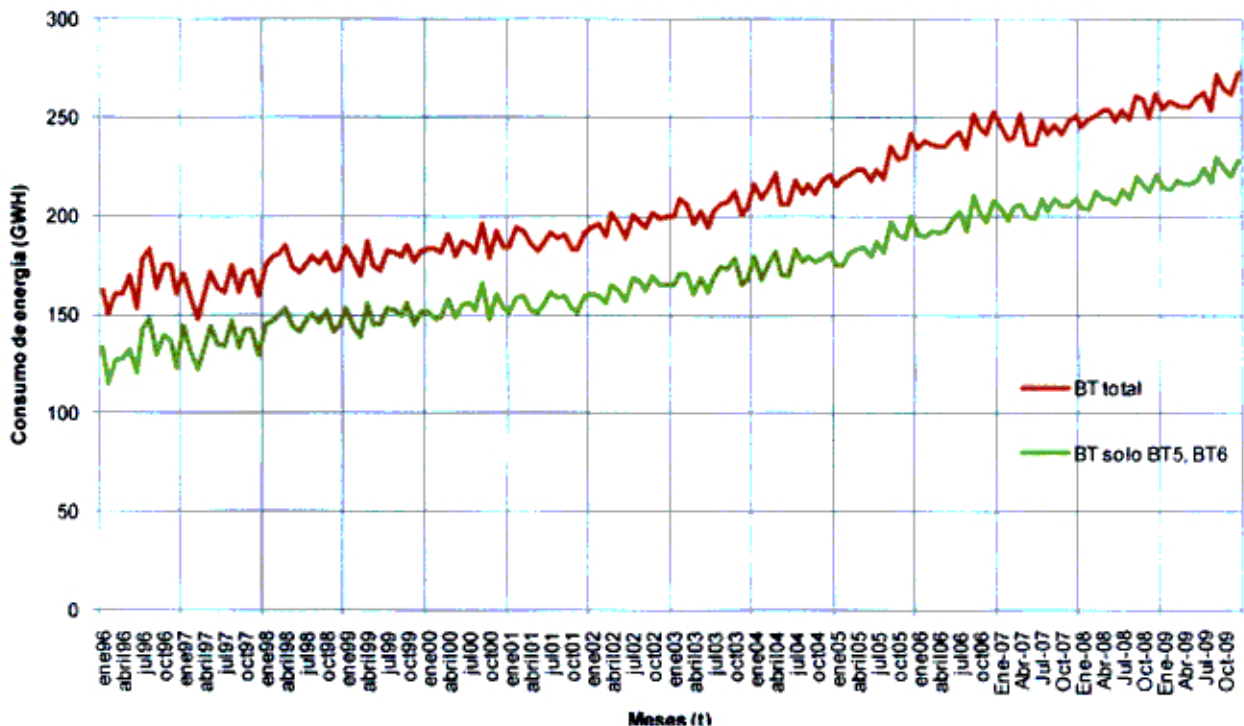


Fig. 3.14 Consumo de Energía Histórico 1996-2009 (GWH)

De otro lado, para el periodo 2007-2009 el consumo de energía en baja tensión tuvo un crecimiento real promedio del 3.32%, con crecimientos anuales de 0.82% para el 2007, 5.18% para el 2008, y 3.92% para el 2009, sin embargo, si consideramos el año 2006 el crecimiento aumenta al 3.71% para el mismo periodo. Por lo tanto, la predicción desarrollada es concordante con el crecimiento real dado en la zona de concesión como por ejemplo el boom de la construcción. La diferencia entre el crecimiento real y estimado en esta tesis, si bien es mínimo se debe principalmente a la irregularidad que presenta la serie, característica inherente a este tipo de series energéticas.

Sobre el particular, las inversiones en infraestructura eléctrica para la zona de análisis crecieron un promedio del 4.50% para el periodo 2006-2009, abarcando las inversiones en redes de media y baja tensión lo que evidencia que la expansión de las redes en BT ha sido coordinado con la MT, y que siempre debe ser coordinado como consecuencia directa que la expansión en BT debe ser respaldada por un sistema de redes y subestaciones de transmisión que tengan la capacidad de soportar el aumento de la demanda eléctrica. Por lo expuesto, se puede afirmar que el proceso de expansión de las redes en BT han seguido el crecimiento de la demanda del último decenio, sin embargo, es recomendable realizar la predicción de la demanda en media tensión tomando la incidencia de las cargas mayores.

## 5.2 CONTRASTACIÓN DE LA HIPÓTESIS.

En esta sección se muestra el análisis comparativo de la calidad predictiva de los valores

obtenidos como resultado de los modelos estimados con la metodología ARIMA y con alisado de Winter para el periodo 1996-2009, para lo cual se desarrollará la evaluación de los modelos de predicción. Para el presente trabajo, como parámetros de evaluación del modelo de predicción se han utilizado los indicadores, SSE y NAPE para cada observación, teniendo así una evolución temporal del error y poder analizar con detalle el comportamiento del modelo durante toda la predicción. El procedimiento de análisis se muestra en el Tabla A.6, cuyo resumen es mostrado a continuación:

**Tabla 3.7 Resumen – Comparación error estadístico**

Indicador	Metodología WINTER	Metodología ARIMA
MAE	3.38	2.92
MAPE	1.80%	1.59%
MSE	16.58	14.16
SSE	1592	1359

Es de precisar que, como base para el cálculo del error estadístico mostrado se tomó todos los valores históricos obtenidos desde Ene.2002 hasta Dic.2009, toda vez que la serie analizada presenta en estos años una estructura vegetativa fuerte o representativa. Asimismo, el nivel de tensión de los suministros no permanece constante en el tiempo, y que los sistemas de distribución o redes eléctricas son dinámicos, es decir, las configuraciones de las redes varían en el tiempo, ya sea por ampliaciones y/o traslados de cargas de las redes.

Así tenemos que con el método ARIMA se obtienen resultados con menor porcentaje promedio de error (1,59%) y menor suma de cuadrado de errores (1359), lo que demuestra la calidad superior de predicción en este tipo de series energéticas; toda vez que, como se observa en el anexo Gráfico A.7, la metodología ARIMA representa y se ajusta mejor a las variaciones de la serie temporal analizada. En el anexo Gráfico A.7 se muestra el análisis gráfico comparativo del consumo histórico respecto de las metodologías utilizadas en el presente trabajo, es decir, metodología ARIMA y suavizado WINTER, unificando las dos figuras antes mostradas, donde se aprecia que los resultados obtenidos con suavizado WINTER presentan mayor distorsión que los resultados ARIMA, evidenciando la superioridad gráfica de ésta última metodología.

Del análisis efectuado se puede concluir que la metodología ARIMA da mejores resultados predictivos que el método alisado de Winter, por lo tanto, la Hipótesis Nula formulada para el presente trabajo (...) *Utilizando el análisis estocástico ARIMA para modelar y predecir la energía residencial, se obtendrá mejor eficiencia y bondad predictiva respecto de los resultados obtenidos por otras técnicas determinísticas como los alisados exponenciales (...), queda demostrada.*

## **CONCLUSIONES**

1. El análisis univariante de series temporales con metodología ARIMA (Autoregressive Integrated Moving Average), debe ser utilizado para describir y predecir el comportamiento a corto plazo de las series energéticas de mayor representatividad, por lo que, debe ser tomada como herramienta para la planificación energética en cualquier zona de concesión, dando respaldo técnico a las decisiones de inversión en infraestructura eléctrica, toda vez que proporciona una alternativa altamente eficaz para describir y predecir el comportamiento futuro de dichas variables.
2. El consumo de energía eléctrica en el sector residencial o carga vegetativa, como serie energética representa dentro de la zona de concesión una variable representativa y de importancia, pues su consumo representa el 39,7% del consumo de energía total en baja tensión para la zona de análisis, siendo por ello necesario la correcta proyección de su crecimiento.
3. Del análisis comparativo efectuado se determinó que la serie Consumo\_Energía\_GWH corresponde a un modelo ARIMA (3,0,0) x (0,1,1). De cuyo desarrollo se obtiene el gráfico de modelación de la serie, donde se observa que el procedimiento seguido y los valores obtenidos tiene buena aproximación, con bajos porcentajes de error calculados.
4. De la proyección efectuada para el periodo 2010-2012 del consumo de energía residencial en la zona de Lima Sur se tiene que, para el año 2010 se tendrá un crecimiento del 3,35%, para el año 2011 del 3,57% y para el 2012 del 3,48%; hechos que constituyen sustento técnico para concluir que el consumo de energía del sector residencial tendrá crecimiento anual del 3,50% en promedio.
5. De los resultados predictivos analizados para serie Consumo\_Energía\_GWH, se demuestra la eficacia superior en el modelamiento y predicción alcanzada con la metodología ARIMA, ya que provee mejores resultados predictivos que el método Alisado de Winter, con lo cual la Hipótesis Nula formulada para el presente trabajo

queda demostrada. Habiéndose obtenido resultados con menor porcentaje promedio de error (1,59%) y menor suma de cuadrado de errores (1359), lo que demuestra la calidad superior de predicción para este tipo de series energéticas.

6. Los límites mínimos y máximos de proyección calculados en el periodo a corto plazo se desarrollan de manera uniforme, dando valores confiables para decisiones de planeamiento energético en la zona de concesión, por lo que, se recomienda realizar el cálculo predictivo cada tres años.
7. Se determinó que la proyección de la carga total en baja tensión (entiéndase residencial, comercial e industrial) tendrá el mismo comportamiento en cuanto a variación porcentual que tendrá la proyección de carga vegetativa o netamente residencial, por lo que, los resultados obtenidos son de utilidad y tienen sustento técnico para orientar las inversiones en la expansión del sistema eléctrico de distribución en baja tensión.

## **RECOMENDACIONES**

1. Desarrollar el análisis tomando como serie de tiempo toda la carga de energía en baja tensión de la zona de concesión, es decir incluyendo carga de oficinas, pequeñas fábricas y demás, a fin de demostrar que su comportamiento y resultados obtenidos son similares a los obtenidos en este trabajo.
2. Desarrollar el análisis predictivo del consumo de energía de la carga vegetativa a nivel de subestaciones de transformación (SET) en la zona de concesión, para poder diferenciar en específico zonas geográficas de mayor o menor consumo y crecimiento asociado, para obtener una mejor calidad predictiva, y con ello determinar porcentajes de crecimiento.
3. Desarrollar el modelamiento y predicción de la demanda eléctrica residencial en la misma zona de análisis mediante la utilización de modelos ARIMA con intervenciones y forma multivariante. Asimismo, desarrollar el modelamiento y predicción con modelos econométricos con variables económico-sociales, tales como, PBI per cápita, población, sectores socio-económicos, etc.
4. Realizar estudios en media tensión sobre modelamiento y predicción aplicando metodología ARIMA considerando la incidencia de las cargas mayores, con la finalidad de establecer la tasa de crecimiento de la demanda en el área de concesión y el programa de expansión de redes de MT y de BT.



## **BIBLIOGRAFÍA**

1. CARRIÓN, A. (1997)  
"Análisis de series temporales y Técnicas de Previsión"  
Departamento de Estadística e Investigación Operativa Aplicadas y Calidad.  
Universidad Politécnica de Valencia.
2. GONZALES, S. (1998)  
"Predicciones de variables energéticas mediante análisis de series temporales"  
Revista TECNIA UNI, Vol 8 N°02
3. GONZALES, S. (1999)  
"Modelos estocásticos ARIMA para predicción de variables energéticas"  
Revista TECNIA UNI, Vol 9 N°01
4. PÉREZ, C. (2001)  
"Técnicas Estadísticas con SPSS"  
Universidad Complutense de Madrid  
Editorial de Prentice Hall de Pearson Education SA
5. TINOCO, O. (2004)  
"Aplicaciones Estadísticas con SPSS"  
Juan Gutemberg Editores – Impresores
6. BOX, JENKINS, REINSEL (1976)  
"Time Series Analysis: Forecasting & Control"  
Wiley Series in probability and Statistics
7. MURILLO, TREJOS, CARBAJAL (2003)  
"Estudio del pronóstico de la demanda de Energía Eléctrica, utilizando modelos de series de tiempo"  
Publicación Scientia et Technica Año IX, No 23

8. GUTIÉRREZ, E. (2003)  
"Tesis doctoral denominada: Demanda Residencial de Energía Eléctrica en la Comunidad Autónoma de Andalucía: Un análisis cuantitativo"  
Universidad de Sevilla.
9. ORBEZO, H. (2005)  
"Trabajo sobre modelamiento y predicción de la Producción de Electricidad (GWH), utilizando la metodología estocástica ARIMA"  
Maestría en Energética – Modelos energéticos  
Universidad Nacional de Ingeniería
10. SECRETARIA DE ENERGÍA (2003)  
"Prospectiva 2002"  
República de Argentina
11. CASTRO, O. (2010)  
"Pronóstico de Demanda, Revisión quinquenal de tarifas de EJESA, Periodo 2011-2016"  
Ba Energy Solutions  
República de Argentina
12. LENTON, CH. (2010)  
"Energía Eléctrica, Proyecciones 2011"  
Business News Americas
13. UNIDAD DE PLANEACIÓN MINERO ENERGÉTICA – UPME. (2003)  
"Proyecciones de Demanda de Energía Eléctrica y Potencia Máxima, 2003-2011"  
Ministerio de Minas y Energía – República de Colombia
14. GRUPO DE DEMANDA ENERGÉTICA. (2011)  
"Proyección de Demanda de Energía Eléctrica y Potencia Máxima"  
Unidad de Planeación Minero Energética  
Ministerio de Minas y Energía – República de Colombia
15. DIRECCIÓN GENERAL DE PLANEACIÓN ENERGÉTICA. (2009)  
"Prospectiva del Sector Eléctrico 2009-2024"  
Secretaría de Energía (SENER) – República de México
16. COMISIÓN NACIONAL DE ENERGÍA. (2009)  
"Modelo de Proyección Demanda Energética Nacional de Largo Plazo"  
Ministerio de Energía – Republica de Chile

17. APEI, CONSULTORÍA Y PROYECTOS EN INGENIERÍA. (2006)  
Estudio "Evaluación del Programa ARIMA"  
Empresa privada – República del Perú
  
18. SANCHEZ DE RIVERA, D. (1994)  
Estadística Modelos y métodos  
Alianza Editorial SA, Madrid – España.
  
19. DIRECCIÓN DE ESTUDIOS Y PROMOCIÓN ELÉCTRICA. (2007)  
"Plan Referencial de Electricidad 2008-2017"  
Dirección General de Electricidad  
Ministerio de Energía y Minas – República del Perú