



**UNIVERSIDAD MICHOACANA
DE SAN NICOLÁS DE HIDALGO**



FACULTAD DE INGENIERÍA MECÁNICA

**DIVISIÓN DE ESTUDIOS DE POSGRADO DE LA
FACULTAD DE INGENIERÍA MECÁNICA.**

**MAESTRÍA EN CIENCIAS EN INGENIERÍA
MECÁNICA**

**“DISEÑO DE UNA RED NEURONAL ARTIFICIAL
PARA LA PREDICCIÓN DE LA DEMANDA
ELÉCTRICA.”**

TESIS

QUE PARA OBTENER EL GRADO DE:

MAESTRO EN CIENCIAS EN INGENIERÍA MECÁNICA

PRESENTA:

ING. RAÚL RAMOS LIRA

ASESOR:

DR. ERASMO CADENAS CALDERÓN

Morelia, Michoacán, Mayo del 2011

ÍNDICE GENERAL

Resumen.	I
Abstract.	II
Hipótesis.	III
Índice de figuras.	IV
Índice de tablas.	VI
CAPÍTULO 1: INTRODUCCIÓN AL PROYECTO	1
1.1. Resumen.	1
1.2. Introducción.	1
1.3. Revisión bibliográfica.	2
1.4. Justificación.	4
1.5. Objetivo general.	4
1.6. Objetivos específicos.	4
CAPÍTULO 2: DEMANDA DE ENERGÍA ELÉCTRICA	5
2.1. Resumen	5
2.2. Importancia del pronóstico de la demanda de energía eléctrica.	5
2.3. Factores que influyen en el pronóstico de la demanda de energía eléctrica.	6
2.3.1. Factores comunes.	6
2.3.2. Factores ocasionales.	7
2.3.3. Factores estacionales.	7
2.4. Horizonte temporal del pronóstico.	7
2.4.1. Corto plazo.	7
2.4.2. Mediano plazo.	8
2.4.3. Largo plazo.	8
2.5. Datos históricos para la elaboración del proyecto.	9

2.6. Series de tiempo.	9
2.6.1. Estacionariedad en la series de tiempo.	9
2.6.2. Estacionalidad en la series de tiempo.	10
2.6.3. Ciclicidad en la series de tiempo.	10
2.6.4. Tendencia en las series de tiempo.	10
2.7. Métodos de predicción.	11
2.7.1. Métodos discrecionales.	11
2.7.2. Métodos cuantitativos.	11
2.7.2.1. <i>Métodos de series temporales.</i>	11
2.7.2.2. <i>Métodos explicativos.</i>	12
2.7.2.3. <i>Métodos de monitoreo.</i>	12
2.7.3. Métodos tecnológicos.	12
2.8. Sumario estadístico.	14
2.8.1. Medidas de tendencia central.	14
2.8.1.1. <i>Media aritmética.</i>	14
2.8.1.2. <i>Mediana.</i>	14
2.8.1.3. <i>Moda.</i>	14
2.8.1.4. <i>Cuartiles</i>	15
2.8.2. Medidas de dispersión.	15
2.8.2.1. <i>Desviación estándar y varianza.</i>	15
2.8.2.2. <i>Intervalo total.</i>	15
2.8.2.3. <i>Error estándar de la desviación estándar.</i>	16
2.8.2.4. <i>Coefficiente de variación.</i>	16
2.8.2.5. <i>Desviación estándar relativa.</i>	16
2.8.2.6. <i>Error estándar de la media.</i>	16
2.8.3. Comentarios de los resultados de las medidas estadísticas.	18

CAPÍTULO 3: REDES NEURONALES ARTIFICIALES	21
3.1. Resumen.	21
3.2. Introducción.	21
3.2.1. Ventajas de las redes neuronales artificiales.	22
3.2.2. Aplicaciones de las redes neuronales artificiales.	23
3.3. Fundamentos de las redes neuronales artificiales.	24
3.3.1. Naturaleza bioeléctrica de la neurona.	25
3.3.2. Modelo de una red neuronal artificial.	25
3.3.2.1. <i>Unidades de proceso</i>	25
3.3.2.1. <i>Valor de activación</i>	26
3.3.2.3. <i>Función de salida o transferencia.</i>	26
3.3.2.4. <i>Conexiones entre neuronas.</i>	27
3.3.2.5. <i>Función o regla de activación.</i>	27
3.3.2.6. <i>Regla de aprendizaje.</i>	28
3.3.3. Estructura de una red neuronal artificial.	28
3.3.3.1. <i>Niveles o capas de neuronas.</i>	28
3.3.3.2. <i>Formas de conexión entre neuronas.</i>	29
3.4. Características de las redes neuronales artificiales.	30
3.4.1. Arquitectura de la red neuronal artificial.	30
3.4.2. Mecanismo de aprendizaje.	30
3.4.2.1. <i>Redes con aprendizaje supervisado.</i>	31
3.4.2.2. <i>Redes con aprendizaje no supervisado.</i>	33
3.4.3. Características de los modelos de red neuronal artificial.	35
3.5. Redes neuronales artificiales con conexiones hacia adelante.	35
3.5.1. El Perceptron.	36
3.5.1.1. <i>Regla de aprendizaje.</i>	36
3.5.2. La red ADALINE.	37
3.5.2.1. <i>Aprendizaje de la red ADALINE.</i>	38
3.5.2.1.1. <i>Regla de aprendizaje de la red ADALINE.</i>	39
3.5.3. La red Backpropagation.	40
3.5.3.1. <i>La regla delta generalizada.</i>	41
3.5.3.2. <i>Regla de aprendizaje de la red Backpropagation.</i>	43

CAPÍTULO 4: PRONÓSTICO DE LA DEMANDA DE ENERGÍA ELÉCTRICA UTILIZANDO REDES NEURONALES ARTIFICIALES.	46
4.1. Resumen.	46
4.2. Modelado de series de tiempo con redes neuronales artificiales.	46
4.2.1. Preparación de datos.	47
4.2.2. Modelado de la red neuronal artificial.	49
4.2.2.1. <i>Inicialización.</i>	49
4.2.2.1.1. Modelos de red neuronal artificial propuestos.	51
4.2.2.2. <i>Entrenamiento.</i>	53
4.2.2.3. <i>Validación.</i>	55
4.2.2.3.1. Método iterativo.	55
4.2.2.4. <i>Resultados.</i>	57
4.2.3. Análisis de resultados.	65
CONCLUSIONES.	75
BIBLIOGRAFÍA.	76

RESUMEN.

La presente tesis aborda la necesidad de conocer con menor incertidumbre el valor de la demanda de electricidad en un futuro, para ello se hace mención de: conceptos generales sobre demanda de electricidad, la técnica de Redes Neuronales Artificiales y el modelado de series de tiempo mediante dicha técnica.

Para empapar al lector en lo concerniente al pronóstico de la demanda de energía eléctrica es necesario conocer la importancia de realizar dicho pronóstico de acuerdo al horizonte temporal que se tenga contemplado, además, de tener en cuenta los factores que influyen en el comportamiento de dicha variable.

Es importante mencionar que para dicho propósito se hace uso de los datos históricos de la carga de energía eléctrica proporcionados por el organismo encargado de prestar el servicio. Por lo anterior es importante conocer conceptos básicos de las series de tiempo como la estacionariedad, estacionalidad, tendencia, ciclicidad; todo esto se muestra de manera detallada en el capítulo número 2 del presente trabajo.

Se presenta en el capítulo 3 todo lo relacionado con la técnica de Redes Neuronales Artificiales con el objetivo de conocer las ventajas, aplicaciones y características que ofrece dicha técnica.

Los modelos de red PERCEPTRON, ADALINE Y BACKPROPAGATION son utilizados comúnmente para realizar el modelado de series temporales, por lo que fue necesario conocer los algoritmos de aprendizaje de dichos modelos.

En el capítulo final se muestra detalladamente el proceso de modelado de la serie de tiempo mediante Redes Neuronales Artificiales, para ello se sigue paso a paso cada una de sus fases: identificación del problema, preparación de datos, modelado de la red neuronal y análisis de resultados.

Cabe mencionar que no se encontraron precedentes de un proyecto similar en la zona al expuesto en esta tesis, la metodología utilizada es innovadora, debido a que se hace énfasis en la fase de preparación de datos (véase capítulo 4) lo cual, representa un avance tecnológico en el pronóstico de la demanda de electricidad.

ABSTRACT

This paper deals the need to meet with less uncertainty the value of the electricity demand in the future, for this is made mention of: general concepts of electricity demand, the technique of artificial neural networks and time series modeling by this technique.

To introduce the reader regarding the forecast of electricity demand is necessary to know the importance of making this forecast according to the time horizon that has referred, in addition to taking into account the factors that influence the performance of that variable.

It is noteworthy that for this purpose makes use of historical data of electric demand, provided by the agency providing the service. Therefore, it is important to understand the basics concepts of time series such as stationary, seasonality, trend, cyclical; all this is shown in detail in Chapter 2 of this work.

It is presented in Chapter 3 all matters related to the technique of artificial neural networks, in order to know the advantages, applications and characteristics offered by this technique.

Network models PERCEPTRON, ADALINE AND BACKPROPAGATION are commonly used for modeling time series so it was necessary to know the learning algorithms of such models.

Finally, in the final chapter is shown in detail the modeling of time series using artificial neural networks, for it follows step by step each of stages, problem identification, data preparation, neural network modeling and analysis results

It is noteworthy that there were no precedents for a similar project in the area described in this paper, the methodology is innovative, which represents a technological breakthrough in forecasting electricity demand.

HIPÓTESIS

Es posible diseñar una red neuronal artificial que pronostique la demanda de energía eléctrica a corto plazo haciendo uso de la información en forma de datos históricos de demanda eléctrica con los que se cuenta y los modelos de Red Neuronal Artificial que se aplican a series de tiempo.

ÍNDICE DE FIGURAS.

Figura 2.1. Curva de la demanda de energía eléctrica de los datos utilizados en el pronóstico correspondiente al día Lunes.	11
Figura 2.2. Curva de la demanda de energía eléctrica de los datos utilizados en el pronóstico correspondiente a cada día de la semana.	19
Figura 2.2 (Continuación). Curva de la demanda de energía eléctrica de los datos utilizados en el pronóstico correspondiente a cada día de la semana.	20
Figura 3.1. Estructura de una neurona biológica típica.	24
Figura 3.2. Estructura de una red multinivel con conexión hacia adelante.	29
Figura 3.3. Conexión entre una neurona de una capa oculta con una neurona de salida.	41
Figura 4.1. Fases de modelado de Redes Neuronales Artificiales.	46
Figura 4.2. Serie de tiempo diferenciada (Domingo).	48
Figura 4.3. Serie de tiempo normalizada (Domingo).	49
Figura 4.4. Función de autocorrelación (Domingo).	50
Figura 4.5. Modelo 1 de Red Neuronal Artificial.	51
Figura 4.6. Modelo 2 de Red Neuronal Artificial.	52
Figura 4.7. Modelo 3 de Red Neuronal Artificial.	52
Figura 4.8. Pronóstico fuera de la muestra (Lunes).	56
Figura 4.9. Resultados de los modelos propuestos para cada día de la semana (Lunes).	57
Figura 4.9. Resultados de los modelos propuestos para cada día de la semana (Continuación.)	58
Figura 4.10. Resultados de los modelos propuestos para cada día de la semana (Martes).	59

Figura 4.11. Resultados de los modelos propuestos para cada día de la semana (Miércoles.)	60
Figura 4.12. Resultados de los modelos propuestos para cada día de la semana (Jueves.)	61
Figura 4.13. Resultados de los modelos propuestos para cada día de la semana (Viernes.)	62
Figura 4.14. Resultados de los modelos propuestos para cada día de la semana (Sábado.)	63
Figura 4.15. Resultados de los modelos propuestos para cada día de la semana (Domingo.)	64
Figura 4.16. Pronóstico fuera de la muestra (Martes).	65
Figura 4.17. Pronóstico fuera de la muestra (Miércoles).	66
Figura 4.18. Pronóstico fuera de la muestra (Jueves).	67
Figura 4.19. Pronóstico fuera de la muestra (Viernes).	68
Figura 4.20. Pronóstico fuera de la muestra (Sábado).	69
Figura 4.21. Pronóstico fuera de la muestra (Domingo).	70
Figura 4.2. (Continuación) Serie de tiempo diferenciada (Lunes, Martes y Miércoles).	71
Figura 4.2. (Continuación) Serie de tiempo diferenciada (Jueves, Viernes y Sábado).	72
Figura 4.3. (Continuación) Serie de tiempo normalizada (Lunes, Martes y Miércoles).	73
Figura 4.3. (Continuación) Serie de tiempo normalizada (Jueves, Viernes y Sábado).	74

ÍNDICE DE TABLAS

Tabla 2.1. Datos históricos de demanda de energía eléctrica utilizados en el modelado de series de tiempo mediante Redes Neuronales Artificiales (Valores en MW)	10
Tabla 2.2. Estadística descriptiva de la demanda de electricidad de la serie de datos utilizada en el pronóstico.	17
Tabla 3.1. Tipos de redes con aprendizaje supervisado más conocidos.	32
Tabla 3.2. Tipos de redes con aprendizaje no supervisado más importantes.	34
Tabla 3.3. Características de los modelos de red neuronal utilizados en el pronóstico.	35
Tabla 4.1. Resumen de corridas realizadas por la red neuronal a diferentes tasas de aprendizaje (Modelo 3, martes).	54
Tabla 4.2. Resultados de error de entrenamiento para cada día de la semana.	54
Tabla 4.3. Exactitud en el pronóstico generada por los modelos (Lunes).	57
Tabla 4.4. Exactitud en el pronóstico generada por los modelos (Martes).	65
Tabla 4.5. Exactitud en el pronóstico generada por los modelos (Miércoles).	66
Tabla 4.6. Exactitud en el pronóstico generada por los modelos (Jueves).	67
Tabla 4.7. Exactitud en el pronóstico generada por los modelos (Viernes).	68
Tabla 4.8. Exactitud en el pronóstico generada por los modelos (Sábado).	69
Tabla 4.9. Exactitud en el pronóstico generada por los modelos (Domingo).	70

CAPÍTULO 1: INTRODUCCIÓN AL PROYECTO

1.1. RESUMEN

En el presente capítulo se muestra un panorama general del desarrollo del presente proyecto, presenta el estado del arte del tema principal del trabajo a través de la revisión bibliográfica, además de mencionar los objetivos que se pretenden cubrir con esta tesis.

1.2. INTRODUCCIÓN

La demanda de energía eléctrica es una característica importante en cualquier sociedad industrializada, siendo de gran relevancia para el desarrollo del país, debido a que, representa un factor preponderante para que los diversos sectores de consumo se desarrollen de manera constante. De acuerdo con el Sistema Eléctrico Nacional el sector industrial y residencial absorben más de cuatro quintas partes de la energía eléctrica [1].

El objetivo primordial de cualquier compañía eléctrica es satisfacer las necesidades energéticas de sus clientes al menor costo posible, con la mayor calidad y fiabilidad que ésta pueda ofrecer. Para lograr dicho objetivo la empresa debe realizar la planeación del despacho de energía con la menor incertidumbre posible.

Tomando en cuenta que la generación de electricidad satisface la demanda instantánea de los consumidores, ya que no es posible almacenar esta energía en cantidades considerables, el pronóstico a corto plazo adquiere mayor relevancia, por lo cual, la compañía que suministra la electricidad debe desarrollar técnicas de previsión de la demanda eléctrica que permitan dar certidumbre al despacho energético.

En este sentido, la intención del trabajo propuesto en el presente documento, es la de contribuir a la ampliación del abanico de técnicas utilizadas para realizar la predicción de la demanda de energía eléctrica en México, permitiendo realizar el despacho energético de manera eficiente.

Los modelos son generados a través de la técnica de Redes Neuronales Artificiales, técnica destacada y utilizada ampliamente en la actualidad para el pronóstico de datos presentados en forma de series temporales, los resultados obtenidos son satisfactorios y la metodología utilizada para la generación de dichos modelos puede ser aplicada en cualquier zona del país garantizando soluciones a dicho problema.

1.3. REVISIÓN BIBLIOGRÁFICA

Un proyecto caracterizado desde una perspectiva diferente es el realizado por G. Mbamalu [2], donde, realiza una descomposición de la demanda eléctrica por sector de consumo; residencial, comercial, municipal e industrial. Para cada uno de estos sectores se consideran sus características operativas y se propone una adecuada estructura de Red Neuronal Artificial. Adicionalmente, la demanda de electricidad para cada sector se divide en cuatro bloques de seis horas con el objetivo de marginar el efecto de las variables meteorológicas en el desempeño de la red neuronal.

M. Kazeminejad [3], en su artículo publicado en 2006 realiza una previsión a corto plazo de la demanda de energía eléctrica utilizando el modelo de red neuronal de Perceptron multicapa. Propone una estructura de red de tres capas, considerando como vectores de entrada los valores de demanda de: el primer día de la semana, los días laborables, fines de semana, además, de incluir los valores de temperatura de tres distintas ciudades de Irán, que representan diversas condiciones climáticas como frío, calor y templado. Los resultados obtenidos son interesantes ya que, maneja un porcentaje de error medio absoluto en el pronóstico menor al error que prevé la compañía que suministra el servicio de electricidad.

Otro estudio similar al propuesto en este proyecto es el realizado por M. Hayati [4] para la zona oeste de Irán, mediante el uso de tres modelos diferentes de red neuronal; Perceptron multicapa, Red Neuronal Recurrente de Elman y Red de base radial. En su modelado utiliza diversas variables como vectores de entrada dentro de las que se tienen: temperatura mínima, temperatura máxima, temperatura promedio, humedad, velocidad del viento, día de la semana y estación del año, además, de introducir los valores de demanda de las 24 horas previas, obteniendo excelentes resultados de acuerdo con el análisis cuantitativo presentado.

En su artículo W. Día [5], presenta una forma de tratar la predicción de series temporales a través de la técnica de Redes Neuronales Artificiales, utilizando como vector de entrada diferentes variables como días laborales, fines de semana, temperatura máxima y mínima, utilizando el modelo de red de Retropropagación con una estructura de tres capas. No realiza un análisis cuantitativo de los resultados, sin embargo, se realiza un análisis cualitativo en donde se muestran resultados satisfactorios.

S. Mishra [6], realiza el pronóstico a corto plazo de la demanda de electricidad combinando diversas técnicas como optimización de enjambre de partículas, algoritmos genéticos y sistemas inmunes artificiales con el modelo de Red Neuronal Artificial Backpropagation (Retropropagación). Sus resultados son interesantes ya que, maneja un promedio de error medio absoluto de alrededor del 5%.

A. Jain [7], muestra una propuesta interesante para la previsión de la demanda de electricidad a corto plazo utilizando Redes Neuronales Artificiales, utiliza el modelo de red Backpropagation, con una estructura que utiliza como datos de entrada los registros de demanda y temperatura de dos años. El estudio toma un conjunto de dichos datos para realizar el entrenamiento de la red y otro tanto para la prueba de esta. Realiza una comparación entre resultados mediante un entrenamiento simple y los resultados que se obtienen al agrupar los datos mediante un valor de umbral determinado entre la carga media diaria del patrón de prueba la carga promedio del patrón de entrenamiento para cada día.

En su artículo A.P. Rewagad [8], menciona el uso de una Red Neuronal Artificial para el pronóstico de la demanda de energía eléctrica utilizando solamente datos históricos de la variable. Dentro de la preparación de datos menciona la transformación de las series temporales mediante un normalizado. La estructura de los modelos de red utilizados radica en veinticuatro nodos de entrada, dos capas intermedias y un objetivo de salida. Realiza su estudio con la ayuda del algoritmo de retropropagación. Al final muestra de manera cualitativa y cuantitativa los resultados.

Una propuesta diferente es la utilizada por J.K. Mandal [9], en su estudio, en donde aplica un modelo Recurrente de Red Neuronal utilizando conexiones entre las neuronas estrictamente hacia adelante, cabe mencionar que este tipo de configuración de red puede utilizar conexiones hacia atrás entre dichas neuronas. Hace uso de datos históricos medidos de manera horaria de todo un año, observando la correlación que existe entre la serie y sus retrasos para conocer los valores en la capa de entrada del modelo, obteniendo resultados satisfactorios de acuerdo con las medidas estadísticas de error expuestas.

Yu-Jun-He [10], presenta en su artículo una combinación interesante de técnicas, ya que, utiliza la clasificación borrosa y el modelo de Red Neuronal Recurrente de Elman para realizar la previsión de la demanda de electricidad. Los resultados son comprados con una predicción realizada mediante la utilización del algoritmo de aprendizaje backpropagation, donde, su propuesta se impone de acuerdo a los valores de error mostrados en el estudio.

J. Murillo [11], en su artículo publicado en 2003, presenta un estudio del pronóstico de la demanda de energía eléctrica utilizando la metodología de Box-Jenkins, realiza un análisis de la serie temporal y con base en esto identifica un modelo autorregresivo integrado con medias móviles (ARIMA) adecuado para dichos datos. Los resultados de la predicción se muestran de manera cuantitativa, mientras que los valores de error se presentan de forma gráfica.

Finalmente, mencionar que los estudios analizados anteriormente son solo algunos de los diversos trabajos relacionados de alguna manera con el tema a desarrollar, es importante mencionar que dichos artículos tienen gran relación con este proyecto, ya que utilizan técnicas similares a la propuesta, sin embargo, existen diferencias sustanciales (Véase capítulo 4) que conllevan a la aportación en el modelado de series de tiempo mediante Redes Neuronales Artificiales.

1.4 JUSTIFICACIÓN

La predicción de la demanda eléctrica es una actividad muy importante para el organismo encargado de proveer de energía eléctrica al país. Ya que si las predicciones resultan ser demasiado bajas o elevadas, los costos de la energía se incrementan considerablemente. Es difícil que en ambos casos el usuario final no resulte perjudicado. Si se produce una carencia de energía eléctrica el usuario tendrá que pagar más por el servicio, en cambio, si existe una abundancia de electricidad los costos asociados con deshabilitar algunas plantas de potencia se verán reflejados en el costo final de esta.

1.5 OBJETIVO GENERAL

Diseñar una Red Neuronal Artificial para pronosticar la demanda de energía eléctrica que trabaje a partir de los datos históricos de dicha variable.

1.6 OBJETIVOS ESPECÍFICOS

1. Contribuir a la integración de una base tecnológica suficiente para la predicción de la demanda de energía eléctrica.
2. Generar un precedente en el modelado de series de tiempo para la predicción de la demanda de energía eléctrica, que pueda ser utilizado como referencia en trabajos de modelado en otras disciplinas similares.

CAPITULO 2. DEMANDA DE ENERGÍA ELÉCTRICA.

2.1. RESUMEN.

En el presente capítulo se muestra la relevancia de realizar una previsión adecuada de la demanda de energía eléctrica, los principales factores que influyen para llevar a cabo esta tarea a corto, mediano o largo plazo, de acuerdo con el alcance que se desee tener. Además, se muestran los datos históricos que serán utilizados para realizar la previsión a corto plazo de la demanda de electricidad, los cuales son proporcionados por la empresa que suministra el servicio de energía eléctrica. Finalmente, se realiza un análisis estadístico de los datos, lo que representa una descripción de cada una de las muestras obtenidas para realizar el pronóstico a corto plazo.

2.2. IMPORTANCIA DEL PRONÓSTICO DE LA DEMANDA DE ENERGÍA ELÉCTRICA.

La importancia de la predicción de la demanda de energía eléctrica nace de la incertidumbre relacionada con el conocer de manera puntual el valor de ésta en un horizonte de tiempo particular, ya sea este a corto, mediano o largo plazo, dependiendo del alcance que la compañía que suministra el servicio desee tener; debido a que, un pronóstico a corto plazo influye de manera directa en el despacho y distribución de electricidad, mientras que un pronóstico a mediano plazo es útil para la planeación del mantenimiento de las plantas de potencia principalmente, una predicción a largo plazo marca una pauta importante para la ampliación, rehabilitación o construcción de una planta de potencia.

No cabe duda que un pronóstico correcto de la demanda de energía eléctrica es un requisito fundamental para que la compañía que suministra el servicio logre las metas previstas de fiabilidad y calidad que se ha planteado, ya que la creciente dependencia de la electricidad aumenta los inconvenientes causados a los usuarios si se producen carencias en el despacho de la energía eléctrica.

En la industria eléctrica, minimizar la incertidumbre de conocer el valor de la demanda eléctrica es primordial para realizar y ejecutar eficientemente programas de planeación para afrontar la demanda de energía eléctrica en un futuro. Es por esto que acertar de manera significativa las predicciones reduce los riesgos que la compañía tomará en cuanto a inversiones innecesarias y despacho eficiente de la electricidad.

2.3 FACTORES QUE INFLUYEN EN EL PRONÓSTICO DE LA DEMANDA DE ENERGÍA ELÉCTRICA.

Existen diversos factores de considerable importancia que influyen en el pronóstico de la demanda de energía eléctrica. La Comisión Federal de Electricidad (CFE) realiza un estudio anual de la demanda de energía por sector de consumo [1]. Así, en el sector industrial se pueden considerar factores importantes, el nivel de empleo en este sector, el nivel de producción industrial, la normativa medio ambiental, etc.

En el sector residencial son factores relevantes el número de personas por vivienda, la utilización de aire acondicionado, la cantidad de electrodomésticos. La demanda de energía eléctrica puede modificarse según la hora del día, pues se consume más energía durante las horas iniciales del día antes de la incorporación de las personas a la jornada laboral, reduciéndose durante el transcurso de la misma e incrementándose el consumo al finalizar la jornada laboral.

Los sectores comercial y de servicios presentan factores tales como: las dimensiones del lugar de trabajo, el nivel de empleo en estos sectores, los niveles de consumo de la población, etc.

Por lo anterior, se puede realizar un estudio de la demanda de electricidad de manera sectorial; seleccionando minuciosamente dichas variables para posteriormente realizar un análisis de correlación y de contribución al pronóstico.

Los factores antes mencionados afectan de manera puntual los sectores de consumo de energía; los cuales a su vez se ven influenciados por factores como: el ritmo de la actividad económica, el crecimiento demográfico, las condiciones climáticas y geográficas, la estructura y niveles tarifarios; además, cambios tecnológicos tales como avances en la eficiencia con la que se utiliza la electricidad en procesos productivos y en los aparatos eléctricos [1]; Estos factores plantean una visión general del problema, por lo cual es factible realizar un desglose de estos aspectos.

2.3.1 FACTORES COMUNES

Las condiciones climáticas y geográficas y el crecimiento demográfico son variables que influyen comúnmente sobre el patrón de la demanda de electricidad. Para ejemplificar estas variables, se tiene: la temperatura, la velocidad del viento, la radiación solar, la humedad relativa, las nevadas, el índice de comodidad, el crecimiento de la población, el producto interno bruto y el número de consumidores conectados [13, 14, 15].

2.3.2 FACTORES OCASIONALES

Existen otros tipos de variables las cuales ocasionalmente se presentan o influyen de manera poco común en la demanda de energía eléctrica; los días de fiesta nacional, los festivales, catástrofes, eventos deportivos, eventos religiosos, huelgas, etc. son ejemplos claros de este tipo de factores [13, 14, 15].

2.3.3 FACTORES ESTACIONALES

La demanda de energía eléctrica puede ser analizada desde la perspectiva estacional. El consumo de energía eléctrica se modifica con respecto a la estación del año o a la hora del día; con las estaciones del año, debido a que en invierno y verano se muestra un incremento de la demanda, mientras que en primavera y otoño se presenta mayor estabilidad en el consumo. En lo que refiere a la hora del día o los días de la semana la demanda de electricidad se ve afectada de acuerdo al día de la semana considerado, ya que la demanda suele ser mayor durante los días laborables que durante los fines de semana o días festivos[13, 14, 15].

2.4 HORIZONTE TEMPORAL DEL PRONÓSTICO.

Realizar un estudio sobre el pronóstico de la demanda de energía eléctrica debe responder a distintas cuestiones relacionadas con la fiabilidad, robustez, computabilidad y optimalidad. Para generar certidumbre a estas interrogantes es necesario conocer el horizonte temporal al que va a ir enfocado el análisis.

2.4.1 CORTO PLAZO

El pronóstico a corto plazo consiste en proporcionar la demanda de energía sobre un intervalo de horas, días y hasta una semana en el futuro, en este lapso el principal objetivo a cubrir por la empresa que suministra el servicio (CFE) es la distribución, de la manera más económica posible, entre los distintos recursos de generación con los que cuente la empresa; además debe ocuparse de la operación correcta de las unidades de generación.

Como ya se ha mencionado la demanda de electricidad se ve afectada por diversos factores, en este sentido, el principal factor que se presenta en este horizonte temporal es el de las condiciones climáticas; es decir, la variación de carga está relacionada principalmente con la variación de la temperatura y la radiación solar en periodos relacionados a la misma época en años recientes. Al realizar una previsión a corto plazo de la demanda de electricidad la exactitud del pronóstico es fundamental, ya que, la puesta en marcha o mantener los generadores en estado de espera depende de este.

2.4.2 MEDIANO PLAZO

Realizar una previsión de la carga de energía a mediano plazo se considera de gran importancia para la empresa con relación a establecer el calendario de mantenimiento de las distintas unidades generadoras, planear el consumo de las plantas de generación, prever los recursos energéticos para dar certidumbre a la demanda futura, así como, el mantenimiento a la red de transmisión. El pronóstico a mediano plazo consiste en proporcionar información en lapsos que comprenden desde semanas, meses e incluso hasta cinco años [16].

Los factores más importantes que intervienen en las estimaciones de la demanda de energía eléctrica en un horizonte temporal de mediano plazo son más variados que en la previsión a corto plazo, ya que, no solo las condiciones climáticas son importantes, en este sentido, las condiciones socioeconómicas se integran al desarrollo del análisis. Generalmente la serie de tiempo en estos periodos muestra cierta ciclicidad, modificándose únicamente los picos de la demanda, debido principalmente al crecimiento de la población y por ende existe una tendencia positiva en la curva que presenta la demanda de electricidad.

2.4.3 LARGO PLAZO

El pronóstico a largo plazo se utiliza principalmente en la planificación de recursos (capital, recursos humanos y tecnología), nuevos productos y servicios. La tarea primordial de la predicción es pronosticar las tendencias a largo plazo y la necesidad de nuevos productos o servicios en caso de que no cambien los patrones o relaciones establecidos. El reto más grande es explicar el impacto de los cambios en la tecnología, la competencia, las materias primas, el precio de los energéticos, la población y la demografía, las actitudes y necesidades del consumidor. Tales cambios pueden tener un impacto duradero en las tendencias o relaciones establecidas y se necesita tomarlos en consideración en la planificación a largo plazo [16]. Desde una perspectiva generalizada, lo anterior, indica los lineamientos a seguir para realizar una previsión correcta; para enfocarlo en relación a la predicción de la demanda de electricidad los factores que principalmente influyen son: el crecimiento demográfico, el ritmo de la actividad económica, el producto interno bruto, los niveles tarifarios; además de los factores mencionados en los previos horizontes temporales.

La previsión a largo plazo para la empresa suministradora del servicio de electricidad suele abarcar periodos de entre cinco y veinticinco años; se utiliza para planificar la construcción o ampliación de plantas generadoras y redes de transmisión. Realizar un pronóstico a largo plazo de la demanda de energía de manera correcta permitirá anticipar la capacidad de generación y la combinación de potencia generadora con la que debe contar un nuevo parque generador.

2.5 DATOS HISTÓRICOS PARA LA ELABORACIÓN DEL PROYECTO.

Los datos históricos que fueron utilizados para la creación de este proyecto fueron proporcionados por la CFE, a través de la Gerencia de Proyectos Geotermoeléctricos, con sede en Morelia, Michoacán. Para la generación de pronósticos a corto plazo y realización de la estadística descriptiva se contó con las mediciones horarias de demanda de energía eléctrica del mes de noviembre del año 2009.

El documento está configurado con la fecha, hora del día y magnitud de demanda de energía eléctrica consumida, por lo que se puede generar la serie de tiempo que representa la zona en forma horaria (Tabla 2.1).

2.6 SERIES DE TIEMPO

Una serie de tiempo es el resultado de observar los valores que registra una variable a lo largo del tiempo en periodos equidistantes (cada hora, cada día, cada mes, cada año, etc.). Desarrollar modelos de previsión, ayuda a estimar, cómo la secuencia de registros continuará en un determinado horizonte temporal.

Las series temporales manifiestan ciertos patrones los cuales ayudan al pronosticador a describirla. Los elementos que se presentan son los siguientes: el horizontal, el estacional, el cíclico y el tendencial.

2.6.1 ESTACIONARIEDAD EN LAS SERIES DE TIEMPO

Una serie de tiempo manifiesta estacionariedad o un patrón horizontal cuando no hay tendencia alguna en los datos, es decir, crecimiento o decremento de estos. Cuando existe este patrón, generalmente se hace referencia a la serie como estacionaria. En otras palabras, los datos oscilan alrededor de una media constante, independiente del tiempo, y la varianza es aproximadamente constante [15, 16, 17].

Para evaluar la estacionariedad en una serie temporal regularmente se utilizan las técnicas gráficas, y observando lo siguiente:

- a) Si la serie no muestra evidencia de un cambio en la media en el tiempo, entonces se dice que la serie es estacionaria en la media.
- b) Si la serie no muestra un cambio significativo en la varianza con respecto del tiempo, entonces se dice que la serie es estacionaria en la varianza

Tabla 2.1. Datos históricos de demanda de energía eléctrica utilizados en el modelado de series de tiempo mediante Redes Neuronales Artificiales (Valores en MW).

DÍA	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	
1	DOM	5186	4924	4816	4761	4704	4753	4727	4588	4706	4890	5069	5103	5136	5158	5083	5032	5074	5194	5761	6250	6165	5948	5653	5146
2	LUN	4870	4537	4532	4439	4572	4622	4838	5078	5470	5680	5850	5896	5976	6014	5896	5838	5794	5689	5920	6324	6243	5940	5688	5482
3	MAR	5125	4931	4850	4876	4918	5012	5442	5737	5901	6206	6301	6537	6647	6743	6748	6647	6583	6431	6585	6920	6714	6370	6128	5865
4	MIÉ	5566	5316	5264	5172	5310	5410	5821	6032	6246	6479	6646	6710	6780	6749	6699	6740	6672	6584	6593	6967	6812	6408	6144	5957
5	JUE	5557	5409	5360	5305	5284	5454	5712	6009	6358	6560	6733	6654	6693	6883	6836	6700	6764	6597	6753	6956	6784	6408	6130	5929
6	VIE	5602	5459	5381	5205	5271	5458	5862	5992	6191	6651	6559	6588	6746	6673	6709	6745	6695	6695	6732	6987	6809	6415	6265	6054
7	SÁB	5677	5465	5312	5320	5263	5319	5531	5420	5853	6184	6292	6321	6524	6443	6269	6153	6039	6094	6514	6563	6257	6232	6145	5683
8	DOM	5394	5163	5000	4889	4895	4913	4858	4843	4935	5105	5206	5354	5275	5392	5190	5259	5278	5400	6064	6577	6495	6304	5832	5311
9	LUN	4949	4819	4753	4694	4681	4932	5361	5611	5808	6039	6089	6263	6655	6654	6596	6716	6425	6409	6821	7008	6840	6282	6099	6035
10	MAR	5702	5451	5367	5337	5306	5472	5609	5854	5982	6329	6505	6475	6562	6634	6531	6738	6572	6499	6971	7143	6928	6544	6348	6117
11	MIÉ	5640	5290	5240	5026	5217	5442	5808	6022	6361	6552	6642	6671	6669	6779	6650	6660	6745	6625	6852	7077	6879	6514	6204	5942
12	JUE	5562	5278	5218	5137	5177	5328	5753	5994	6336	6513	6820	6737	6659	6720	6785	6769	6705	6699	6903	7126	6939	6638	6401	6222
13	VIE	5695	5398	5271	5188	5210	5566	5938	6190	6402	6682	6782	6888	6935	6641	6727	6824	6662	6692	6914	7059	6844	6438	6327	6148
14	SÁB	5665	5431	5323	5255	5272	5473	5564	5541	5911	6265	6379	6393	6397	6303	6150	5898	5809	5767	6374	6451	6008	5996	6003	5622
15	DOM	5493	5230	5000	4825	4884	4962	4878	4815	5189	5400	5623	5735	5627	5477	5429	5364	5411	5570	6155	6594	6617	6428	5937	5464
16	LUN	4972	4756	4627	4602	4588	4587	4681	4748	5241	5625	5842	5851	5977	5967	5804	5682	5743	5789	6583	6972	6827	6442	6109	5498
17	MAR	5028	4857	4712	4694	4699	4909	5503	5786	5953	6313	6370	6596	6759	6692	6589	6643	6686	6729	7086	7167	6912	6572	6174	5760
18	MIÉ	5315	5036	4902	4851	4861	5172	5564	5810	6057	6443	6575	6581	6565	6600	6568	6536	6520	6421	6816	7096	6840	6508	6137	5729
19	JUE	5434	5181	5029	5038	5094	5235	5617	6079	6306	6564	6740	6921	6925	6865	6767	6687	6636	6526	6897	7125	6889	6521	6297	5949
20	VIE	5503	5387	5249	5204	5261	5456	5880	6123	6356	6536	6553	6511	6707	6988	6722	6627	6593	6480	6691	6904	6760	6396	6237	6066
21	SÁB	5674	5447	5359	5259	5235	5383	5610	5616	5960	6251	6398	6516	6503	6460	6328	6128	6161	6135	6777	6772	6346	6218	6135	5796
22	DOM	5290	4996	4964	4847	4806	4913	4888	4820	4990	5263	5275	5353	5385	5410	5306	5272	5255	5279	6027	6553	6500	6320	5949	5326
23	LUN	4925	4806	4743	4637	4736	4891	5322	5758	6042	6437	6547	6604	6721	6700	6583	6499	6624	6519	6790	7019	6837	6513	6189	5962
24	MAR	5447	5379	5234	5190	5276	5437	5807	5896	6224	6454	6501	6806	6896	6819	6734	6813	6791	6525	6725	7063	6879	6692	6381	5895
25	MIÉ	5330	5329	5213	5247	5265	5469	5694	6032	6366	6548	6670	6737	6813	6772	6743	6709	6665	6623	6940	7029	6873	6528	6263	5943
26	JUE	5516	5371	5245	5139	5227	5408	5844	5975	6257	6632	6680	6771	6851	6792	6695	6597	6736	6788	7057	7084	6915	6506	6267	6004
27	VIE	5549	5254	5243	5211	5274	5463	5930	6181	6500	6800	6827	6835	6797	6882	6808	6693	6805	6847	6919	7060	6873	6534	6339	6061
28	SÁB	5681	5487	5345	5311	5326	5442	5662	5848	6037	6416	5954	6021	6437	6283	6051	5921	5987	5954	6610	6844	6562	6430	6349	5854
29	DOM	5369	5132	5001	4928	4893	4907	4708	4873	5217	5157	5318	5435	5350	5446	5281	5365	5357	5083	5942	6600	6794	6561	5989	5443
30	LUN	5066	4896	4773	4817	4772	5018	5504	5909	6189	6491	6683	6763	6751	6835	6799	6782	6814	6803	7040	7133	7000	6608	6340	6068

2.6.2 ESTACIONALIDAD EN LAS SERIES DE TIEMPO

Una serie temporal muestra un patrón estacional cuando varía en intervalos de tiempo iguales. Los intervalos de tiempo pueden ser los meses o las cuatro estaciones del año, pero también pueden ser las horas del día, los días de la semana o los días del mes. La serie de tiempo utilizada para cada día de la semana muestra un patrón estacional, el cual, se manifiesta cada 24 horas. Para ejemplificar lo antes expuesto, la Figura 2.1 muestra la serie temporal utilizada para la generación del modelo de pronóstico correspondiente al día lunes. Al final del capítulo se muestran cada una de las series temporales utilizadas para la previsión del resto de los días de la semana (Figura 2.2).

2.6.3 CICLICIDAD EN LAS SERIES DE TIEMPO

Un patrón cíclico es semejante al patrón estacional, la diferencia radica en la duración del tiempo, generalmente la duración de un ciclo único es mayor a un año. El patrón cíclico es difícil de pronosticar, debido a que no se repite en intervalos constantes de tiempo y su duración no es uniforme [15, 16, 17].

2.6.4 TENDENCIA EN LAS SERIES DE TIEMPO

Se presenta un patrón tendencial en una serie de tiempo cuando existe un incremento o decremento general del valor de la variable a lo largo del tiempo.

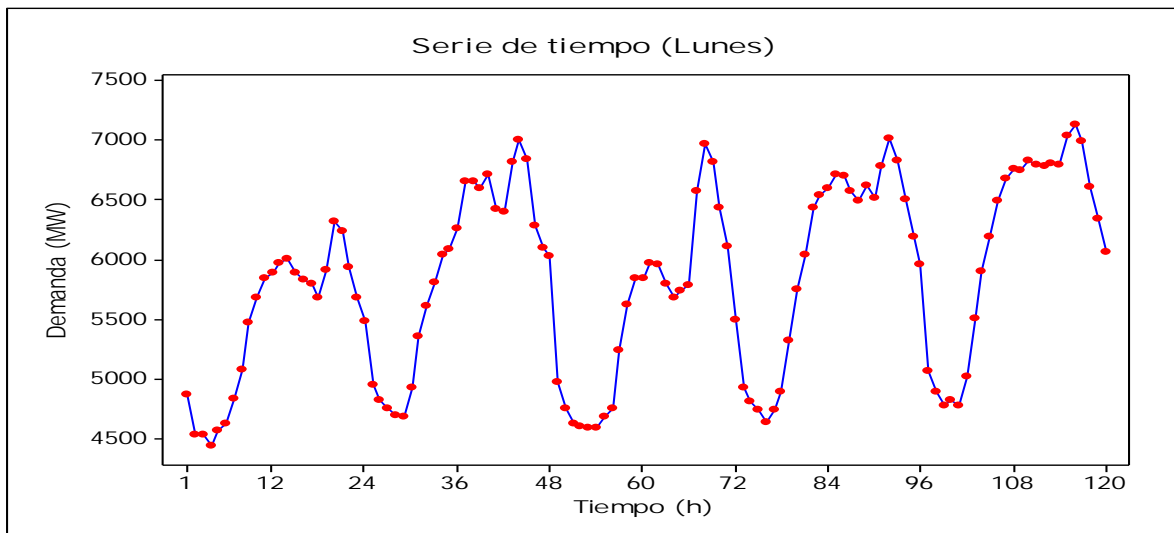


Figura 2.1. Curva de la demanda de energía eléctrica de los datos utilizados en el pronóstico correspondiente al día Lunes.

Para llevar a cabo la generación de los modelos de pronóstico en el presente trabajo, se hizo uso de las series temporales horarias correspondientes al mes de noviembre del año 2009, las cuales, fueron divididas por cada día de la semana, para la fase de entrenamiento de los modelos se utilizó el 75% de la serie mientras que, para la etapa de prueba del modelo se utilizó el restante 25%.

2.7 MÉTODOS DE PREDICCIÓN

Elaborar pronósticos es importante en un amplio campo de situaciones de planificación o toma de decisiones. Los modelos de previsión de la demanda de electricidad han evolucionado de manera notable de la década de 1970 a la actualidad. Existe una gran diversidad de técnicas de pronóstico y su aplicación depende fundamentalmente del problema que se aborde. Existen diferentes esquemas de clasificación al considerar los principales enfoques para pronosticar, el más utilizado divide a esos métodos en tres categorías: discretionales, cuantitativos y tecnológicos [16].

2.7.1 MÉTODOS DISCRECIONALES

Los métodos discretionales son los que se utilizan más comúnmente en las empresas y organizaciones gubernamentales. Los pronósticos de este tipo se hacen muy a menudo como juicios individuales o por acuerdos o decisiones de un comité.

2.7.2 MÉTODOS CUANTITATIVOS

Los métodos cuantitativos, usan la extrapolación de la tendencia, la estacionalidad, la estacionariedad, basándose en métodos estadísticos. Fundamentados en modelos probabilísticos que también usan la extrapolación, pero con la idea de minimizar el error del pronóstico. Los métodos cuantitativos a su vez, se clasifican en tres categorías: los métodos de series temporales, los métodos explicativos y los métodos de monitoreo.

2.7.2.1 *MÉTODOS DE SERIES TEMPORALES*

Los métodos de series temporales buscan identificar patrones históricos (empleando el tiempo como referencia) para posteriormente pronosticar, utilizando una extrapolación basada en el tiempo de esos patrones. La necesidad de identificar el comportamiento de la serie, hace necesario utilizar la generación de gráficas como una herramienta fundamental en la identificación de patrones que afectan a la serie.

El análisis de los datos se refiere básicamente a responder los siguientes planteamientos:

- a) ¿Existen patrones consistentes?
- b) ¿Existe tendencia significativa?
- c) ¿Existe evidencia de tendencia cíclica?
- d) ¿Existe la presencia de valores desviados?

Basándose en lo mencionado anteriormente es posible definir un espectro de técnicas adecuadas para aplicar a la serie temporal, entre las más utilizadas se encuentran las siguientes:

- 1) Modelos de regresión.
- 2) Métodos de suavizado exponencial.
- 3) Modelos ARIMA (Box-Jenkins).
- 4) Modelos no lineales (Redes Neuronales Artificiales)
- 5) Modelos multivariable.

2.7.2.2 *MÉTODOS EXPLICATIVOS*

Los métodos explicativos tratan de identificar las relaciones que conducen a resultados observados (causados) en el pasado y luego pronosticar mediante la aplicación de tales relaciones al futuro.

2.7.2.3 *MÉTODOS DE MONITOREO*

Los métodos de monitoreo, que todavía no alcanzan un uso muy extendido, buscan identificar cambios en los patrones y relaciones. Básicamente se utilizan para indicar cuándo no es apropiada la extrapolación de patrones o relaciones pasadas.

2.7.3 *MÉTODOS TECNOLÓGICOS*

Los métodos tecnológicos tienen que ver con los problemas de largo plazo de la naturaleza tecnológica, social, económica o política [16].

El presente trabajo aborda la metodología de series de tiempo utilizando la técnica de Redes Neuronales Artificiales (RNA) para la obtención de los modelos de pronóstico. Los modelos de RNA son capaces de manejar las incertidumbres que aparecen en la solución a problemas relacionados con el mundo real, ofreciendo soluciones robustas y fáciles de implementar comparado con soluciones tradicionales. El pronóstico ha resultado ser una de sus más importantes aplicaciones.

2.8 SUMARIO ESTADÍSTICO

A continuación, se presenta una descripción de las medidas estadísticas utilizadas para el análisis de los datos históricos. Los resultados se muestran en la Tabla 2.2 al final del capítulo.

2.8.1 MEDIDAS DE TENDENCIA CENTRAL

La tendencia central se refiere al punto medio de una distribución. Se conocen también como parámetros de localización [18]. Existen varias medidas de tendencia central, a continuación se describe cada una de ellas:

2.8.1.1 *MEDIA ARITMÉTICA*

La media es el parámetro más común para expresar la tendencia central en un conjunto de datos. Se recomienda calcular el valor de la media de acuerdo con la siguiente ecuación:

dónde:

Media aritmética.

Termino i -ésimo de la sumatoria.

2.8.1.2 *MEDIANA*

La mediana es el valor central de un arreglo ordenado de datos. Se utiliza comúnmente como medida descriptiva del centro cuando se desea minimizar los cálculos o eliminar los efectos de valores desviados.

Cuando las observaciones son número par, la mediana es la media aritmética de los dos valores centrales.

2.8.1.3 *MODA*

La moda es el valor observado más frecuentemente en un conjunto de datos y se puede estimar a partir de un arreglo ordenado de datos o a partir del análisis de un histograma.

2.8.1.4 *CUARTILES*

Cuando los datos han sido ordenados del menor al mayor valor, se define como el cuartil superior como el dato que representa el 25% de las medidas y el cuartil inferior como el dato que representa el 75% de las medidas, estas dan origen a otra medida de tendencia central que es el segundo cuartil, el cual se define:

2.8.2 *MEDIDAS DE DISPERSIÓN*

Para conocer la dispersión o escala de la serie de tiempo obtenida, es necesaria la utilización de los parámetros que permitan medirlas.

2.8.2.1 *DESVIACIÓN ESTÁNDAR Y VARIANZA*

La desviación estándar y la varianza son los parámetros de dispersión que se utilizan con mayor frecuencia para conocer la distancia promedio de cualquier observación de un conjunto de datos con respecto a la media de distribución.

A la varianza también se le conoce como el segundo momento:

2.8.2.2 *INTERVALO TOTAL*

El intervalo total o la dispersión total se define como la diferencia entre la observación más alta y la más baja.

2.8.2.3 *ERROR ESTÁNDAR DE LA DESVIACIÓN ESTÁNDAR*

—

Esta ecuación señala que, por lo general, a medida que se incrementa el número de las observaciones n , el error estándar de la desviación estándar s_e disminuye, significando que la desviación estándar s se ha determinado mejor que con pocas observaciones.

2.8.2.4 *COEFICIENTE DE VARIACIÓN*

El coeficiente de variación s_v se define como:

—

Esta medida representa el valor relativo de la desviación estándar con respecto a la media.

2.8.2.5 *DESVIACIÓN ESTÁNDAR RELATIVA*

2.8.2.6 *ERROR ESTÁNDAR DE LA MEDIA*

El error estándar de la media es un parámetro importante para expresar la dispersión de la media. Éste se define como:

—

Tabla 2.2. Estadística descriptiva de la demanda de electricidad de la serie de datos utilizada en el pronóstico.

DÍA/#DATOS	MEDIDAS DE TENDENCIA CENTRAL					MEDIDAS DE DISPERSIÓN						
	Media (kWh)	Mediana (kWh)	Q1 (kWh)	Q2 (kWh)	Q3 (kWh)	Desv_est (kWh)	Varianza	Int_Tot. (kWh)	Sv	Error_Est.	%RSD	Sx
Domingo (120)	5366.7	5273.6	4962.3	5225.8	5489.2	522.7	273192.9	2206.3	10.3	33.7	1026.8	47.7
Lunes (120)	5829.1	5914.2	4954.7	5764.6	6574.4	792.3	627701.1	2694.8	7.4	51.1	735.7	72.3
Martes (96)	6127.5	6370.1	5479.8	6085.1	6690.4	690.2	476386.4	2473.4	8.9	49.8	887.8	70.4
Miércoles (96)	6186.3	6493.5	5584.3	6128.0	6671.6	634.8	403030.2	2245.3	9.7	45.8	974.5	64.8
Jueves (96)	6249.2	6523.1	5575.9	6171.2	6766.4	641.3	411222.2	2096.7	9.7	46.3	974.5	65.4
Viernes (96)	6271.2	6505.6	5736.4	6241.1	6745.8	594.7	353633.4	1872.6	10.5	42.9	1054.6	60.7
Sábado (96)	5962.5	6005.6	5575.2	5949.3	6323.3	433.3	187792.1	1609.3	13.8	31.3	1375.9	44.2

Q1= Primer cuartil.

Q2= Segundo cuartil.

Q3= Tercer cuartil.

Sv= Coeficiente de variación

%RSD= Desviación estándar relativa

Sx= Error estándar de la media.

2.8.3. COMENTARIOS DE LOS RESULTADOS DE LAS MEDIDAS ESTADÍSTICAS.

Con respecto al sumario estadístico que se realizó a los datos crudos de la Tabla 2.2, se observa lo siguiente:

- a) Las medidas de tendencia central, muestran valores parecidos para cada uno de los días, esto representa normalidad en las series de datos.
- b) Las medidas de dispersión, de acuerdo con la magnitud de los datos manejados no son elevados, en cuanto a los errores obtenidos éstos parecen altos.
- c) Finalmente, no será necesario realizar pruebas de normalidad a las muestras, debido a que la metodología utilizada por las Redes Neuronales Artificiales es robusta; lo cual, implica no realizar un análisis sobre la presencia de valores desviados en las muestras.

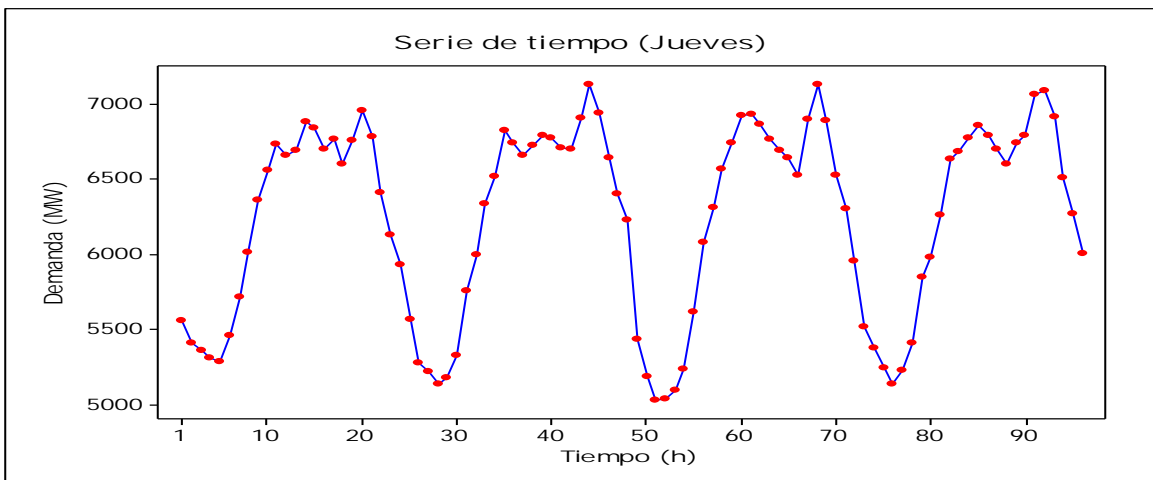
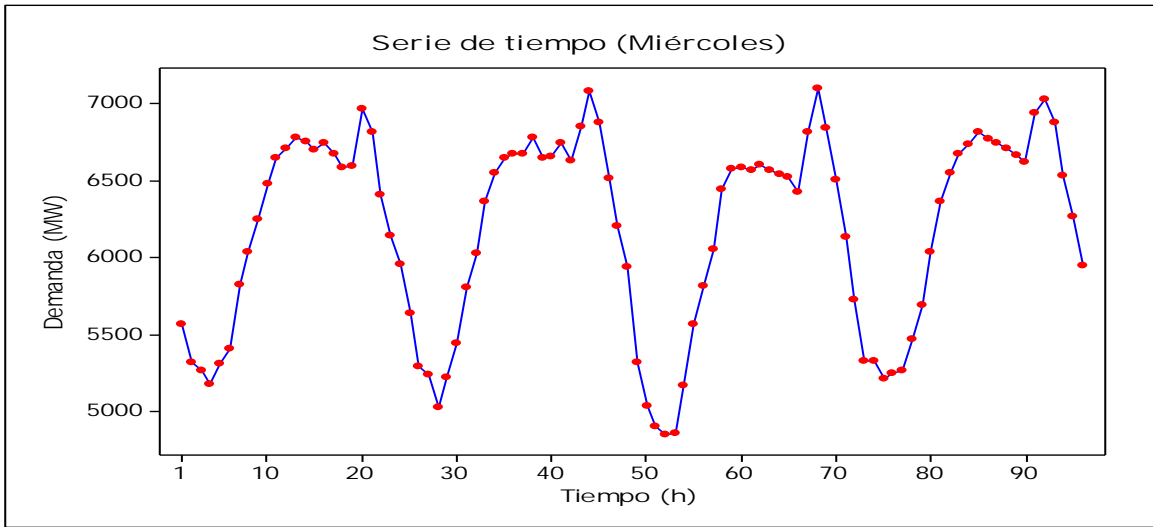
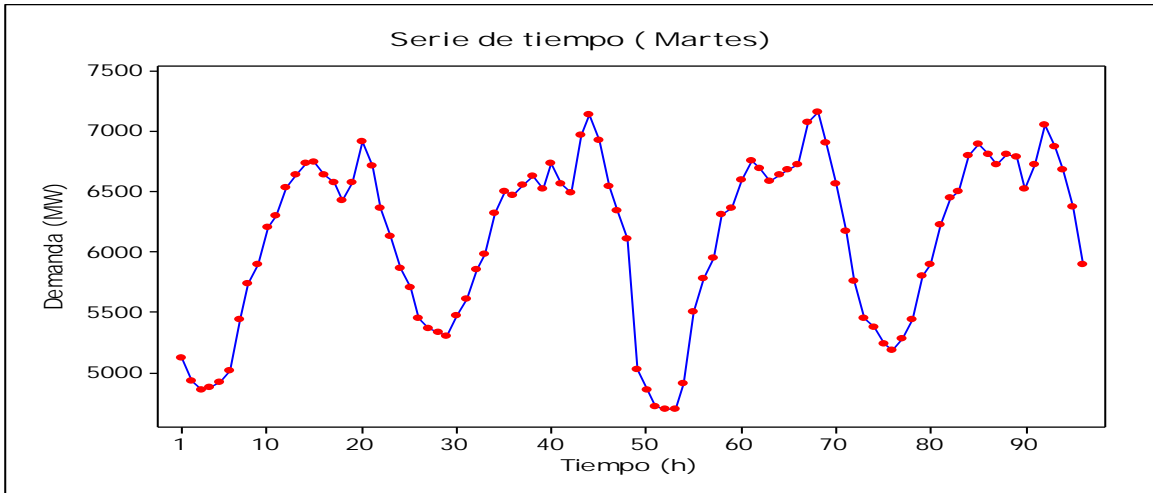


Figura 2.2 Curva de la demanda de energía eléctrica de los datos utilizados en el pronóstico correspondiente a cada día de la semana.

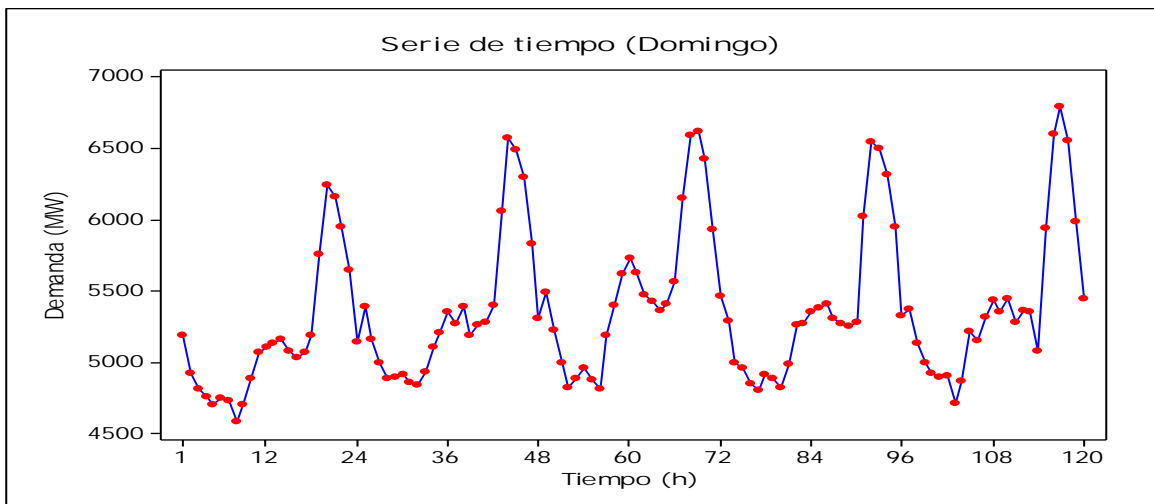
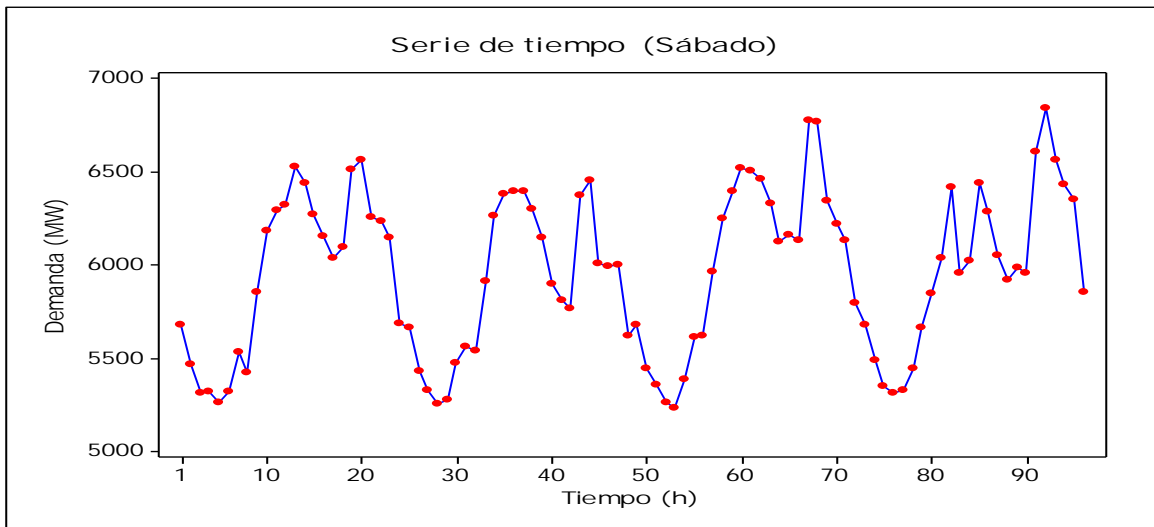
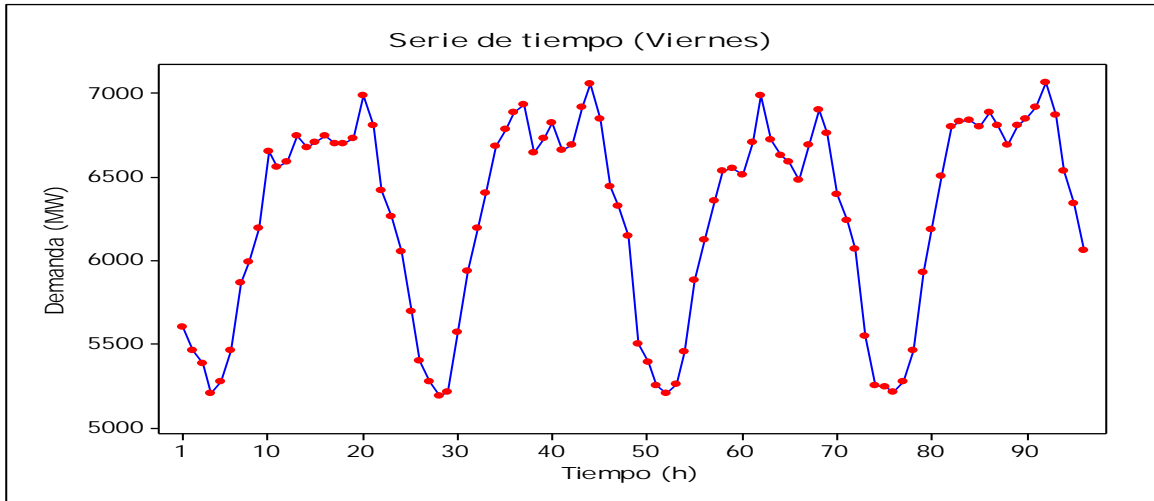


Figura 2.2 (Continuación). Curva de la demanda de energía eléctrica de los datos utilizados en el pronóstico correspondiente a cada día de la semana.

CAPITULO 3: REDES NEURONALES ARTIFICIALES.

3.1. RESUMEN.

En este capítulo se presenta un visión general de las redes neuronales; se introducen los conceptos, ventajas, aplicaciones y características fundamentales de la redes neuronales, además, se profundiza en los modelos de red PERCEPTRON, ADALINE Y BACKPROPAGATION (RETROPROPAGACIÓN), los cuales fueron utilizados como base para el desarrollo del diseño de una Red Neuronal Artificial para el pronóstico de la demanda de energía eléctrica.

3.2. INTRODUCCIÓN.

En la actualidad, las computadoras resultan ineficaces en el procesamiento de tareas de bajo nivel, como son: reconocimiento de patrones, percepción, control, etc. Debido a que no fueron diseñadas para ello. Para este tipo de problemas la naturaleza dotó a los seres vivos con una excelente solución: el procesamiento autoorganizado que surge de la interacción de numerosos procesadores elementales.

Las Redes Neuronales Artificiales parten de la idea de emular la estructura de funcionamiento del cerebro, creando un sistema compuesto por múltiples neuronas ampliamente interconectadas, y estudiar sí a partir de su autoorganización pueden reproducirse sus capacidades.

Existen diversas formas de definir lo que son las redes neuronales, desde las definiciones cortas y genéricas hasta las que intentan explicar detalladamente lo que significa red neuronal:

- a) Una nueva forma de computación, inspirada en modelos biológicos.

- b) Un modelo matemático compuesto por un gran número de elementos procesales organizados en niveles.

- c) Un sistema de computación hecho por un gran número de elementos simples, elementos de proceso muy interconectados, los cuales procesan información por medio de su estado dinámico como respuesta a entradas externas.

En las redes biológicas, las células neuronales (neuronas) constituyen los elementos de proceso. Las interconexiones se realizan por medio de las ramas de salida (axones) que producen un número variable de conexiones (sinapsis) con otras neuronas. Las redes neuronales son sistemas de simples elementos de proceso muy interconectados [19, 20, 21].

3.2.1. VENTAJAS DE LAS REDES NEURONALES

Debido a su constitución y a sus fundamentos, las Redes Neuronales Artificiales presentan un gran número de características semejantes a las del cerebro. Por ejemplo son capaces de aprender de la experiencia, de generalizar de casos anteriores a nuevos casos, de abstraer características esenciales a partir de entradas que representan información irrelevante, etc. Esto hace que ofrezcan numerosas ventajas y que este tipo de tecnología se esté aplicando en diversas áreas:

- a) *Aprendizaje adaptativo*. Capacidad de aprender a realizar tareas basadas en un entrenamiento o una experiencia inicial.
- b) *Autoorganización*. Una red neuronal puede crear su propia organización o representación de la información que recibe mediante una etapa de aprendizaje.
- c) *Tolerancia a fallos*: La destrucción parcial de una red conduce a una degradación de su estructura; sin embargo, algunas capacidades de la red se pueden retener, incluso sufriendo un gran daño.
- d) *Operación en tiempo real*: Los computadores neuronales pueden ser realizados, en paralelo y se diseñan y fabrican máquinas con hardware especial para obtener esta capacidad.
- e) *Fácil inserción dentro de la tecnología ya existente*: se pueden obtener chips especializados para redes neuronales que mejoran su capacidad en ciertas tareas. Ello facilitará la integración modular en los sistemas existentes.

3.2.2. APLICACIONES DE LAS REDES NEURONALES ARTIFICIALES.

Las primeras aplicaciones de las redes neuronales a problemas reales surgieron en los años sesenta; desde entonces se aplica el modelo de red lineal ADALINE como filtro adaptativo de manera rutinaria (por ejemplo, en los módems convencionales). Los modelos no lineales de redes neuronales permiten abordar un abanico de aplicaciones mucho más amplio, pero debido a los recursos de cómputo requeridos sus aplicaciones reales no surgen hasta finales de los años ochenta.

En la actualidad, la aplicación de las redes neuronales puede considerarse que ha llegado a su madurez, y en muchas ocasiones ya ni se cita que en determinado producto se empleen estas técnicas, por ser algo rutinario.

Hay muchos tipos diferentes de redes neuronales, cada uno de los cuales tiene una aplicación particular más apropiada. A continuación, se muestra un listado de aplicaciones comerciales separándolas en grupos según las distintas disciplinas:

- ❖ Biología
 - Aprender más acerca del cerebro y otros sistemas.
 - Obtención de modelos de la retina.
- ❖ Empresa
 - Evaluación de probabilidad de formaciones geológicas y petrolíferas.
 - Identificación de candidatos para posiciones específicas.
 - Explotación de bases de datos.
 - Optimización de plazas y horarios de líneas de vuelo.
 - Reconocimiento de caracteres escritos.
- ❖ Medio ambiente
 - Analizar tendencias y patrones.
 - Pronóstico del tiempo.
- ❖ Finanzas
 - Pronóstico de la evolución de los precios.
 - Valoración del riesgo de los créditos.
 - Identificación de falsificaciones.
 - Interpretación de firmas.
- ❖ Manufacturación
 - Robots automatizados y sistemas de control.
 - Control de producción en líneas de proceso.
 - Inspección de calidad.

- ❖ Medicina
 - Analizadores del habla para la ayuda de audición de sordos profundos.
 - Diagnostico y tratamiento a partir de síntomas y/o de datos analíticos.
 - Monitorización en cirugía.
 - Predicción de reacciones adversas a los medicamentos.
 - Lectores de rayos X.
- ❖ Militares
 - Clasificación de señales de radar.
 - Creación de armas inteligentes.
 - Optimización del uso de recursos escasos.
 - Reconocimiento y seguimiento de tiro al blanco.

3.3 FUNDAMENTOS DE LAS REDES NEURONALES ARTIFICIALES.

Los fundamentos de las RNA se basan en la estructura neuronal del cerebro para tratar de reproducir sus capacidades, donde la neurona es el elemento fundamental. Es por esto, que es necesario plasmar algunos conceptos básicos de los sistemas neuronales biológicos, para poder establecer el paralelismo que hay entre ambos.

Se estima que el sistema nervioso contiene alrededor de cien mil millones de neuronas. Vistas desde el microscopio, este tipo de células puede presentarse en múltiples formas, aunque muchas de ellas presentan un aspecto similar (Figura 3.1), con un cuerpo celular o soma (de entre 10 y 80 micras de longitud), del que surge un denso árbol de ramificaciones (árbol dendrítico) compuesto por las dendritas, y del cual parte una fibra tubular denominada axón (cuya longitud varía desde 100 micras hasta el metro en el caso de las neuronas motoras), que también se ramifica en su extremo final para conectar con otras neuronas.

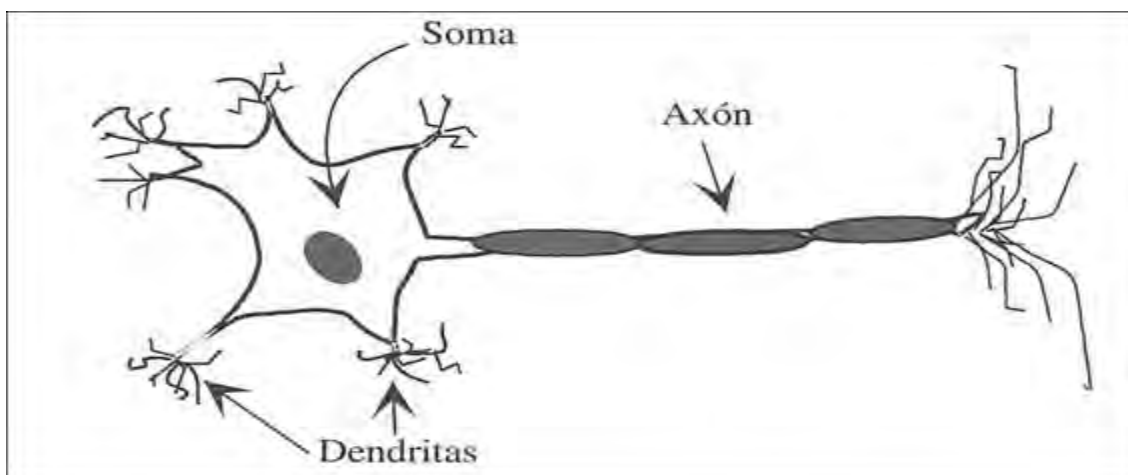


Figura 3.1. Estructura de una neurona biológica típica.

Desde una perspectiva funcional, las neuronas constituyen procesadores de información sencillos. Como todo sistema de este tipo, poseen un canal de entrada de información, las dendritas, un órgano de cómputo, el soma, y un canal de salida, el axón [20, 21, 22].

3.3.1 NATURALEZA BIOELÉCTRICA DE LAS NEURONAS

Las neuronas utilizan dos tipos de señales para comunicarse, éstas son de distinta naturaleza. Eléctrica y química. La señal que genera la neurona en su núcleo y que es transportada a lo largo del axón es de carácter eléctrico, mientras que la señal de origen químico es aquella que se presenta entre los terminales axónicos de una neurona y las dendritas de la siguiente neurona; ésta comunicación se realiza mediante moléculas de sustancias transmisoras (neurotransmisores) que fluyen a través de las sinapsis.

Se conoce como sinapsis a la unión entre dos neuronas. En el tipo de sinapsis más común no existe contacto físico entre las neuronas, sino que éstas permanecen separadas por un pequeño vacío (espacio sináptico) de entre 50 y 200 Angstroms. En relación a la sinapsis, se habla de neuronas presinápticas (las que envían las señales) y postsinápticas (las que reciben las señales). Existen dos tipos de sinapsis: a) las sinapsis excitadoras, cuyos neurotransmisores provocan disminuciones de potencial en la membrana de la célula postsináptica, facilitando la generación de impulsos a mayor velocidad, y b) las sinapsis inhibitoras, cuyos neurotransmisores tienden a estabilizar el potencial de la membrana, dificultando la emisión de impulsos.

3.3.2. MODELO DE UNA RED NEURONAL ARTIFICIAL.

La neurona artificial pretende emular las características de una neurona biológica, así, cada neurona i -ésima está caracterizada en todo momento por un valor numérico llamado valor de activación x_i ; asociado a cada unidad, existe una función de salida f , que transforma el valor actual de activación en una señal de salida y_i . Dicha señal es enviada a través de los canales de comunicación unidireccionales a otras unidades de la red; en estos canales la señal se modifica de acuerdo con la sinapsis (el peso, w_{ij}) asociados a cada uno de ellos según una determinada regla. A continuación se describen brevemente los conceptos básicos de los distintos componentes de un modelo de Red Neuronal Artificial.

3.3.2.1 UNIDADES DE PROCESO

Si se tienen N unidades (neuronas), pueden ser ordenadas arbitrariamente y designar la j -ésima unidad como U_j . Su labor radica en recibir entradas de las unidades vecinas y calcular un valor de salida para ser enviado a las unidades restantes. Existen tres tipos de unidades: entrada, salida y ocultas. Las unidades de entrada reciben información del entorno; las unidades de salida envían la señal fuera del sistema; las unidades ocultas son aquellas cuyas entradas y salidas se encuentran dentro del sistema.

3.3.2.2 *VALOR DE ACTIVACIÓN*

Además de un conjunto de unidades, la representación necesita los valores del sistema en un tiempo t . Esto se especifica con un vector de N números reales, que representa el valor de activación del conjunto de unidades de procesamiento. Cada elemento del vector representa la activación de una unidad en el tiempo t . La activación de una unidad en el tiempo t se designa por; es decir:

El procesamiento que realiza la red se ve como la evolución de un patrón de activación en el conjunto de unidades que lo componen a través del tiempo [20, 21, 22].

3.3.2.3 *FUNCIÓN DE SALIDA O DE TRANSFERENCIA*

Entre las unidades o neuronas que conforman una Red Neuronal Artificial existe un conjunto de conexiones que unen unas con otras. Cada unidad transmite señales a aquellas que están conectadas con su salida. Asociada con cada unidad hay una función de salida, que transforma el valor actual de activación en una señal de salida; es decir:

El vector que contiene todas las salidas de todas las neuronas en un instante t es:

3.3.2.4 CONEXIONES ENTRE NEURONAS

Las neuronas que unen a las neuronas tienen asociado un peso, que es el que hace que la red aprenda. Se considerará w_{ij} como el valor de salida de una neurona i en un instante dado. Una neurona recibe un conjunto de señales que le dan información del valor de activación de todas las neuronas con las que se encuentra conectada. Cada conexión entre la neurona i y la neurona j esta ponderada por un peso w_{ij} . Normalmente, como simplificación, se considera que el efecto de cada señal es aditivo, de tal forma que la entrada neta que recibe la neurona j es la suma del producto de cada señal individual por el valor de la sinapsis que conecta ambas neuronas:

Esta regla muestra el procedimiento a seguir para combinar los valores de entrada a una unidad con los pesos de las conexiones que llegan a esa unidad y es conocida como regla de propagación.

3.3.2.5 FUNCIÓN O REGLA DE ACTIVACIÓN

Así como es necesaria una regla que combine las entradas a una neurona con los pesos de las conexiones, también se requiere una regla que combine las entradas con el valor actual de la neurona para producir un nuevo valor de activación. Esta función F produce un nuevo valor de activación en una neurona a partir del estado Net_j que existía y la combinación de las entradas con los pesos de las conexiones Net_j .

Dado el estado de activación Net_j de la unidad j y la entrada total que llega a ella, Net_j , el estado de activación siguiente, Net_j , se obtiene aplicando una función F llamada función de activación.

En la mayoría de los casos, F es la función identidad, por lo que el estado de activación de una neurona $t+1$ coincidirá con el Net de la misma en t . En este caso, el parámetro que se le pasa a la función de salida, f , de la neurona será directamente el Net . El estado de activación anterior no se tiene en cuenta. Según esto, la salida de una neurona i quedará según la siguiente expresión:

3.3.2.6 *REGLA DE APRENDIZAJE*

Biológicamente, se acepta que la información memorizada en el cerebro está más relacionada con los valores sinápticos de las conexiones entre las neuronas que con ellas mismas; es decir, el conocimiento se encuentra en la sinapsis. En el caso de las Redes Neuronales Artificiales, se puede considerar que el conocimiento se encuentra representado en los pesos de las conexiones entre neuronas. Todo proceso de aprendizaje implica cierto número de cambios en estas conexiones. En realidad, puede decirse que se aprende modificando los valores de los pesos de la red.

3.3.3. ESTRUCTURA DE UNA RED NEURONAL ARTIFICIAL

La estructura de una Red Neuronal Artificial se compone principalmente por los siguientes elementos:

- a) Número de niveles o capas.
- b) Número de neuronas por nivel.
- c) Patrones de conexión.
- d) Flujo de información.

3.3.3.1 *NIVELES O CAPAS DE NEURONAS*

La configuración de la neuronas dentro de una red se realiza formando niveles o capas de un número determinado de neuronas cada una. A partir de su distribución dentro de la red, se distingues tres tipos de capas:

- a) **Capa de entrada:** es la capa que recibe directamente la información proveniente de las fuentes externas.
- b) **Capa oculta:** son internas a la red y no tienen contacto directo con el entorno exterior. El número de niveles ocultos puede estar entre cero y un número elevado. Las neuronas de las capas ocultas pueden estar interconectadas de distintas maneras, lo que determina, junto con su número, las distintas topologías de redes neuronales.
- c) **Capa de salida:** Transfieren información de la red hacia el exterior.

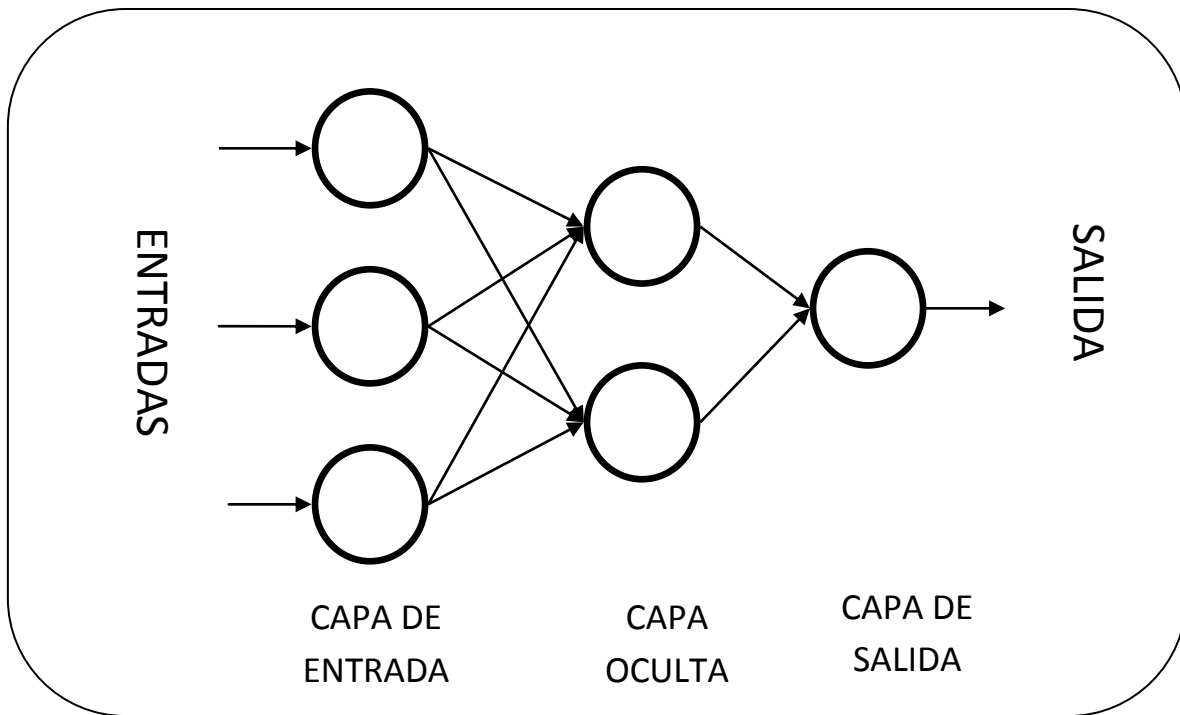


Figura 3.2. Estructura de una red multinivel con conexión hacia adelante

En la Figura 3.2 se muestra el esquema de la estructura de una posible red multicapa en la que cada nodo o neurona únicamente está conectada con las neuronas de un nivel superior.

3.3.3.2. FORMAS DE CONEXIÓN ENTRE NEURONAS

La conexión entre los nodos de una Red Neuronal Artificial está asociada con la forma en que las salidas de las neuronas están canalizadas para convertirse en entradas de otras neuronas. La señal de salida de un nodo puede ser una entrada de otro elemento de proceso, o incluso ser una entrada de sí mismo (conexión autorrecurrente).

Cuando ninguna salida de las neuronas es entrada de neuronas del mismo nivel o precedentes, la red se describe como de propagación hacia adelante (Figura 3.2). Cuando las salidas pueden ser conectadas como entradas de neuronas de niveles previos o del mismo nivel, incluyéndose ellas mismas, la red es de propagación hacia atrás.

3.4 CARACTERÍSTICAS DE LAS REDES NEURONALES ARTIFICIALES.

Los aspectos más importantes en la configuración de una Red Neuronal Artificial son: su estructura, el mecanismo de aprendizaje, tipo de asociación realizada entre la información de entrada y salida y la forma de representación de estas informaciones.

3.4.1. ARQUITECTURA DE LAS REDES NEURONALES ARTIFICIALES.

La arquitectura o topología de las redes neuronales consiste en la organización y disposición de las neuronas en la red formando capas o agrupaciones de neuronas más o menos alejadas de la entrada y la salida de la red. En este sentido, los parámetros fundamentales de la red son: el número de capas, el número de neuronas por capa, el grado de conectividad y el tipo de conexiones entre neuronas.

Cuando se realiza una clasificación de las redes en términos topológicos, se suele distinguir entre las redes con una sola capa de neuronas y las redes multicapa.

3.4.2. MECANISMO DE APRENDIZAJE

El aprendizaje es el proceso por el cual una red neuronal modifica sus pesos en respuesta a una información de entrada. Los cambios que se producen durante el proceso de aprendizaje se reducen a la destrucción, modificación y creación de conexiones entre las neuronas. En los sistemas biológicos existe una continua creación y destrucción de conexiones. En los modelos de Redes Neuronales Artificiales, la creación de una conexión implica que el peso de la misma pasa a tener un valor distinto de cero. De la misma forma, una conexión se destruye cuando su peso pasa a ser cero.

Un aspecto importante respecto al aprendizaje en las redes neuronales es el conocer como se modifican los valores de los pesos; es decir, cuales son los criterios que se siguen para cambiar el valor asignado a las conexiones cuando se pretende que la red aprenda una nueva información. Estos criterios determinan lo que se conoce como regla de aprendizaje de la red. Se suelen considerar dos tipos de reglas: la correspondiente a un aprendizaje supervisado, y las que manifiestan un aprendizaje no supervisado. La diferencia fundamental entre ambos estriba en la existencia o no de un agente externo (supervisor) que controle el proceso de aprendizaje de la red.

Otro criterio que suele utilizarse para diferenciar las reglas de aprendizaje se basa en considerar si la red puede aprender durante su funcionamiento habitual o si el aprendizaje supone una desconexión de la red; es decir su inhabilitación hasta que el proceso termine. El primer caso se trataría de un aprendizaje ON LINE, mientras que el segundo caso se le conoce como aprendizaje OFF LINE.

Cuando el aprendizaje es OFF LINE, se distingue entre una fase de aprendizaje o entrenamiento y una fase de operación o funcionamiento, existiendo un conjunto de datos de entrenamiento y un conjunto de datos de prueba que serán utilizados en la correspondiente fase. En las redes con aprendizaje OFF LINE, los pesos de las conexiones permanecen fijos después de que termina la etapa de entrenamiento de la red. Debido precisamente a su carácter estático, estos sistemas no presentan problemas de estabilidad en su funcionamiento.

En las redes con aprendizaje ON LINE no se distingue entre fase de entrenamiento y de operación, de tal forma que los pesos varían dinámicamente siempre que se presente una nueva información al sistema [20, 21, 22].

3.4.2.1. *REDES CON APRENDIZAJE SUPERVISADO*

El aprendizaje supervisado se caracteriza porque el proceso de aprendizaje se realiza mediante un entrenamiento controlado por un agente externo (supervisor, maestro) que determina la respuesta que debería generar la red a partir de una entrada determinada. El supervisor comprueba la salida de la red y en el caso de que ésta no coincida con la deseada, se procederá a modificar los pesos de las conexiones, con el fin de conseguir que la salida obtenida se aproxime a la deseada (Tabla 3.1).

En este tipo de aprendizaje se suele considerar, a su vez, tres formas de llevarlo a cabo que dan lugar a los siguientes aprendizajes supervisados:

- a) Aprendizaje por corrección de error.
- b) Aprendizaje por refuerzo.
- c) Aprendizaje estocástico.

El aprendizaje por corrección de error consiste en ajustar los pesos de las conexiones de la red en función de la diferencia entre los valores deseados y los obtenidos en la salida de la red; es decir, en función del error cometido en la salida.

Una regla o algoritmo simple de aprendizaje por corrección de error podría ser el siguiente:

Siendo

: Variación en el peso de la conexión entre las neuronas i y j

- : Valor de salida de la neurona i .
- : Valor de salida deseado de la neurona j .
- : Valor de salida obtenido en la neurona j .
- α : Factor de aprendizaje ($0 < \alpha \leq 1$) que regula la velocidad del aprendizaje.

El aprendizaje por refuerzo se trata de un aprendizaje supervisado, que se basa en la idea de no disponer de un ejemplo completo del comportamiento deseado; es decir, de no indicar durante el entrenamiento exactamente la salida que se desea que proporcione la red ante una determinada entrada.

El aprendizaje estocástico consiste básicamente en realizar cambios aleatorios en los valores de los pesos de las conexiones de la red y evaluar su efecto a partir del objetivo deseado y de distribuciones de probabilidad.

Tabla 3.1. Tipos de redes con aprendizaje supervisado más conocidos

TIPO DE APRENDIZAJE SUPERVISADO		MODELO DE RED
APRENDIZAJE POR CORRECCION DE ERROR	OFF LINE	PERCEPTRON
		ADALINE/MADALINE
		BACKPROPAGATION (RETROPROPAGACIÓN)
		BRAIN-STATE-IN-A-BOX
		COUNTERPROPAGATION
APRENDIZAJE POR REFUERZO	ON LINE	LINEAR REWARD PENALTY
		ASSOCIATIVE REW. PENALTY
		ADAPTIVE HEURISTIC CRITIC
APRENDIZAJE ESTOCASTICO	OFF LINE	BOLTZMANN MACHINE
		CAUCHY MACHINE

3.4.2.2. *REDES CON APRENDIZAJE NO SUPERVISADO*

Las redes con aprendizaje no supervisado no requieren influencia externa para ajustar los pesos de las conexiones entre neuronas. La red no recibe ninguna información por parte del entorno que le indique si la salida generada en respuesta a una determinada entrada es o no correcta; por ello, suele decirse que estas redes son capaces de autoorganizarse.

Estas redes deben encontrar las características, regularidades, correlaciones o categorías que se puedan establecer entre los datos que se presenten en su entrada. Puesto que no hay un supervisor que indique a la red la respuesta que debe generar ante una entrada correcta, cabría preguntarse precisamente por lo que la red genera en estos casos. Existen varias posibilidades en cuanto a la interpretación de la salida de estas redes, que dependen de su estructura y del algoritmo de aprendizaje empleado.

En algunos casos, la salida representa el grado de familiaridad o similitud entre la información que se le está presentando en la entrada y las informaciones que se le han mostrado hasta entonces (en el pasado). En otro caso, podría realizar una clusterización (clustering) o establecimiento de categorías, indicando la red a la salida a qué categoría pertenece la información presentada a la entrada, siendo la propia red quien debe encontrar las categorías apropiadas a partir de correlaciones entre las informaciones presentadas.

En cuanto a los algoritmos de aprendizaje no supervisado, en general se suelen considerar dos tipos que dan lugar a los siguientes aprendizajes [20, 21, 22]:

- a) Aprendizaje hebbiano.
- b) Aprendizaje competitivo y cooperativo.

En el primer caso, normalmente se pretende medir la familiaridad o extraer las características de los datos de entrada, mientras que el segundo suele orientarse hacia la clusterización o clasificación de dichos datos (Tabla 3.2).

Tabla 3.2. Tipos de redes con aprendizaje no supervisado más importantes.

TIPO DE APRENDIZAJE NO SUPERVISADO		MODELO DE RED
APRENDIZAJE HEBBIANO	OFF LINE	HOPFIELD
		LEARNING MATRIX
		TEMPORAL ASSOC. MEMORY
		LINEAR ASSOCIATIVE MEMORY (LAM)
		OPTIMAL LAM
		DRIVE-REINFORCEMENT
		FUZZY ASSOCIATIVE MEMORY
	ON LINE	ADDITIVE GROSSBERG
		SHUNTING GROSSBERG
		BIDIRECTIONAL ASSOCIATIVE MEMORY (BAM)
APRENDIZAJE COMPETITIVO /COOPERATIVO	OFF LINE	LEARNING VECTOR QUANTIZER
		COGNITRON/NEOCOGNITRON
		TOPOLOGY PRESERVING MAP
	ON LINE	ADAPTIVE RESONANCE THEORY

3.4.3. CARACTERÍSTICAS DE LOS MODELOS DE RED NEURONAL ARTIFICIAL.

En la Tabla 3.3 se muestran los modelos de Red Neuronal Artificial utilizados en la elaboración de este proyecto, junto con sus autores y las características más relevantes de cada uno:

- a) Topología.
- b) Mecanismo de aprendizaje.
- c) Señales de entrada/salida.
- d) Autor (es).

Tabla 3.3. Características de los modelos de red neuronal utilizados en el pronóstico.

MODELO DE RED	TOPOLOGÍA	APRENDIZAJE			AUTORES
		ON/OFF LINE	SUPERVISADO/NO SUPERVISADO	REGLA	
PERCEPTRON	2 CAPAS (CONEXIÓN HACIA ADELANTE)	OFF LINE	SUPERVISADO	CORRECCIÓN DE ERROR	ROSENBLAT 1958
ADALINE	2 CAPAS (CONEXIÓN HACIA ADELANTE)	OFF LINE	SUPERVISADO	CORRECCIÓN DE ERROR (MÍNIMO ERROR CUADRADO MEDIO)	WIDROW HOFF 1960
BACKPROPAGATION (RETROPROPAGACIÓN)	N CAPAS (CONEXIÓN HACIA ADELANTE)	OFF LINE	SUPERVISADO	CORRECCIÓN DE ERROR (REGLA DELTA GENERALIZADA)	RUMELHART 1986

3.5 REDES NEURONALES ARTIFICIALES CON CONEXIONES HACIA ADELANTE.

En las redes con conexión hacia adelante (feedforward), todas las señales neuronales se propagan hacia adelante a través de las capas de la red. No existen conexiones hacia atrás (ninguna salida de neuronas de una capa i se aplica a la entrada de neuronas de capas $i-1$, $i-2$, ...), y normalmente tampoco autorrecurrentes (salida de una neurona aplicada a su propia entrada), ni laterales (salida de una neurona aplicada a la entrada de neuronas de la misma capa).

Este apartado se enfocará en tres modelos en especial; PERCEPTRON, ADALINE, y BACKPROPAGATION (RETROPROPAGACIÓN) ya que, estos modelos son los que fueron utilizados para el desarrollo de este proyecto.

3.5.1 EL PERCEPTRON

Este fue el primer modelo de red neuronal desarrollado por Rosenblatt en 1958. Despertó un enorme interés en los años 60's, debido a su capacidad para aprender patrones sencillos.

3.5.1.1 REGLA DE APRENDIZAJE DEL PERCEPTRON

El algoritmo de aprendizaje del Perceptron es de tipo supervisado, lo cual requiere que sus resultados sean evaluados y se realicen oportunas modificaciones del sistema si fuera necesario. Los valores de los pesos pueden determinar, como se ha dicho, el funcionamiento de la red; estos valores se pueden fijar o adaptar utilizando diferentes algoritmos de entrenamiento de la red.

A continuación se abordará el algoritmo de convergencia de ajuste de pesos para realizar el aprendizaje de un Perceptron (Aprendizaje por corrección de error) con N elementos procesales de entrada y un único elemento procesal de salida:

1. Inicialización de pesos y del umbral

Inicialmente se asignan valores aleatorios a cada uno de los pesos (w_{ij}) de las conexiones y al umbral (θ).

2. Presentación de un nuevo par (Entrada, Salida esperada).

Presentar un nuevo patrón de entrada (x_1, x_2, \dots, x_n) junto con la salida esperada (y) .

3. Cálculo de la salida actual

Siendo f la función de transferencia escalón.

4. Adaptación de los pesos

Donde y_d representa la salida deseada. En estas ecuaciones, α es un factor de ganancia en el rango de 0.0 a 1.0. Este factor debe ser ajustado de forma que satisfaga tanto los requerimientos de aprendizaje rápido como la estabilidad de las estimaciones de los pesos. Este proceso se repite hasta que el error que se produce para cada uno de los patrones (diferencia entre el valor de salida deseado y obtenido) es cero o bien menor que un valor preestablecido. Obsérvese que los pesos no se cambian si la red ha tomado la decisión correcta.

5. volver al paso 2

Este algoritmo es extensible al caso de múltiples neuronas en la capa de salida. El Perceptron será capaz de aprender a clasificar todas sus entradas, en un número finito de pasos, siempre y cuando el conjunto de los patrones de entrada sea linealmente separable. En tal caso, puede demostrarse que el aprendizaje de la red realiza un número finito de pasos.

3.5.2. LA RED ADALINE

La red ADALINE (ADAPtive LINear Element) fue desarrollada por Bernie Widrow en la Universidad de Stanford poco después de que Rosenblatt desarrollara el Perceptron. La estructura de ADALINE es esencialmente la misma que la del Perceptron [20, 21, 22]. La diferencia fundamental respecto al Perceptron se refiere al mecanismo de aprendizaje.

La red ADALINE utiliza la denominada regla Delta de Widrow-Hoff o regla del mínimo error cuadrado medio (LMS), basada en la búsqueda del mínimo de una expresión del error entre la salida deseada y la salida lineal obtenida antes de aplicarle la función de activación escalón. Debido a esta nueva forma de evaluar el error, estas redes pueden procesar información analógica, tanto de entrada como de salida, utilizando una función de activación lineal o sigmoideal.

3.5.2.1 APRENDIZAJE DE LA RED ADALINE

La red ADALINE utiliza un aprendizaje OFF LINE con supervisión denominado LMS (Least Mean Squared) o regla del mínimo error cuadrado medio. También se conoce como regla Delta porque trata de minimizar una delta o diferencia entre el valor observado y el deseado en la salida de la red; como ocurre con el Perceptron, solo que ahora la salida considerada es el valor previo a la aplicación de la función de activación lineal.

La regla de aprendizaje de los mínimos cuadrados (Least Mean Squared) es un método para hallar un vector de pesos W deseado, el cual deberá ser único y asociar con éxito cada vector del conjunto de vectores o patrones de entrada con su correspondiente valor de salida correcto (o deseado). Nótese que el problema de hallar un conjunto de pesos W que para un único vector de entrada X de lugar a un valor de salida correcto resulta sencillo, lo que no ocurre cuando se dispone de un conjunto de vectores de entrada, cada uno con su propio valor de salida asociado. El entrenamiento de la red consiste en adaptar los pesos a medida que se vayan presentando los patrones de entrenamiento y salidas deseadas para cada uno de ellos. Para cada combinación entrada-salida se realiza un proceso automático de pequeños ajustes en los valores de los pesos hasta que se obtienen las salidas correctas.

La primera cuestión que se debe resolver es la de definir qué significa obtener el mejor vector de pesos obtenido a partir unas parejas de valores ejemplo de forma que, una vez encontrado, desearemos que al aplicar todos los vectores de entrada se obtenga como resultado el valor de salida correcto. Se trata de eliminar o, por lo menos, minimizar la diferencia entre la salida deseada y salida real para todos los vectores de entrada.

Concretamente, la regla de aprendizaje LMS minimiza el error cuadrado medio, definido como:

—

Donde L es el número de vectores de entrada (patrones) que forman el conjunto de entrenamiento, y d_k la diferencia entre la salida deseada y la obtenida cuando se introduce el patrón k -ésimo, que, en el caso de la red ADALINE, se expresa como $d_k = y_k - \hat{y}_k$.

3.5.2.1.1 *REGLA DE APRENDIZAJE DE LA RED ADALINE*

1. Se aplica un vector o patrón de entrada, \mathbf{p} , en las entradas del ADALINE.
2. Se obtiene la salida lineal \mathbf{y} y se calcula la diferencia con respecto a la deseada \mathbf{d} .
3. Se actualizan los pesos
4. Se repiten los pasos del 1 al 3 con todos los vectores de entrada (L).
5. Si el error cuadrado medio:

—

Es un valor reducido aceptable, termina el proceso de aprendizaje; si este valor no es aceptable, se repite otra vez desde el paso 1 con todos los patrones.

El valor del parámetro α tiene una gran influencia sobre el entrenamiento. Si α es demasiado grande, la convergencia es posible que no se produzca, debido a que se darán saltos en torno al mínimo sin alcanzarlo. Si α es demasiado pequeño, se alcanza la convergencia, pero a costa de una etapa de aprendizaje más larga.

En cuanto al momento en el que se debe detener el entrenamiento, éste depende, sobre todo, de los requisitos de salida del sistema: se detiene el entrenamiento cuando el error observado es menor que el valor admisible en la señal de salida de forma sostenida. Se suele tomar el error cuadrático medio como la magnitud que determina el instante en el que un sistema ha convergido.

3.5.3 LA RED BACKPROPAGATION (RETROPROPAGACIÓN)

En 1986, Rumelhart, Hinton y Williams, basándose en los trabajos de otros investigadores formalizaron un método para que una red neuronal aprendiera la asociación que existe entre los patrones de entrada a la misma y las clases correspondientes, utilizando más niveles de neuronas que los que utilizó Rosenblatt para desarrollar el Perceptron. Este método, conocido en general como backpropagation (propagación del error hacia atrás), está basado en la generalización de la regla delta y, a pesar de sus propias limitaciones, ha ampliado de forma considerable el rango de aplicaciones de las redes neuronales.

El algoritmo de propagación hacia atrás, o retropropagación, es una regla de aprendizaje que se puede aplicar en modelos de redes con más de dos capas de células. Una característica importante de este algoritmo es la representación interna del conocimiento que es capaz de organizar en la capa intermedia de las células para conseguir cualquier correspondencia entre la entrada y la salida de la red.

De forma simplificada, el funcionamiento de una red backpropagation (backpropagation net, BPN) consiste en un aprendizaje de un conjunto predefinido de pares de entradas-salidas dados como ejemplo, empleando un ciclo propagación-adaptación de dos fases: primero se aplica un patrón de entrada como estímulo para la primera capa de neuronas de la red, se va propagando a través de todas las capas superiores hasta generar una salida. Se compara el resultado obtenido en las neuronas de salida con la salida que se desea obtener y se calcula un valor del error para cada neurona de salida. A continuación, estos errores se transmiten hacia atrás, partiendo de la capa de salida, hacia todas las neuronas de la capa intermedia que contribuyan directamente a la salida, recibiendo el porcentaje de error aproximado a la participación de la neurona intermedia en la salida original. Este proceso se repite, capa por capa, hasta que todas las neuronas de la red hayan recibido un error que describa su aportación relativa al error total. Basándose en el valor del error recibido, se reajustan los pesos de conexión de cada neurona, de manera que en la siguiente vez se presente el mismo patrón, la salida esté más cercana a la deseada; es decir, el error disminuya.

La importancia de la red backpropagation consiste en su capacidad de autoadaptar los pesos de las neuronas de las capas intermedias para aprender la relación que existe entre un conjunto de patrones dados como ejemplo y sus salidas correspondientes. Para poder aplicar esa misma relación, después del entrenamiento, a nuevos vectores de entrada con ruido o incompletas, dando una salida activa si la nueva entrada es parecida a las presentadas durante el aprendizaje. Esta característica importante, que se exige a los sistemas de aprendizaje, es la capacidad de generalización, entendida como la facilidad de dar salidas satisfactorias a entradas que el sistema no ha visto nunca en su fase de entrenamiento. La red debe encontrar una representación interna que le permita generar salidas deseadas cuando se le dan las entradas de entrenamiento, y que pueda aplicar, además, a entradas no presentadas durante la etapa de aprendizaje para clasificarlas según las características que compartan con los ejemplos de entrenamiento.

3.5.3.1 LA REGLA DELTA GENERALIZADA

La regla propuesta por Widrow en 1960 (regla delta) ha sido extendida a redes con capas intermedias (regla delta generalizada) con conexiones hacia adelante (feedforward) y cuyas células tienen funciones de activación continuas, dando lugar a un algoritmo de retropropagación (backpropagation). Estas funciones continuas son no decrecientes y derivables [20, 21, 22].

Funcionamiento del algoritmo:

El método que sigue la regla delta generalizada para ajustar los pesos es exactamente el mismo que el de la regla delta utilizada en el Perceptron y ADALINE; es decir, los pesos se actualizan de forma proporcional a la delta, o diferencia entre la salida deseada y la obtenida ($\epsilon = \text{sal. deseada} - \text{sal. obtenida}$).

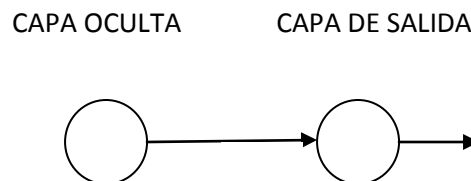


Figura 3.3. Conexión entre una neurona de una capa oculta con una neurona de salida

Dada una neurona (unidad j) y la salida que produce, y_j (Figura 3.3), el cambio que se produce en el peso de la conexión que une la salida de dicha neurona con la unidad i para un patrón de aprendizaje p determinado es:

En donde el subíndice p se refiere al patrón de aprendizaje concreto y α es la constante o tasa de aprendizaje.

En el punto en donde difieren la regla delta generalizada de la regla delta es el valor concreto de δ_j . Por otro lado, en las redes multinivel, a diferencia de las redes sin neuronas ocultas, en principio no se puede conocer la salida deseada de las neuronas de las capas ocultas para poder determinar los pesos en función del error cometido. Sin embargo, inicialmente sí se puede conocer la salida deseada de las neuronas de salida. Según esto, si se considera la unidad j de salida (Figura 3.3), entonces se define:

Donde y_j^p es la salida deseada de la neurona j para el patrón p y x_j es la entrada neta que recibe la neurona j .

3.5.3.2. *REGLA DE APRENDIZAJE DE LA RED BACKPROPAGATION*

A continuación se presentan, a modo de síntesis, los pasos y ecuaciones a utilizar para aplicar el algoritmo de entrenamiento:

1. Inicializar los pesos de la red con valores pequeños aleatorios.
2. Presentar un patrón de entrada, $\{x_p\}$, y especificar la salida deseada que debe generar la red: $\{d_p\}$.
3. Calcular la salida actual de la red, para ello se presentan las entradas a la red y se calcula la salida que presenta cada capa hasta llegar a la capa de salida ésta será la salida de la red $\{o_p\}$.

- Se calculan las entradas netas para las neuronas ocultas procedentes de las neuronas de entrada.

Para una neurona j oculta:

En donde el índice h se refiere a magnitudes de la capa oculta (hidden); el subíndice p , al p -ésimo vector de entrenamiento, y j a la j -ésima neurona oculta.

- Se realizan los mismos cálculos para obtener las salidas de las neuronas de salida (*capa o: output*):

4. Calcular los términos de error para todas las neuronas.

Si la neurona k es una neurona de la capa de salida, el valor de la delta es:

Mientras que para las neuronas ocultas:

Donde se observa que el error en las capas ocultas depende de todos los términos de error de la capa de salida. De aquí el término de *propagación hacia atrás*.

5. Actualización de pesos

Para los pesos de las neuronas de la capa de salida:

Y para los pesos de las neuronas de la capa oculta:

6. El proceso se repite hasta que el termino error

-

Resulta aceptablemente pequeño para cada uno de los patrones aprendidos.

CAPITULO 4: PRONÓSTICO DE LA DEMANDA DE ENERGÍA ELÉCTRICA UTILIZANDO REDES NEURONALES ARTIFICIALES.

4.1. RESUMEN

En presente capitulo se lleva a cabo la conclusión del proyecto, el cual, tiene como objetivo primordial el diseñar una Red Neuronal Artificial para el pronóstico de la demanda de energía eléctrica. Se presentan y desarrollan las fases principales para alcanzar dicho objetivo.

4.2. MODELADO DE SERIES DE TIEMPO CON REDES NEURONALES ARTIFICIALES.

La dificultad de relacionar diferentes variables con el comportamiento cambiante de la demanda de electricidad, hace necesario el uso de series temporales en la tarea de pronosticar la misma. El pronóstico emplea sólo el comportamiento pasado de la variable.

El modelado de series temporales mediante la técnica de Redes Neuronales Artificiales consta de cuatro fases principales. (Figura 4.1), planteamiento del problema, preparación de datos, modelado de la Red Neuronal Artificial y análisis de resultados [23].

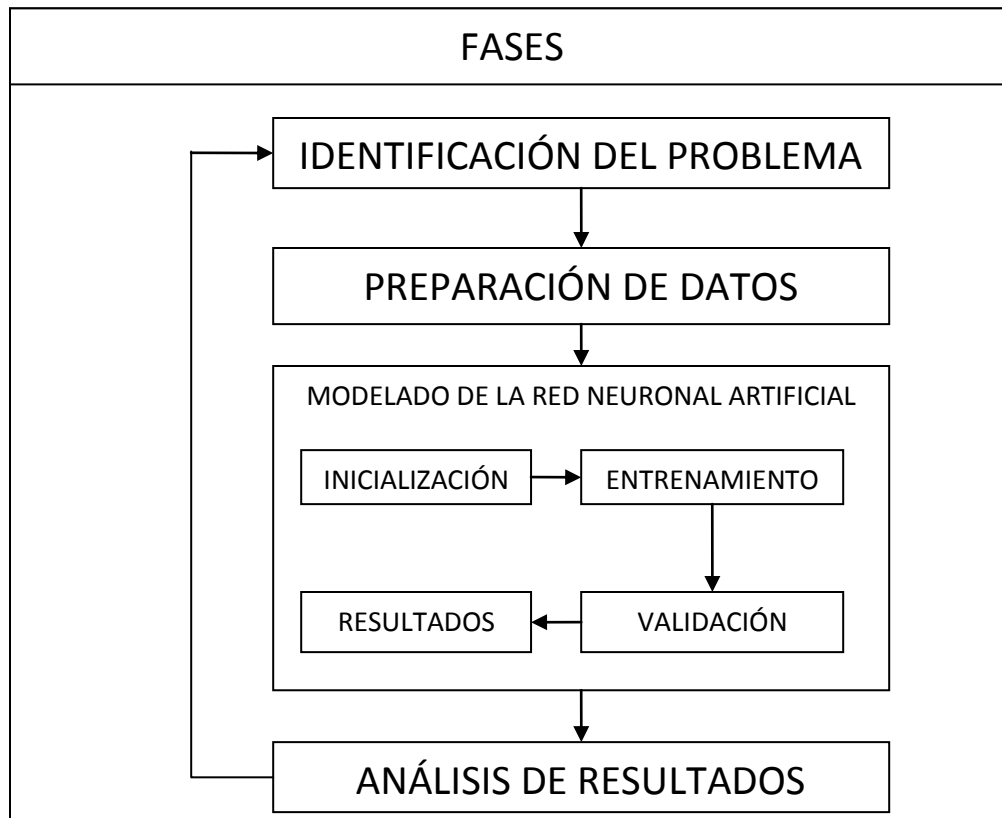


Figura 4.1. Fases de modelado de Redes Neuronales Artificiales.

La primera fase, la cual se enfoca en la identificación del problema conlleva a resaltar la importancia de realizar el modelado de series de tiempo con Redes Neuronales Artificiales dirigido a un problema específico. En este caso, el presente proyecto aborda el problema relacionado con la predicción de la demanda de la energía eléctrica, haciendo uso de las mediciones horarias de demanda de electricidad. La importancia de realizar una previsión correcta de esta variable ya se ha mencionado de manera amplia en el capítulo referente a la demanda de energía eléctrica.

4.2.1. PREPARACIÓN DE DATOS

La preparación de datos es un importante y crítico paso en el modelado de series de tiempo mediante redes neuronales, ya que, la calidad de los valores de entrada puede influenciar de manera importante el resultado que arroje el modelo de red neuronal. Generalmente, los datos que se preparan correctamente son fáciles de manejar. Por otro lado, no preparar los datos puede dificultar la tarea de análisis, o simplemente hacerla imposible.

Como ya se ha mencionado, los datos utilizados para realizar el presente proyecto muestran dos patrones importantes; tendencia y estacionalidad, lo cual, en el modelado de series de tiempo mediante redes neuronales enfocadas al pronóstico puede tener efectos no deseados en los resultados que se obtengan, por lo tanto, eliminar la tendencia y desestacionalizar las series de tiempo es imperativo en este análisis. Una alternativa para remover este tipo de patrones, es el método de la diferenciación. La diferenciación en la series de tiempo se define como el cambio entre cada observación de la serie original, la expresión 4.1 ejemplifica lo anterior.

La serie diferenciada cuenta con $n-1$ valores, debido a que no es posible calcular una diferencia del primer valor. Para ejemplificar, a continuación se muestra la serie de tiempo transformada para el día domingo (Figura 4.2). Al final del capítulo se presentan las series transformadas para los días restantes de la semana.

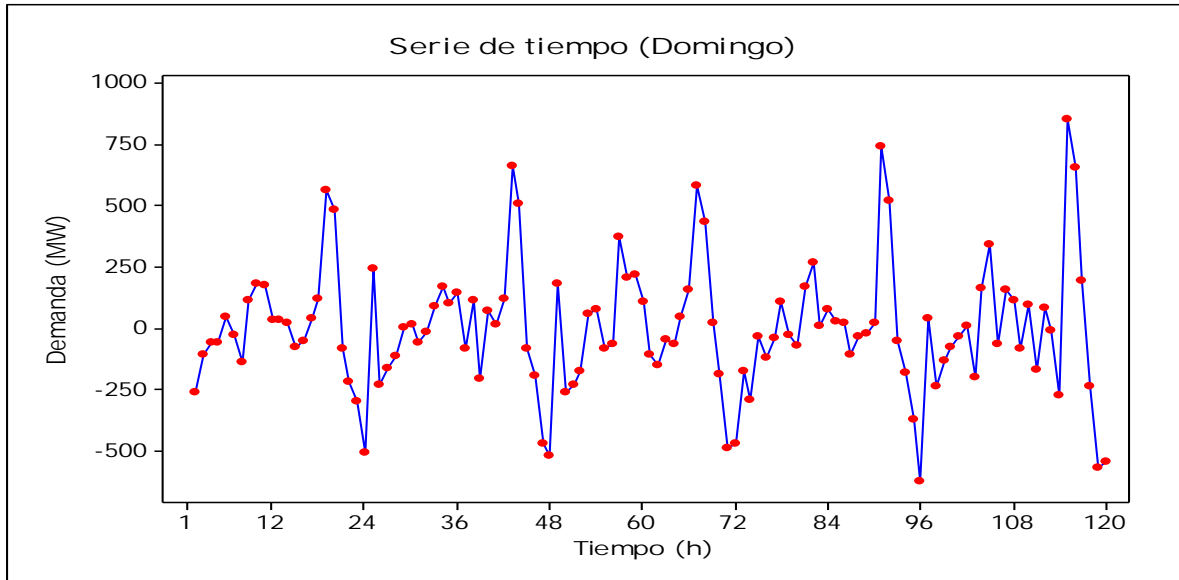


Figura 4.2. Serie de tiempo diferenciada (Domingo).

Otra cuestión importante a resolver es la multi-escala de los valores de la variable, debido a que, durante el aprendizaje de la red este tipo de valores conducen a una inestabilidad en el modelo. Los datos utilizados para la inserción en las neuronas de entrada típicamente tienen una escala entre -1 y 1 ó 0 y 1. Para el tratamiento de la multi-escala es necesaria la normalización de los datos [24]. Este proceso se realiza mediante la siguiente expresión:

Donde:

= Valor normalizado.

v = Valor correspondiente a la variable x .

= Valor mínimo y máximo de la variable x .

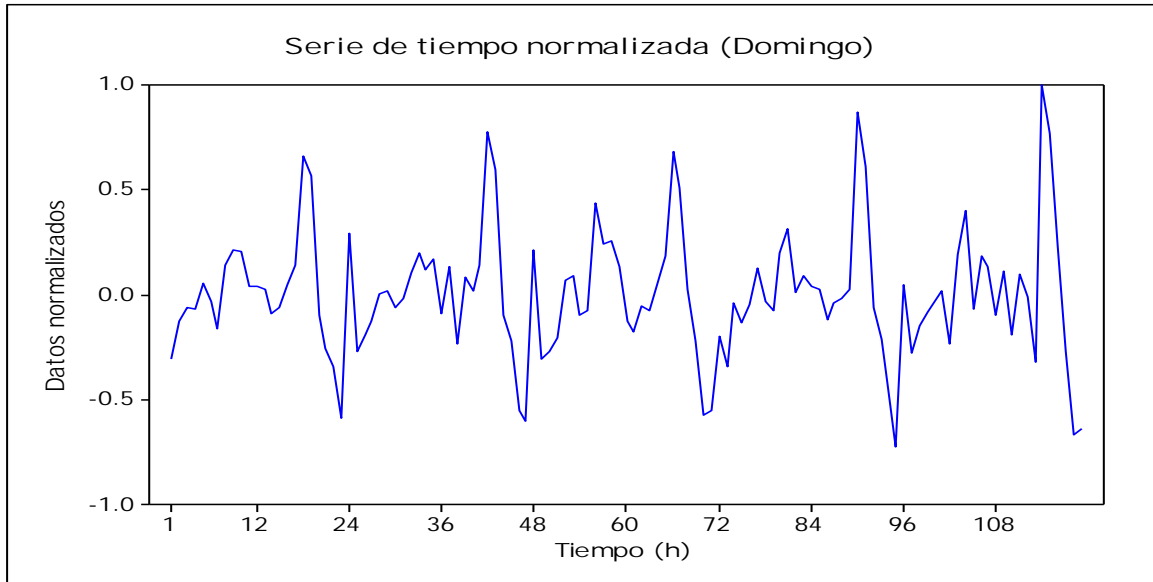


Figura 4.3. Serie de tiempo normalizada (Domingo)

En la Figura 4.3 se muestra un ejemplo de la serie de tiempo normalizada utilizada para el desarrollo de los modelos de red neuronal. Al final del capítulo se muestran las series de tiempo normalizadas para los restantes días de la semana.

4.2.2. MODELADO DE LA RED NEURONAL ARTIFICIAL.

La siguiente fase para pronosticar la demanda de energía es el modelado de la Red Neuronal Artificial, esta fase, consta de cuatro etapas principales: inicialización, entrenamiento, validación y resultados.

4.2.2.1. INICIALIZACIÓN.

Como ya se plasmó en la etapa anterior, los datos ya se encuentran caracterizados para ser incluidos en las neuronas de entrada de cada modelo. El siguiente paso para seguir es el de la estructura que tendrán los modelos, es decir, el número de capas, número de neuronas de entrada, número de neuronas en la capa oculta y el número de neuronas en la capa de salida. Para conocer la cantidad de neuronas idóneas en la capa de entrada existen herramientas estadísticas que ayudan a orientar esta decisión, tal es el caso de la función de autocorrelación. El coeficiente de autocorrelación se define como la correlación de la serie con ella misma en sus retardos de tiempo, la expresión que la define es la siguiente:

Al obtener ρ_1 , se encuentra la relación lineal existente entre el valor presente x_t y su primer retardo, de la misma manera ocurre con ρ_2 , respecto del retardo numero 2, y así sucesivamente. El valor de los coeficientes siempre se encuentra entre -1 y 1, y no pueden sobrepasar dicho valor. Al obtener los valores sucesivos de ρ_k , se conforma la función de autocorrelación, la cual además de representar una herramienta cuantitativa es una gráfica visual de mucha ayuda para determinar el número de vectores de entrada a la red. (Figura 4.4).

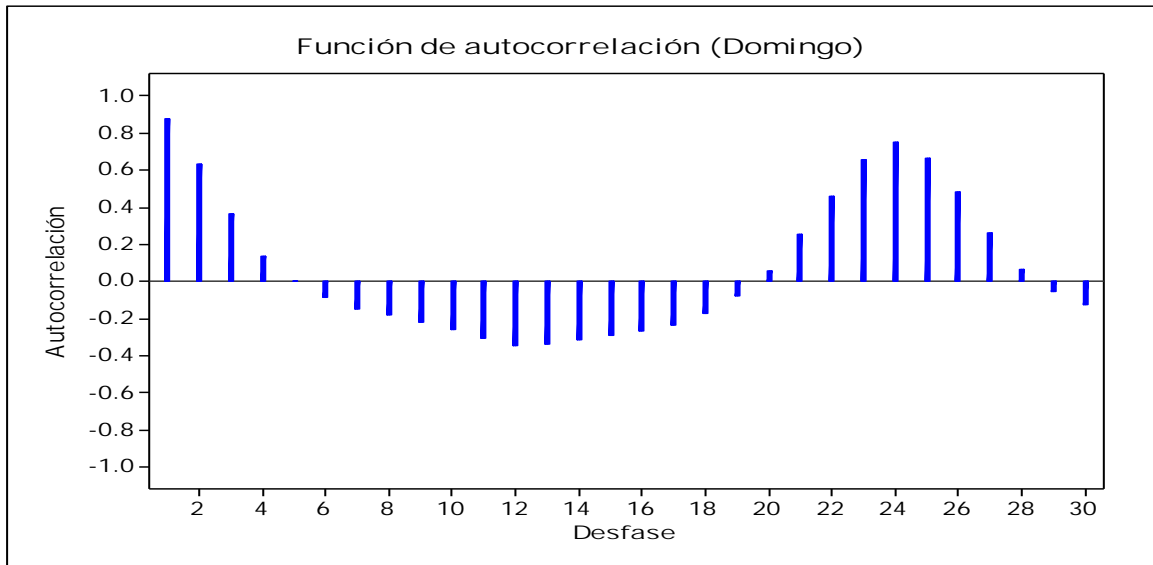


Figura 4.4. Función de autocorrelación (Domingo).

En lo que concierne a los parámetros restantes que debe presentar el modelo de red, no existe una regla general para integrar estos parámetros, sin embargo, es aceptado que una red con tres capas, conexión hacia adelante, con una función de transferencia identificada en la unidad de salida y una función lógica en las unidades de la capa intermedia puede aproximar cualquier función continua de una manera satisfactoria.

4.2.2.1.1. MODELOS DE REDES NEURONALES ARTIFICIALES PROPUESTOS.

Con base en lo anterior las configuraciones propuestas para realizar el proyecto son las siguientes:

1. Modelo 1: Tres capas, dos neuronas en la capa de entrada, dos neuronas en la capa oculta y una neurona en la capa de salida. (Figura 4.5).

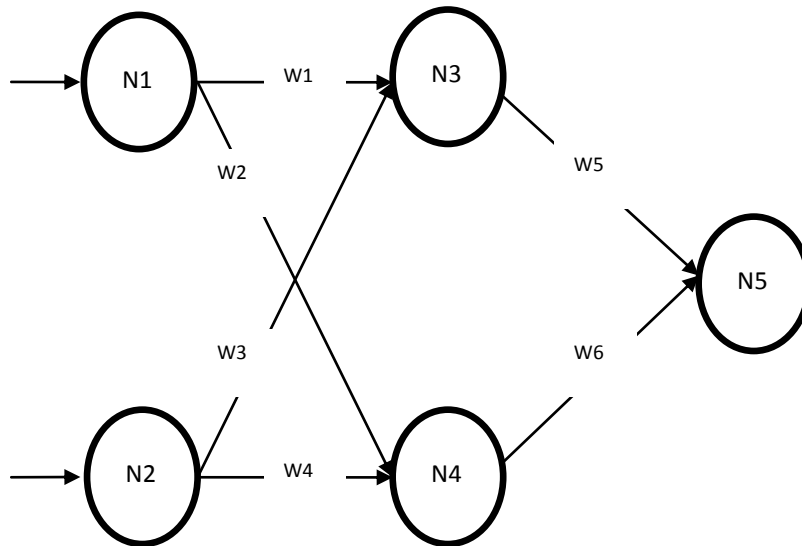


Figura 4.5. Modelo 1 de Red Neuronal Artificial.

2. Modelo 2: Tres capas, dos neuronas en la capa de entrada, tres en la capa oculta y una neurona en la capa de salida (Figura 4.6).

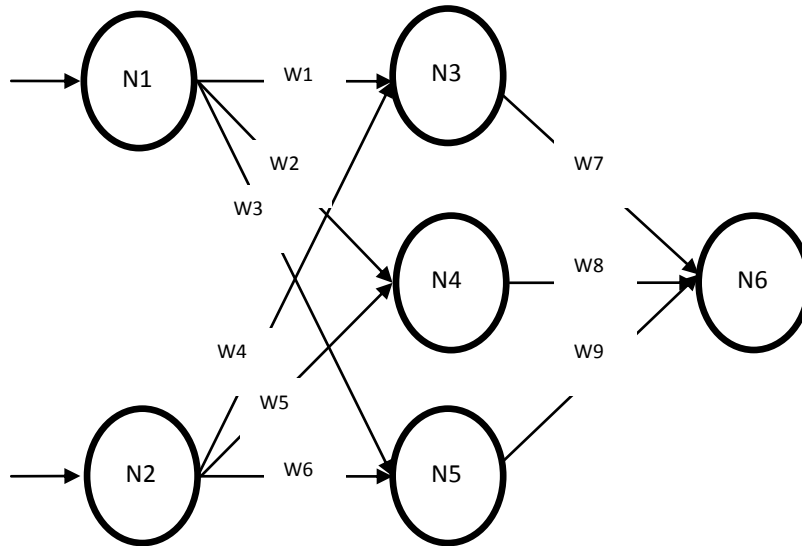


Figura 4.6. Modelo 2 de Red Neuronal Artificial.

3. Modelo 3: Dos capas, dos neuronas en la capa de entrada y una neurona en la capa de salida (Figura 4.7).

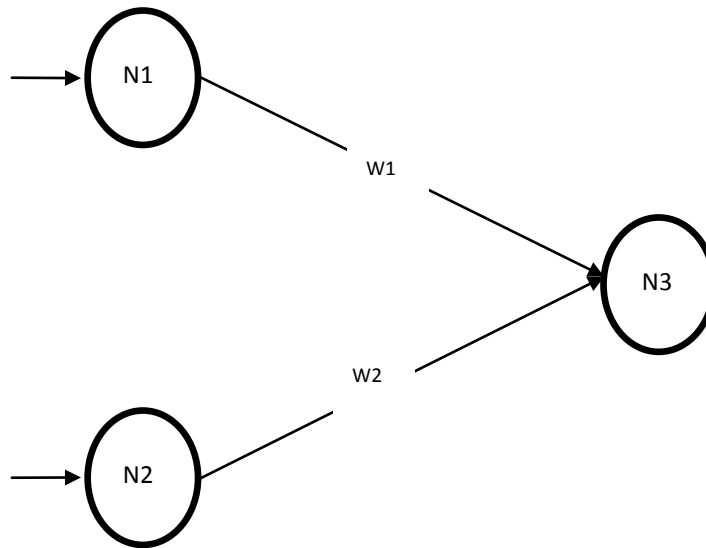


Figura 4.7. Modelo 3 de Red Neuronal Artificial.

Continuando con la etapa de inicialización, como ya se mencionó anteriormente se hace uso de los modelos de red con conexión hacia adelante, ADALINE y BACKPROPAGATION (RETROPROPAGACIÓN). La serie de pasos a seguir son los que proporcionan los algoritmos de aprendizaje de dichos modelos. En este caso para el modelo de red de dos capas se utiliza el algoritmo de la red ADALINE, mientras que para los modelos de tres capas se utiliza el algoritmo de la red de RETROPROPAGACIÓN.

4.2.2.2. *ENTRENAMIENTO.*

En este apartado se desglosa el proceso iterativo de aprendizaje de los algoritmos de red. Para el caso de la red ADALINE la cual es representada por el modelo número 3. Se inicia con la introducción de los dos vectores de entrada , que en este caso serían los datos preparados en la fase anterior, cabe señalar que este proceso de entrenamiento se realizó con la ayuda de una hoja de cálculo, en donde se realizó la programación básica de los algoritmos utilizados. Posteriormente a la introducción de los vectores de entrada es necesario inicializar los pesos con un valor pequeño aleatorio, en este caso se utiliza el valor de 0.5 para cada uno de los pesos. Siguiendo con el algoritmo de aprendizaje se calcula la salida lineal y se calcula el error , se actualizan los pesos mediante la siguiente expresión en donde el parámetro α determina la estabilidad y la velocidad de convergencia del vector de pesos hacia el valor de error mínimo, este parámetro se modifica en un intervalo de valores entre 0 y 1, así, se dice que el modelo de red se encuentra entrenado cuando se observa el valor mínimo de error a diferentes tasas de aprendizaje (Tabla 4.1). Este procedimiento se realiza para cada uno de los modelos diarios propuestos. Es importante mencionar que para la etapa de entrenamiento de los modelos se utilizó el 75% de los datos diarios lo que equivale a las primeras tres semanas del mes mientras que para la etapa de prueba del modelo se utilizó el restante 25% de los datos. El proceso de entrenamiento en el caso de los modelos de tres capas que utilizan el algoritmo de aprendizaje de la red de RETROPROPAGACIÓN es similar al descrito anteriormente, es decir, seguir paso a paso el desarrollo del algoritmo de dicha red. Los resultados de entrenamiento para cada día de la semana se presentan en la Tabla 4.2.

Tabla 4.1. Resumen de corridas realizadas por la red neuronal a diferentes tasas de aprendizaje (Modelo 3, martes)

Tasa	w1	w2	error
0.010000	0.444016	0.117686	0.072202
0.020000	0.483588	0.062952	0.070741
0.030000	0.493752	0.048849	0.070388
0.050000	0.491602	0.044787	0.070438
0.060000	0.488410	0.045279	0.070595
0.070000	0.485073	0.046174	0.070791
0.080000	0.481749	0.047317	0.071014
0.090000	0.478483	0.048667	0.071255
0.100000	0.475287	0.050215	0.071511
0.200000	0.446798	0.076039	0.074572
0.300000	0.420927	0.117659	0.078520
0.400000	0.391976	0.170165	0.083847
0.500000	0.356835	0.229603	0.091449
0.600000	0.315527	0.292623	0.102762
0.700000	0.270789	0.356255	0.120096
0.800000	0.227366	0.418107	0.147096
0.900000	0.190911	0.476447	0.189314
1.000000	0.166739	0.530033	0.254840

Tabla 4.2 Resultados de error de entrenamiento para cada día de la semana

Día	ERROR DE ENTRENAMIENTO		
	Modelo 1	Modelo 2	Modelo 3
LUNES	0.02861	0.01942	0.04463
MARTES	0.07153	0.03397	0.07039
MIÉRCOLES	0.06278	0.05650	0.10122
JUEVES	0.04365	0.05993	0.08055
VIERNES	0.02573	0.03744	0.08917
SÁBADO	0.01000	0.01248	0.05711
DOMINGO	0.01520	0.01608	0.04336

4.2.2.3. VALIDACIÓN

Para realizar la validación de los modelos es necesario realizar el pronóstico fuera de la muestra, es decir, utilizar los valores obtenidos en la etapa de entrenamiento y aplicarlos al 25% de datos de las series de tiempo que se guardaron para dicho propósito.

4.2.2.3.1 MÉTODO ITERATIVO

Existe una técnica intuitiva, denominada método iterativo, la cual, ha sido utilizada por varios investigadores para generar pronósticos [24, 25]. Un ejemplo de ésta es asumir un modelo de pronóstico con dos retardos, como lo muestra la siguiente expresión:

Para estimar dos periodos adelante, el modelo seria:

De la misma forma para tres periodos se tiene:

Es común usar la técnica anterior en situaciones en donde el retardo domina el horizonte de pronóstico sobre otras correlaciones pasadas. Para ejemplificar el pronóstico de manera cualitativa se muestra la Figura 4.8, en donde, se aprecia la previsión de cada modelo para el día lunes.

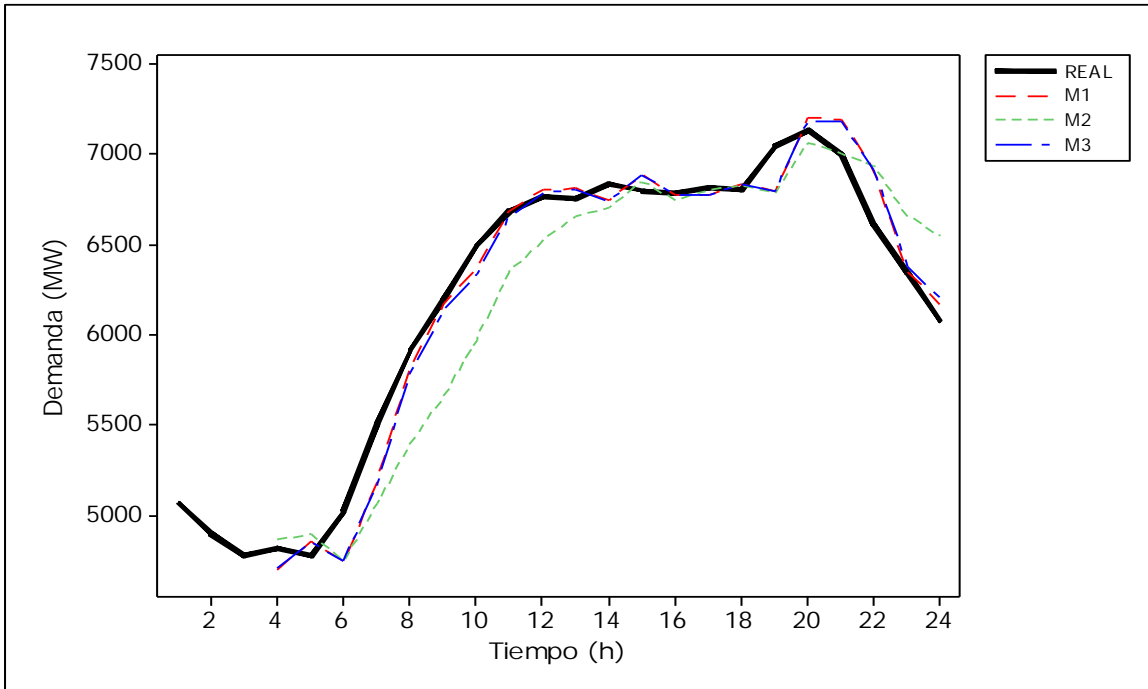


Figura 4.8. Pronóstico fuera de la muestra (Lunes)

En la Figura 4.8 se observa el resultado de realizar la previsión de la demanda de energía eléctrica para el día lunes. Es evidente que los modelos propuestos siguen el comportamiento de la serie de manera similar, destacando el modelo 1 el cual, consta de 5 neuronas, dos en la capa de entrada, dos en la capa intermedia y una en la capa de salida.

Para observar la validación de manera cuantitativa es necesario hacer uso del porcentaje de error medio absoluto (PEMA). Si O_t es la observación actual para un periodo de tiempo t , y P_t es el pronóstico para el mismo periodo, entonces el error del pronóstico está definido como:

Para definir el PEMA, primero se define el porcentaje de error relativo como:

$$e_t = \frac{O_t - P_t}{O_t} \times 100$$

Entonces el PEMA queda:

$$PEMA = \frac{1}{n} \sum_{t=1}^n |e_t|$$

Tabla 4.3. Exactitud en el pronóstico generada por los modelos (Lunes)

MODELO	PEMA
1	1.590 %
2	3.274 %
3	1.687 %

Como se observó en la Figura 4.8 el modelo 4.1 muestra una mayor sensibilidad para ajustarse a la curva real. Este modelo cuenta con la medida estadística de error más pequeña.

4.2.2.4. *RESULTADOS*

En la presente etapa se observan los resultados obtenidos para cada día de la semana, se ofrecen los valores de pesos obtenidos en la fase de entrenamiento (Figura 4.9), para posteriormente en la fase de análisis de resultados llevar a cabo el pronóstico fuera de la muestra y realizar los comentarios oportunos.

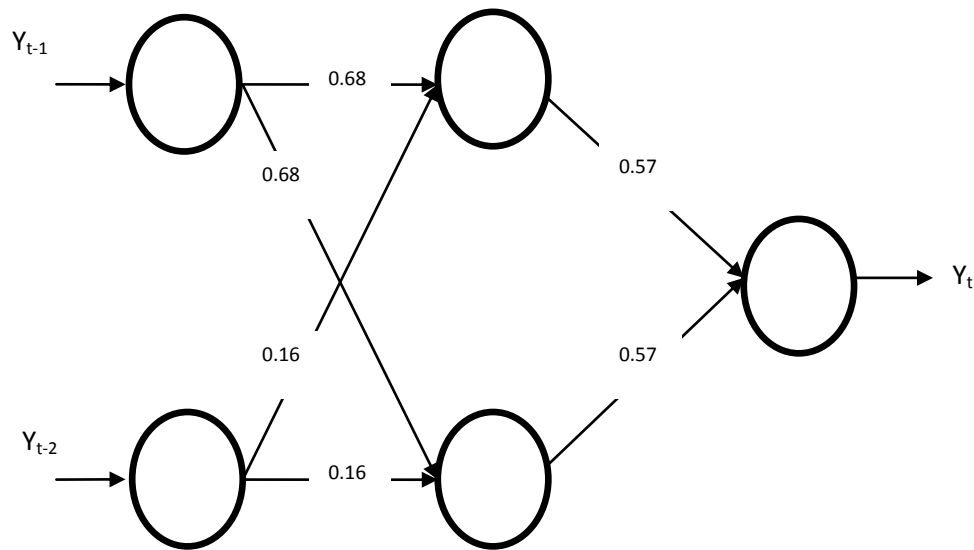
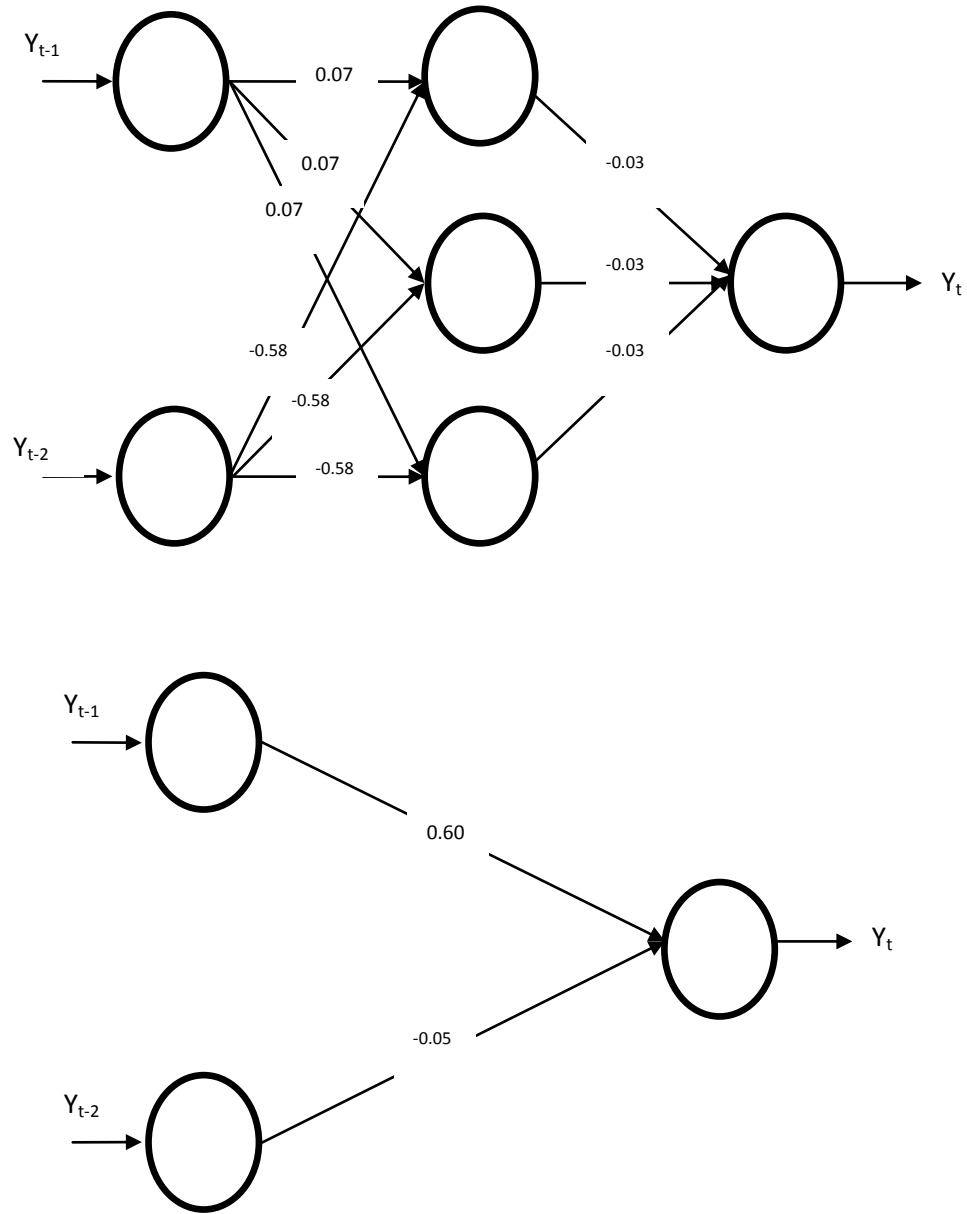


Figura 4.9. Resultados de los modelos propuestos para cada día de la semana (Lunes).



LUNES

Figura 4.9. Resultados de los modelos propuestos para cada día de la semana (Continuación.)

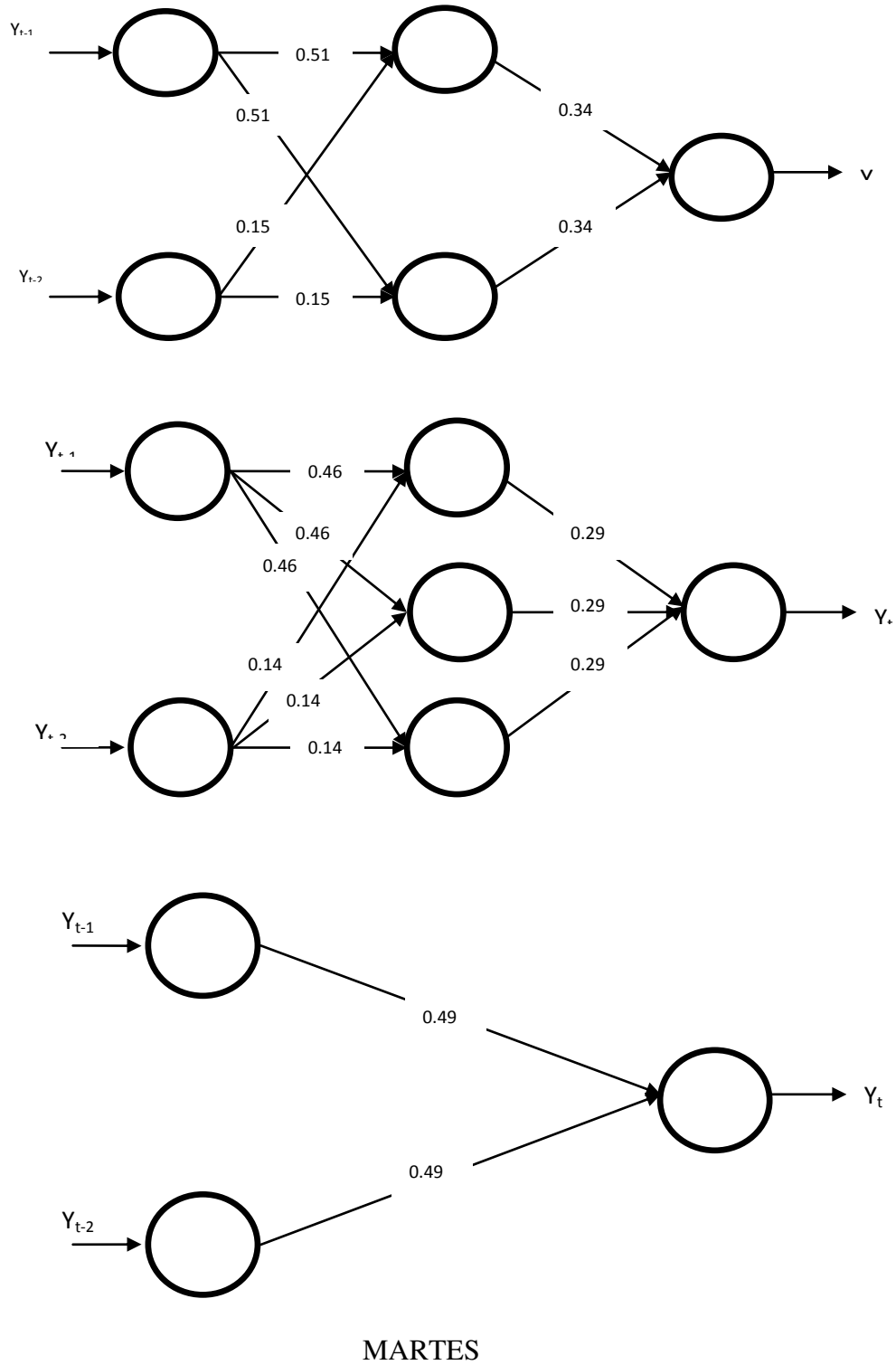


Figura 4.10. Resultados de los modelos propuestos para cada día de la semana (Martes).

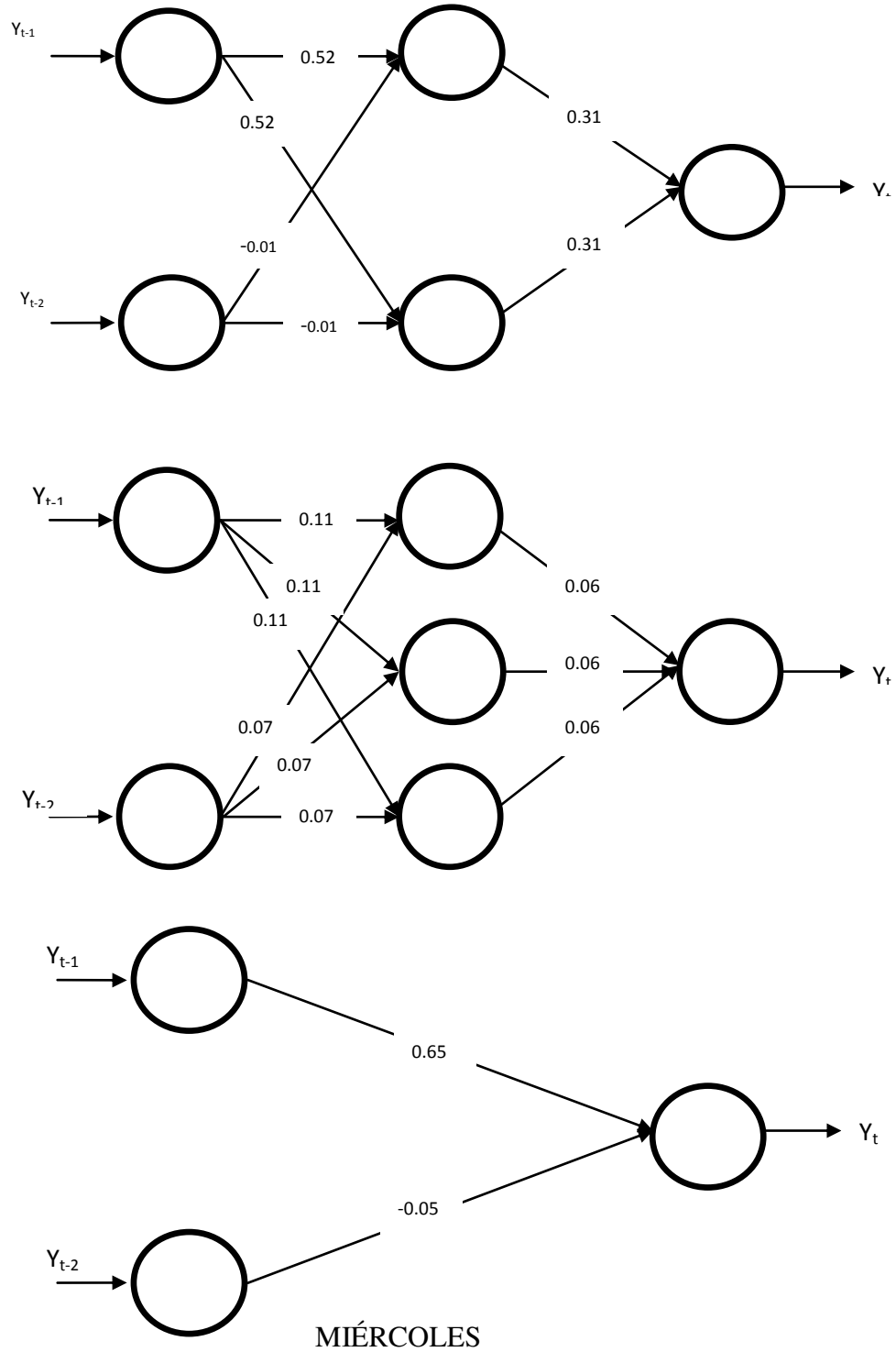


Figura 4.11. Resultados de los modelos propuestos para cada día de la semana (Miércoles.)

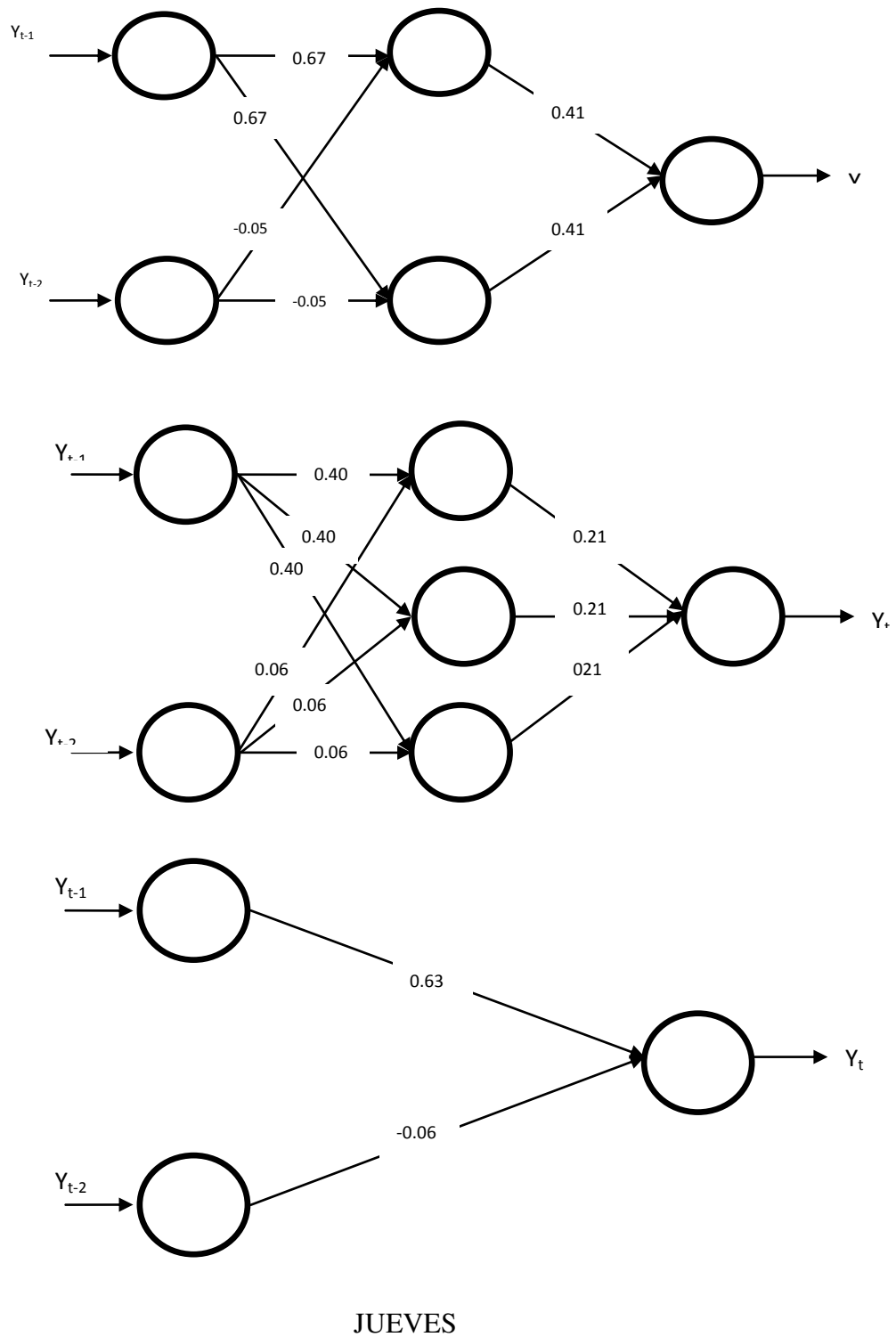


Figura 4.12. Resultados de los modelos propuestos para cada día de la semana (Jueves.).

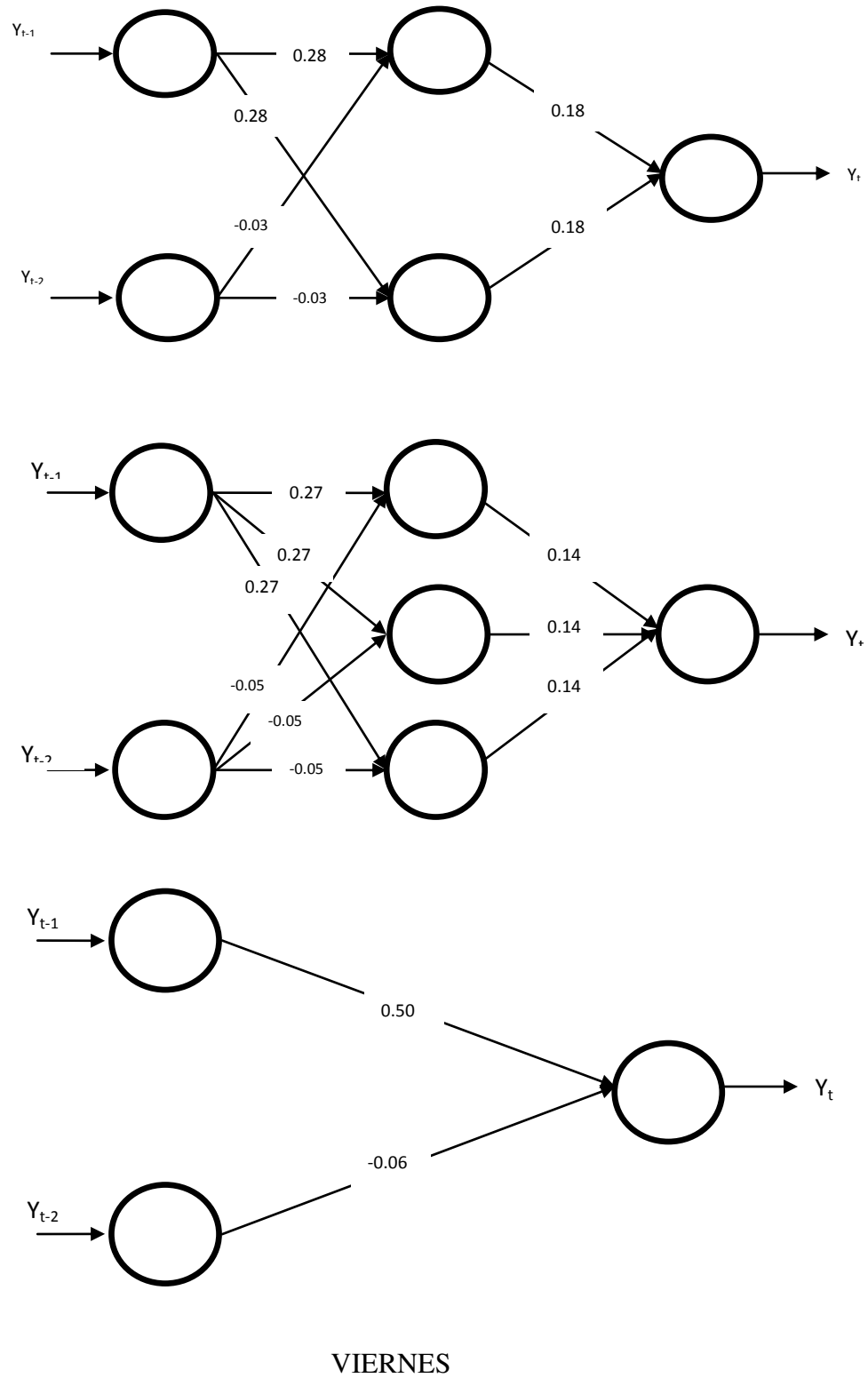


Figura 4.13. Resultados de los modelos propuestos para cada día de la semana (Viernes.).

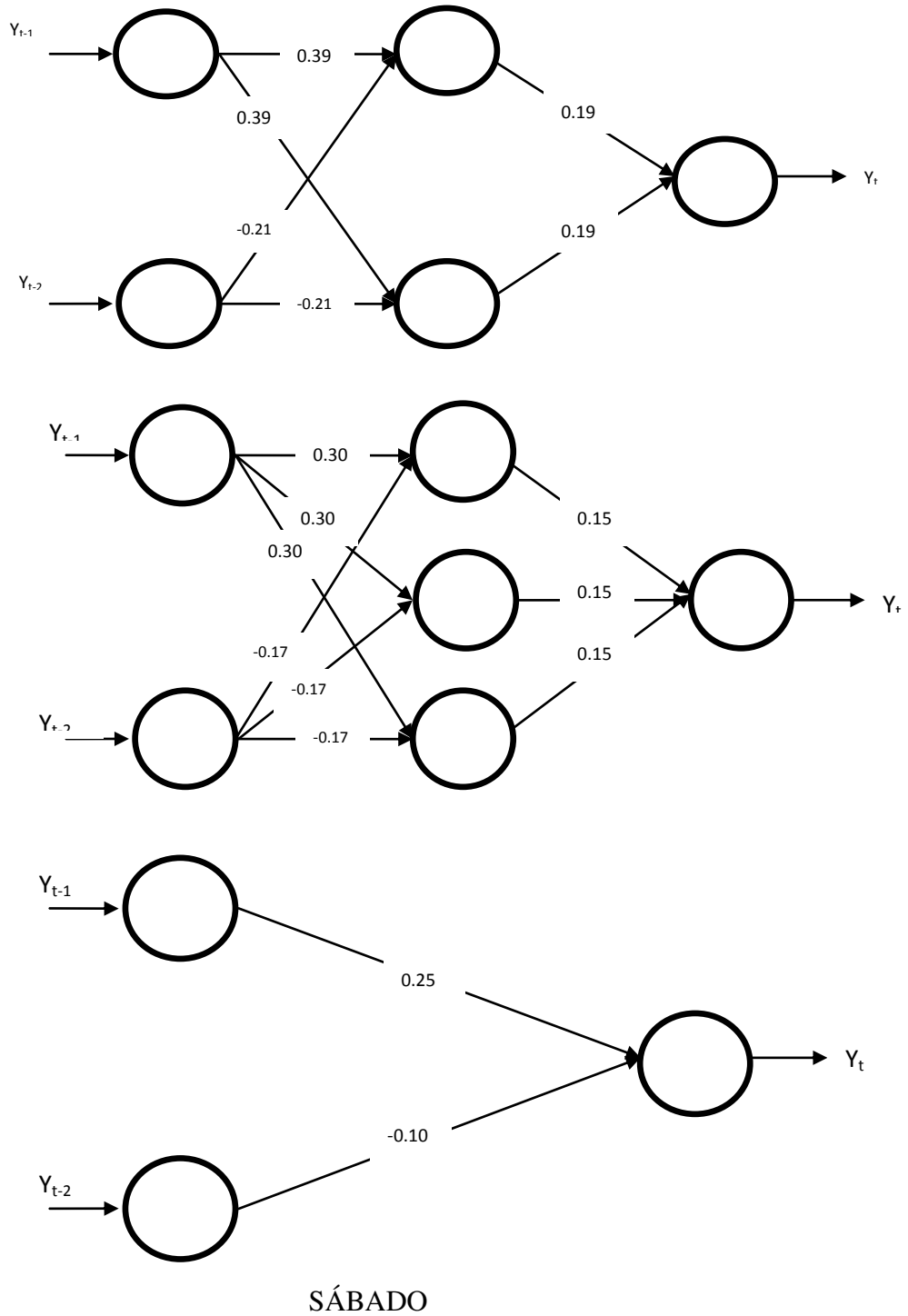


Figura 4.14. Resultados de los modelos propuestos para cada día de la semana (Sábado.).

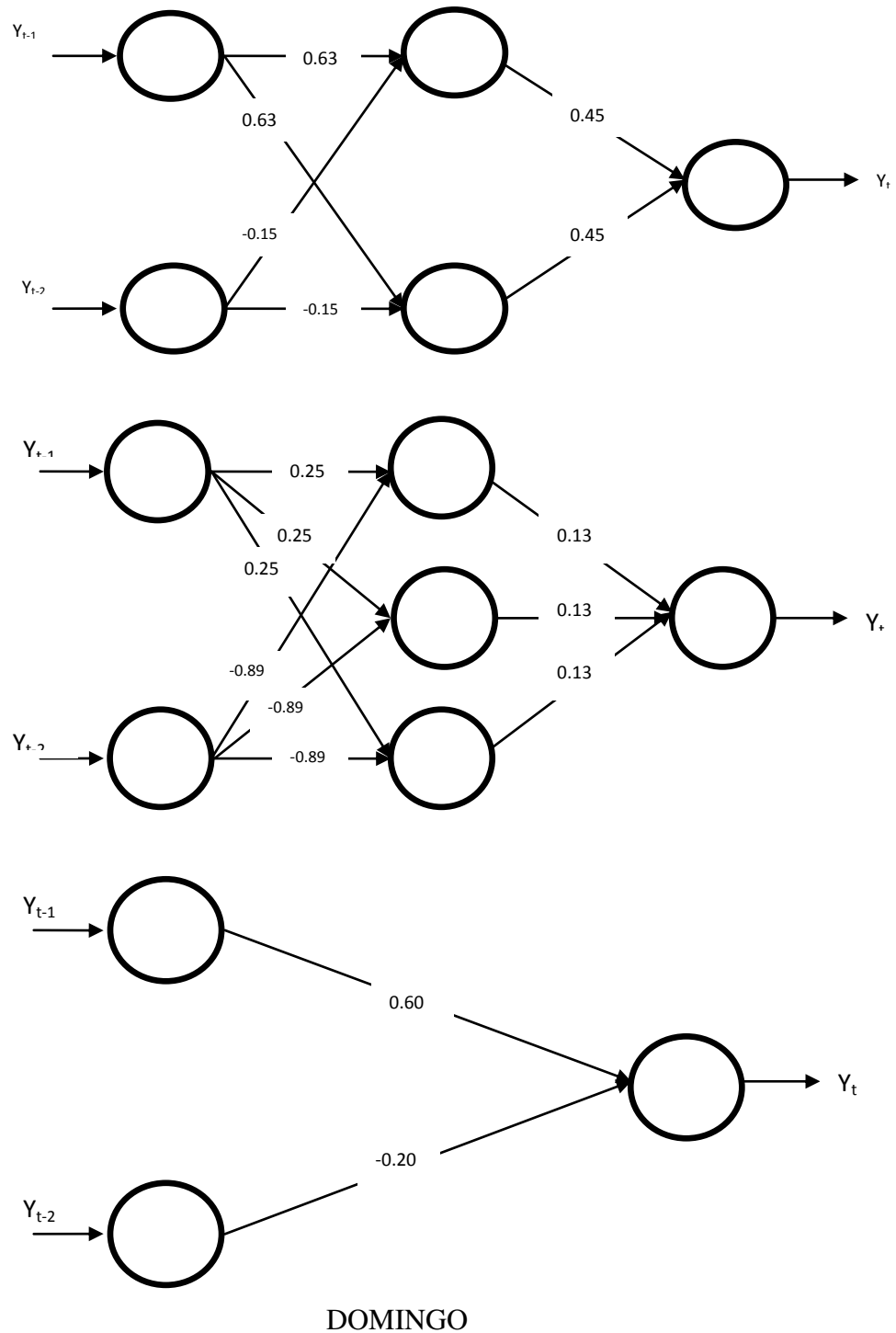


Figura 4.15. Resultados de los modelos propuestos para cada día de la semana (Domingo.).

4.2.3. ANÁLISIS DE RESULTADOS

En la presente etapa, se muestran y se analizan los resultados obtenidos de la validación de cada uno de los modelos para cada día de la semana, el día lunes será omitido en esta sección ya que sus resultados fueron observados para ejemplificar una de las fases anteriores.

Para el día martes se tiene:

Tabla 4.4. Exactitud en el pronóstico generada por los modelos (Martes)

MODELO	PEMA
1	2.396 %
2	2.291 %
3	2.357 %

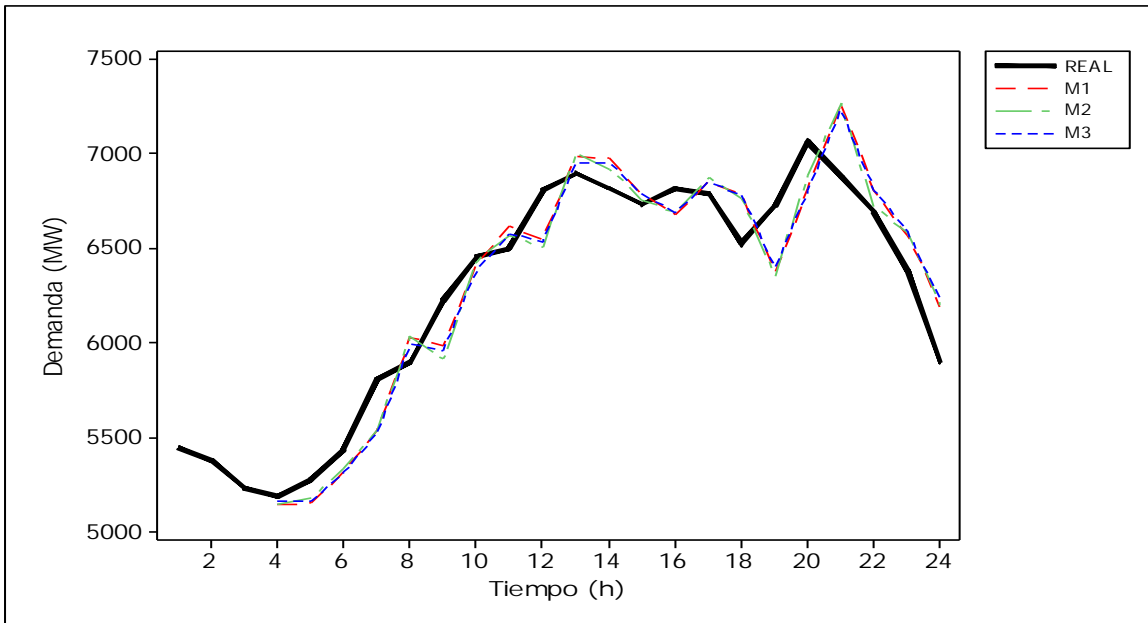


Figura 4.16. Pronóstico fuera de la muestra (Martes).

El comportamiento de las series de los diferentes modelos es muy similar como se observa en la Figura 4.16. A diferencia de la serie del día lunes, esta serie presenta cambios bruscos a lo largo del día, por lo cual, se puede decir que es un patrón difícil de predecir. Sin embargo, el error en el pronóstico de los modelos es bajo, resaltando el modelo 2 que consta de tres capas, con dos neuronas en la capa de entrada, tres en la capa intermedia y una en la capa de salida.

Para el día miércoles se tiene:

Tabla 4.5. Exactitud en el pronóstico generada por los modelos (Miércoles)

MODELO	PEMA
1	1.340 %
2	1.370 %
3	1.372 %

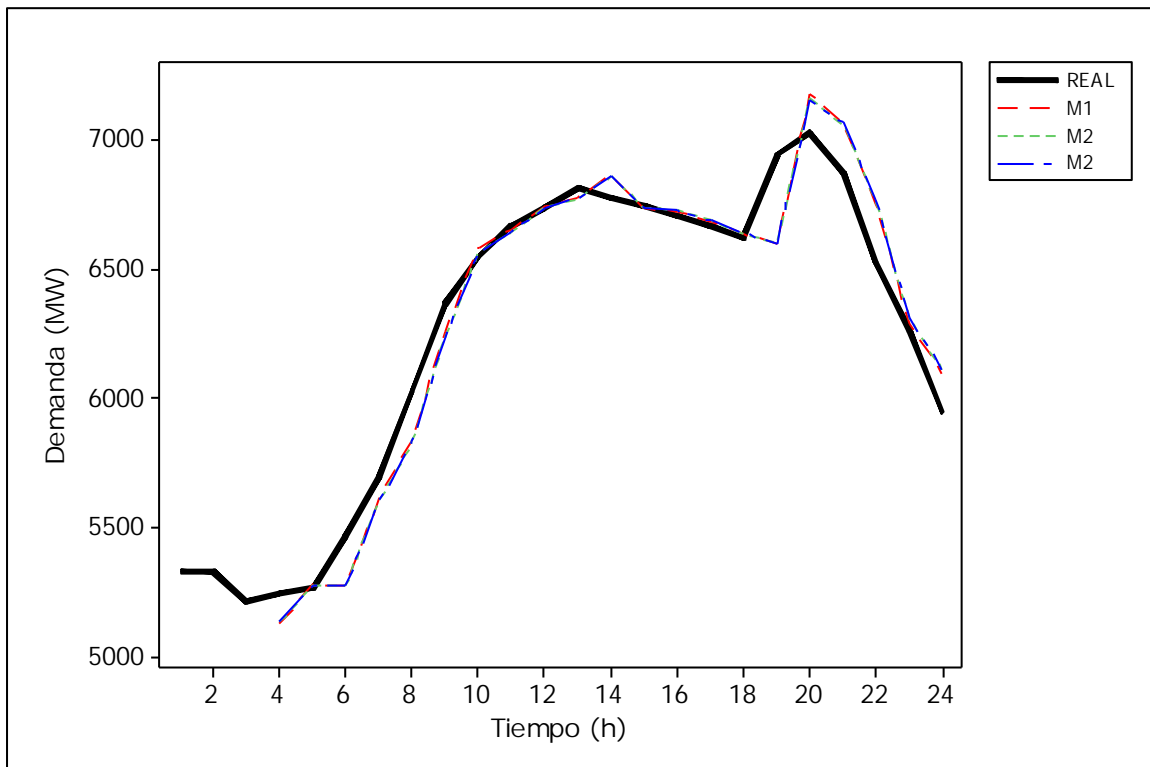


Figura 4.17. Pronóstico fuera de la muestra (Miércoles).

La respuesta de los modelos propuestos para el día miércoles es bastante similar como se observa de manera cualitativa en la Figura 4.17, mientras que de manera cuantitativa la Tabla 4.5 muestra la similitud de los errores en el pronóstico, siendo el modelo número 1 el que menor error presenta.

Para el día jueves se tiene:

Tabla 4.6. Exactitud en el pronóstico generada por los modelos (Jueves)

MODELO	PEMA
1	1.828 %
2	1.860 %
3	1.904 %

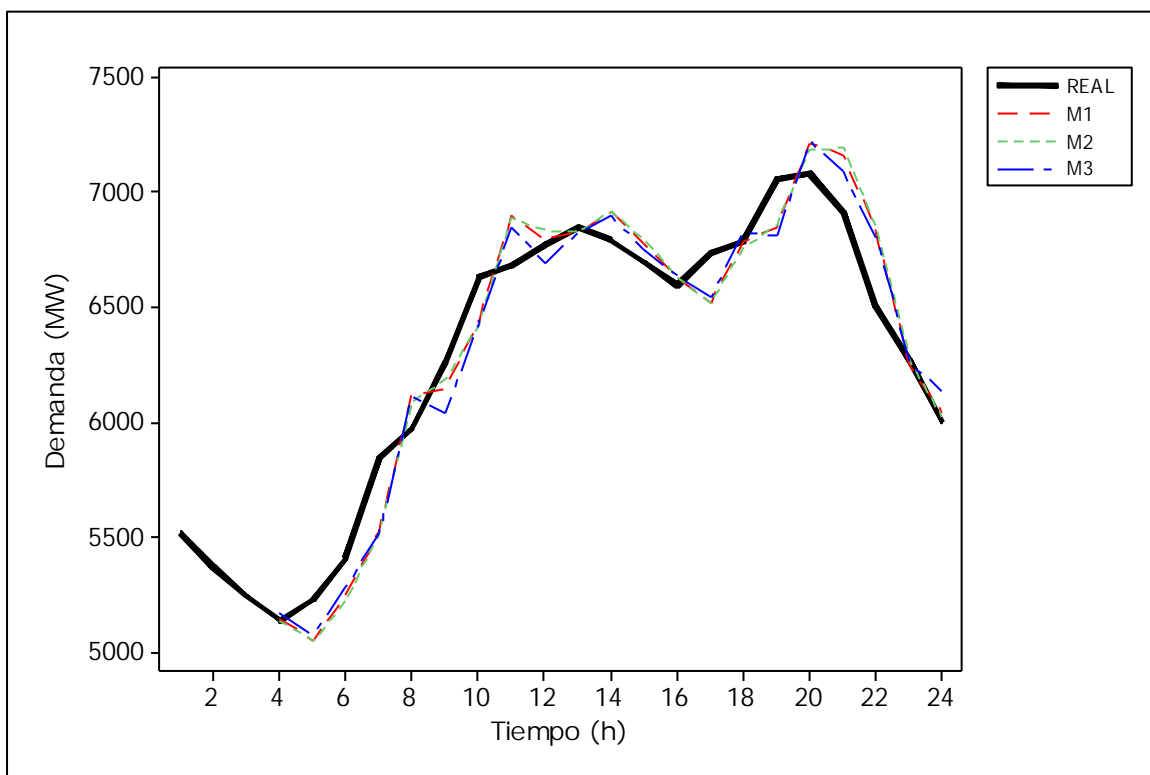


Figura 4.18. Pronóstico fuera de la muestra (Jueves).

La Figura 4.18 muestra el pronóstico realizado con los modelos de red propuestos, donde, nuevamente se presenta un patrón complicado. La Tabla 4.6 indica que el modelo con menor error en la fase de prueba es el modelo numero 1.

Para el día viernes se tiene:

Tabla 4.7. Exactitud en el pronóstico generada por los modelos (Viernes)

MODELO	PEMA
1	1.611 %
2	1.714 %
3	1.634 %

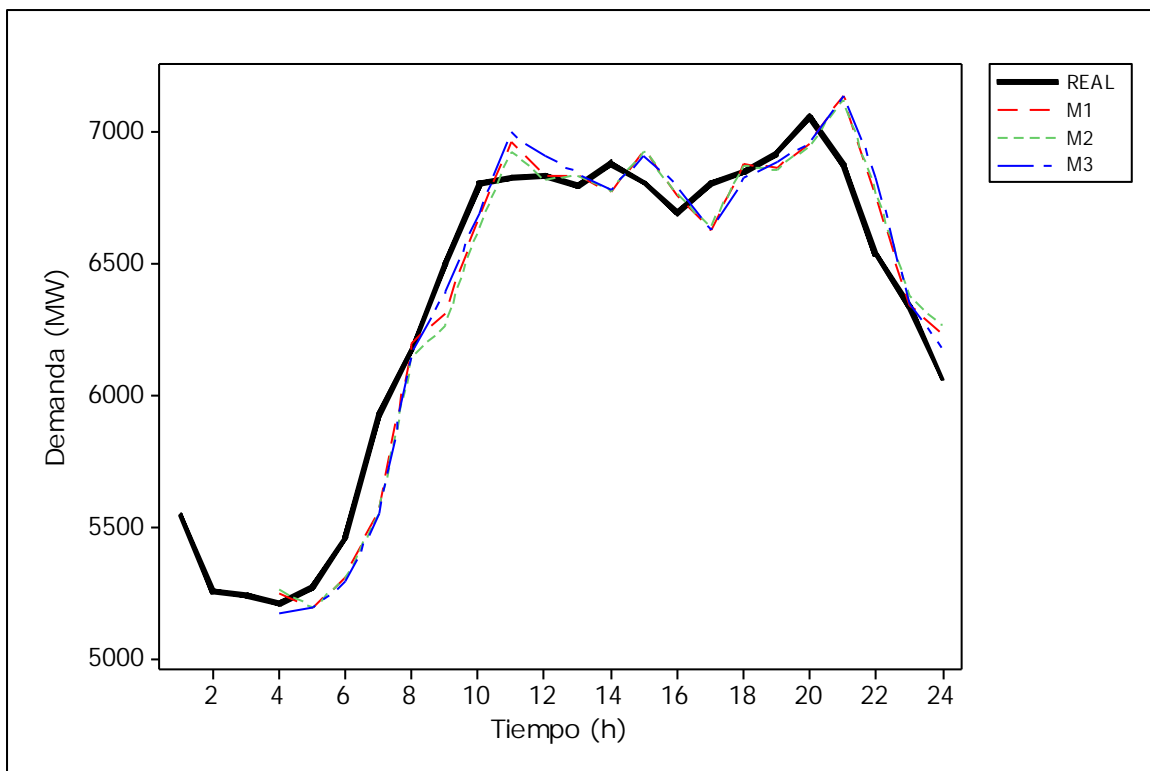


Figura 4.19. Pronóstico fuera de la muestra (Viernes).

El patrón de la serie para el día viernes es un patrón más estable, lo que permite un menor error en la etapa de pronóstico, se observa en la Figura 4.19 como los tres modelos acompañan al patrón de la serie de buena manera. La Tabla 4.7 revela al modelo número 1 como el modelo que presenta la mayor exactitud en el pronóstico.

Para el día sábado se observa:

Tabla 4.8. Exactitud en el pronóstico generada por los modelos (Sábado).

MODELO	PEMA
1	2.908 %
2	2.888 %
3	2.901 %

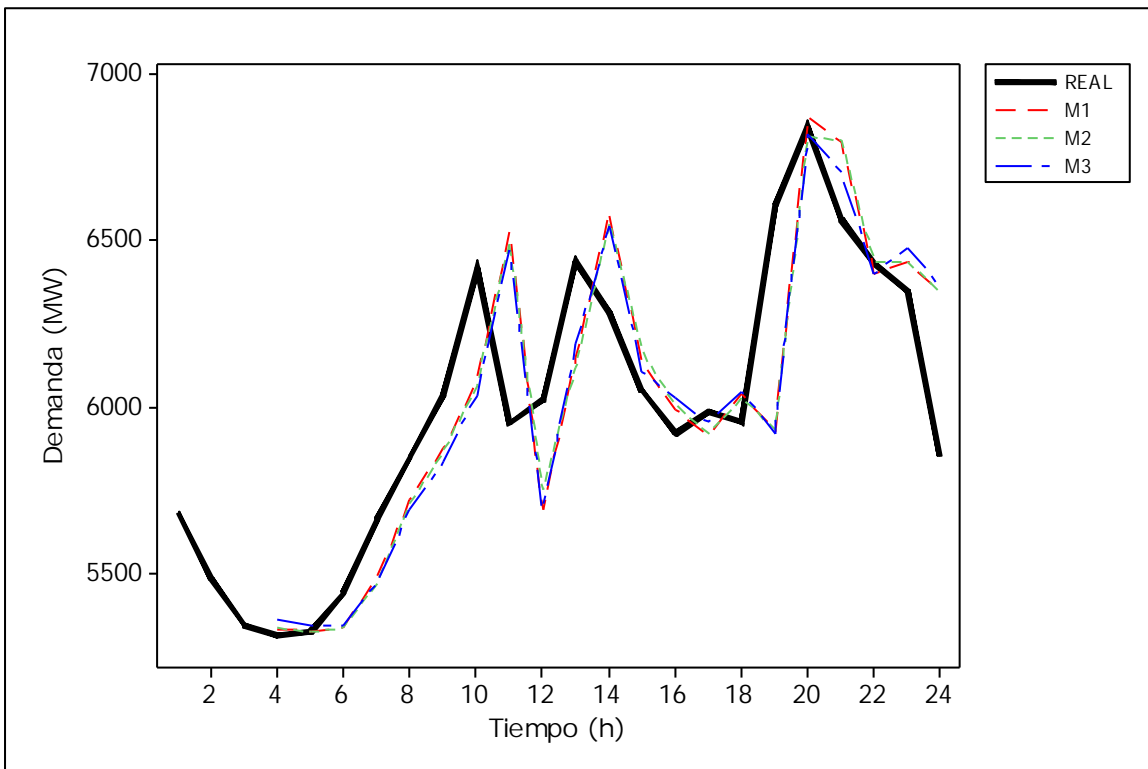


Figura 4.20. Pronóstico fuera de la muestra (Sábado).

Los resultados que arrojan los modelos propuestos para el día sábado al ser evaluados cualitativamente mediante la Figura 4.20 muestran un mayor desfase entre el patrón de la serie real y el pronosticado, como ya se ha comentado el comportamiento inestable de los patrones reduce la exactitud en el pronóstico de la redes neuronales. La Tabla 4.8 muestra los resultados de manera cuantitativa, en este caso el modelo numero 2 presenta la mayor exactitud del pronóstico.

Para el día domingo se tienen los siguientes resultados:

Tabla 4.9. Exactitud en el pronóstico generada por los modelos (Domingo).

MODELO	PEMA
1	3.311 %
2	3.314 %
3	3.331 %

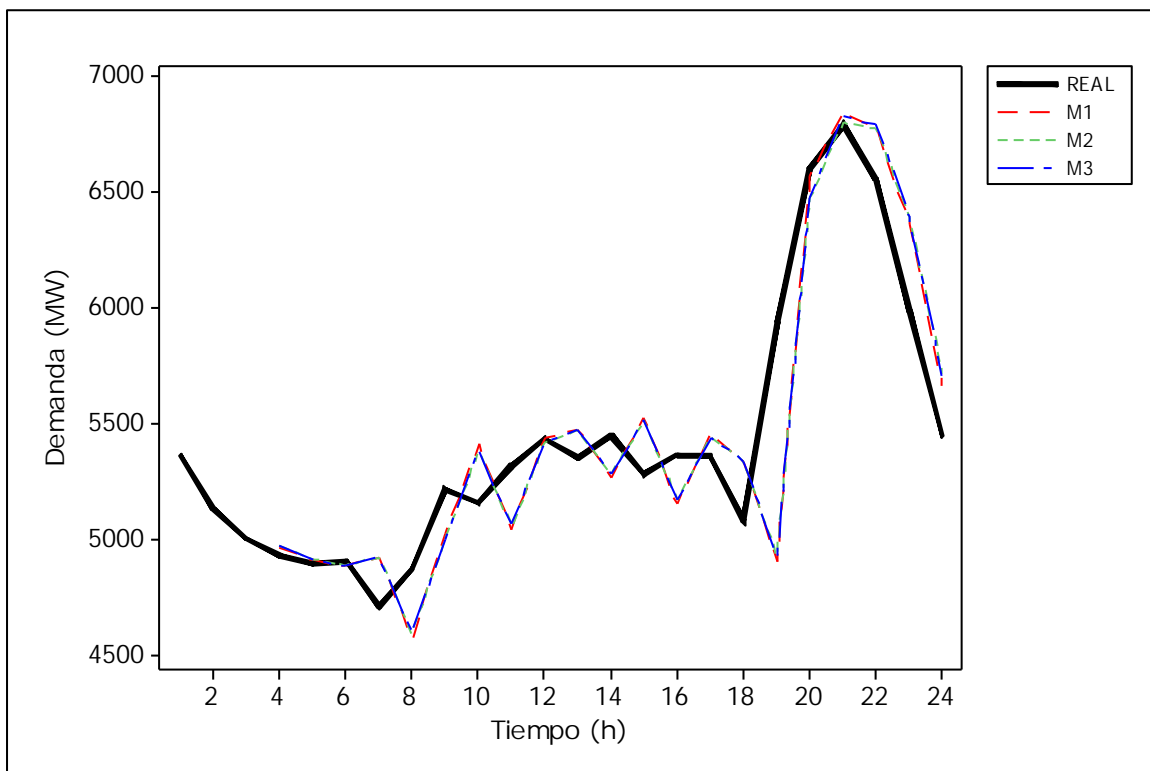


Figura 4.21. Pronóstico fuera de la muestra (Domingo).

El comportamiento de la serie del día domingo es realmente especial. Lo importante a destacar de este resultado es como los modelos de red tratan de seguir el patrón tan extraño de esta serie (Figura 4.21). En la Tabla 4.9 se muestra que el modelo numero 1 nuevamente es la mejor opción debido a que el valor de error en el pronóstico es menor para este modelo.

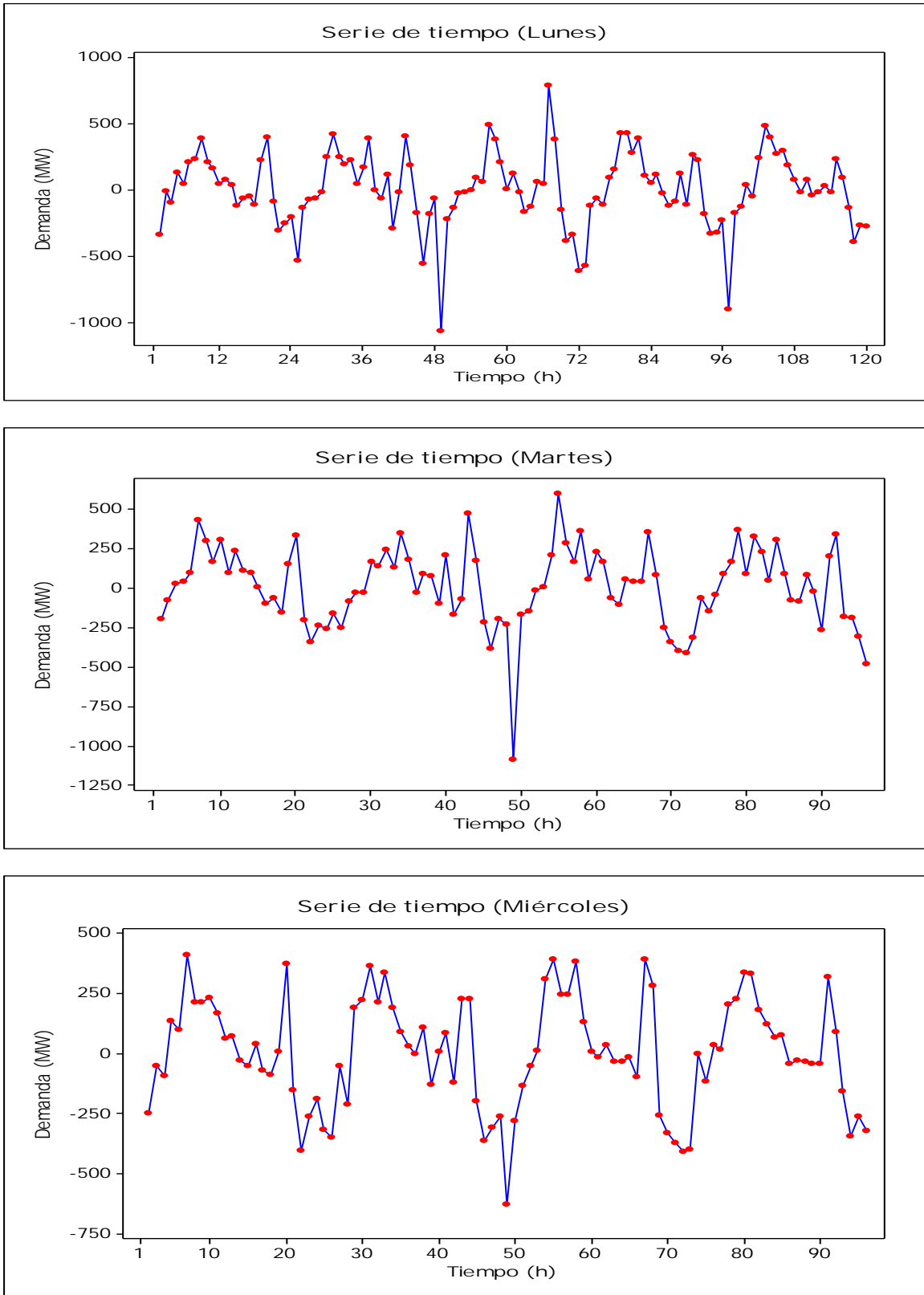


Figura 4.2. (Continuación) Serie de tiempo diferenciada (Lunes, Martes y Miércoles).

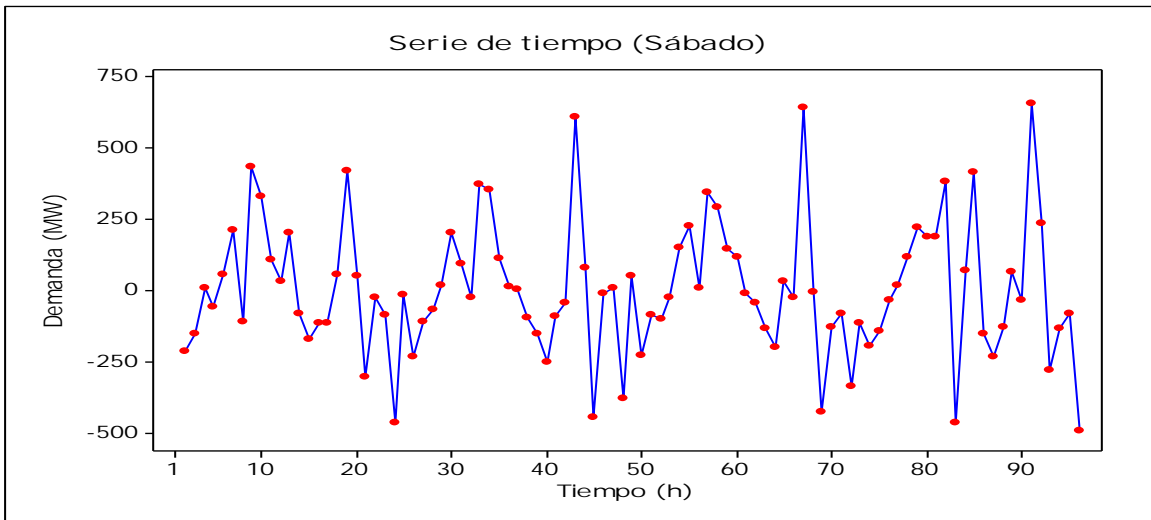
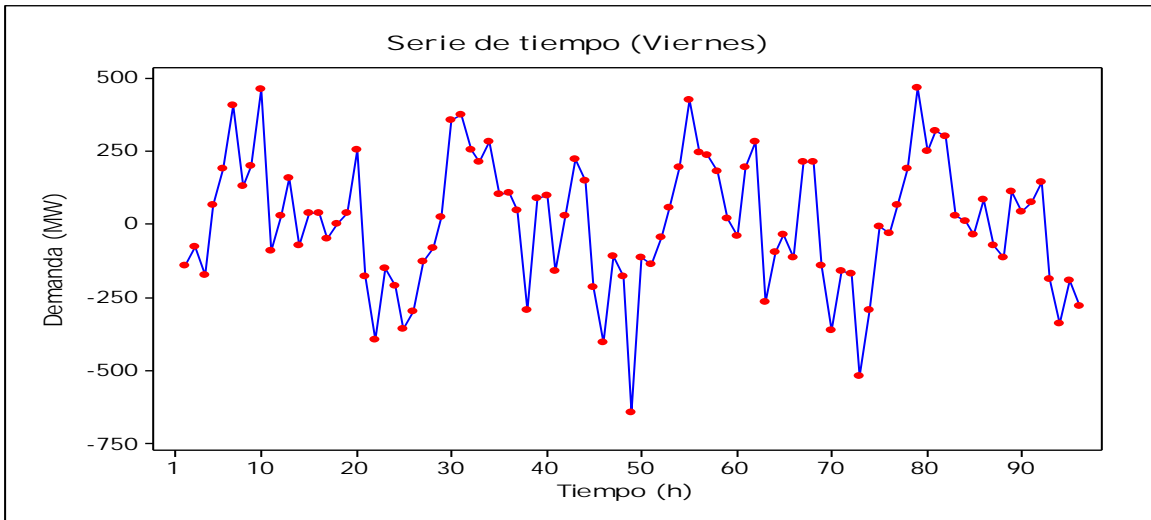
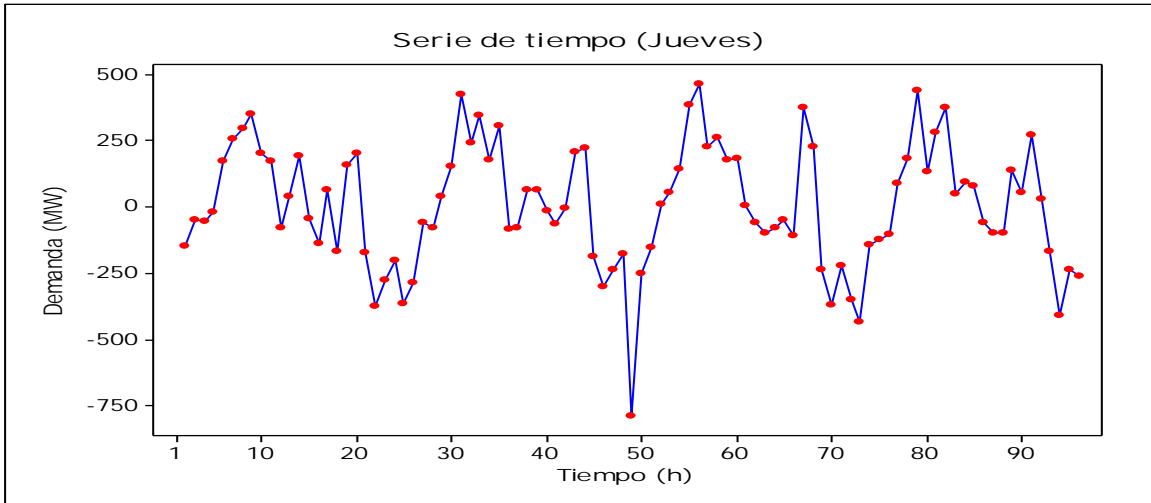


Figura 4.2. (Continuación) Serie de tiempo diferenciada (Jueves, Viernes y Sábado).

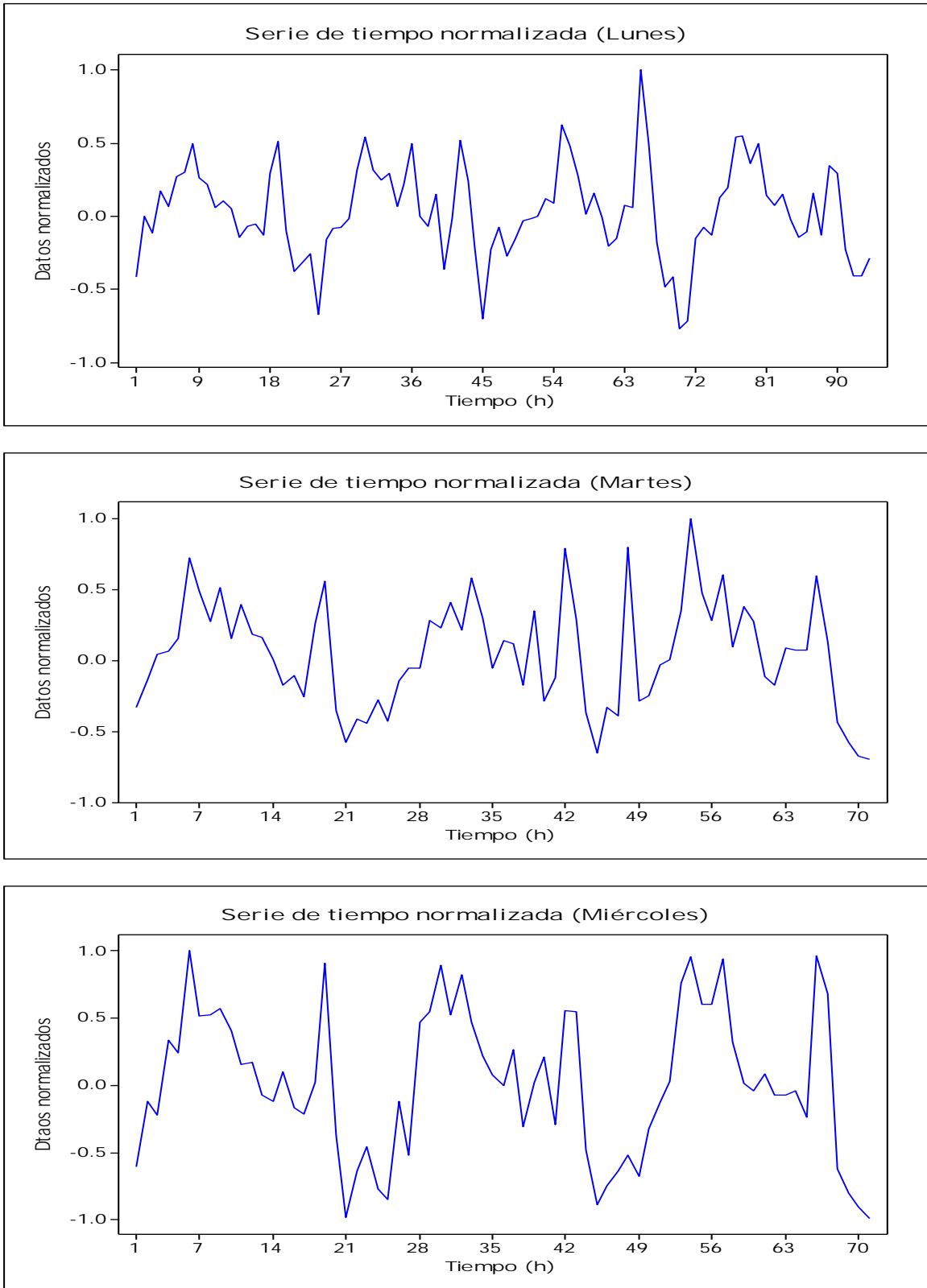


Figura 4.3. (Continuación) Serie de tiempo normalizada (Lunes, Martes y Miércoles).

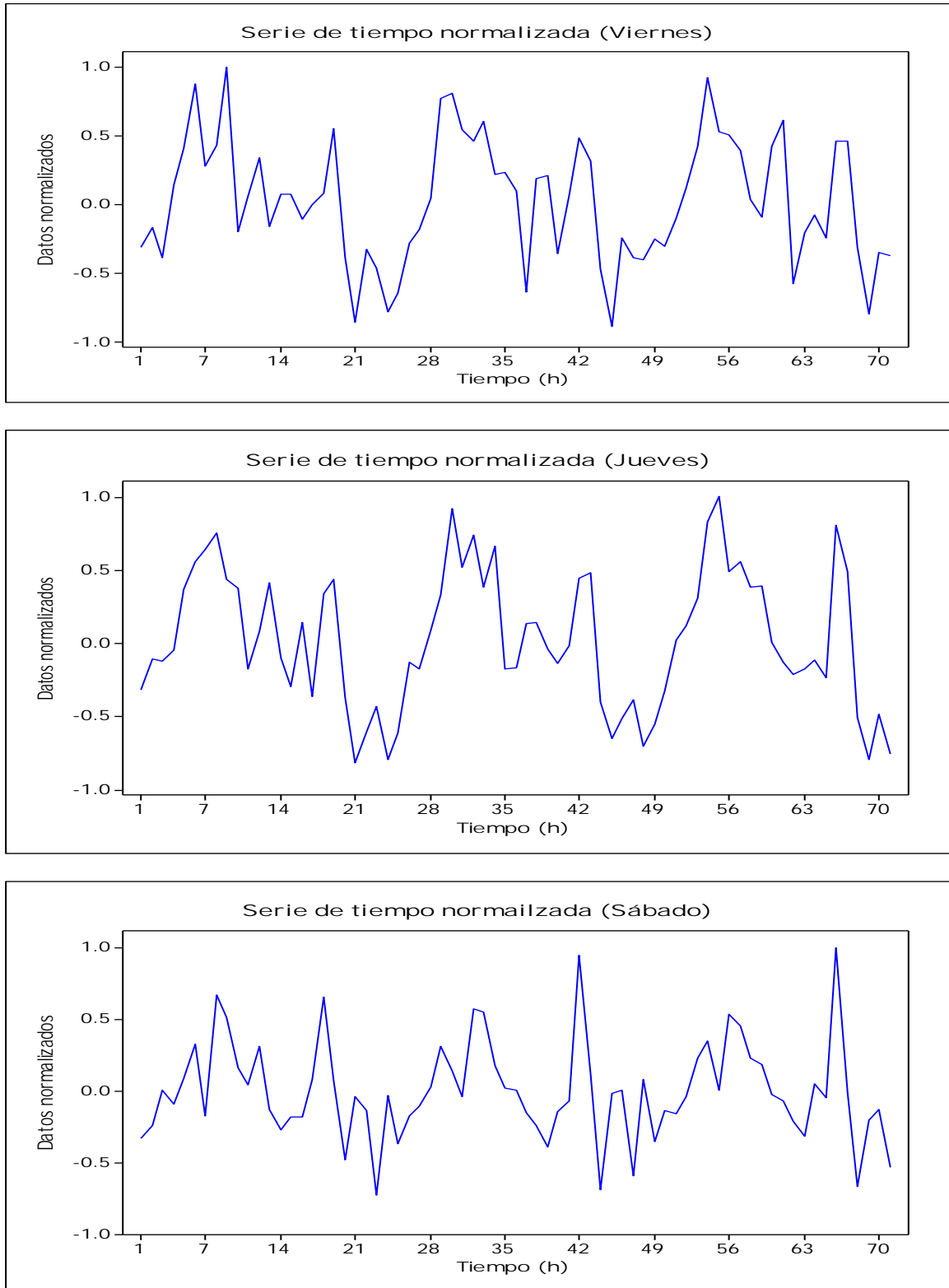


Figura 4.3. (Continuación) Serie de tiempo normalizada (Jueves, Viernes y Sábado).

CONCLUSIONES.

Realizar detalladamente el proceso de modelado de series de tiempo mediante redes neuronales conlleva a tener resultados satisfactorios. En este caso es importante mencionar que para llegar a los resultados expuestos, en un primer intento de realizar el pronóstico de la demanda de energía eléctrica se hizo el modelado sin realizar la caracterización de las series diarias, es decir, no se analizaron los patrones mencionados de estacionalidad y tendencia, obteniendo resultados con mayor incertidumbre, todo esto guiados de alguna manera por la revisión de distintos artículos relacionados con el tema en donde no se hace mención de la influencia de dichos patrones en los resultados.

Por lo anterior es de gran relevancia concluir que la preparación de los datos que fueron utilizados en el modelado de la Red Neuronal Artificial de acuerdo con los patrones que presentaron, permitió obtener los resultados esperados para el presente proyecto debido a que, la calidad de los valores de entrada influyó de manera concreta, permitiendo facilitar la tarea de análisis.

De acuerdo a los datos proporcionados por la compañía eléctrica, el modelo que permite una menor incertidumbre en el pronóstico a corto plazo es el que incluye la siguiente configuración: tres capas, dos neuronas en la capa de entrada, dos neuronas en la capa intermedia y una neurona en la capa de salida. Aunque, los modelos propuestos restantes son capaces de acercarse e incluso uno de ellos (modelo 2) superar en algunos casos al modelo propuesto, a lo que se puede concluir que es posible utilizar ambos modelos en la previsión de la demanda de electricidad.

Finalmente es preciso decir que se cumplió con los objetivos que se mencionan en un principio, teniendo resultados que compiten e incluso son mejores que algunos expuestos en diversos foros internacionales, tales como los presentados en la revisión bibliográfica del presente proyecto, es importante mencionar que dicha bibliografía solo hace énfasis en la estructura de los modelos de red omitiendo la importancia de realizar una adecuada preparación de la serie de tiempo utilizada.

BIBLIOGRAFÍA

- [1] CFE. *Información básica de control de demanda eléctrica*.
- [2] G. Mbamalu, E. El-Hawary. *A decomposition approach to forecasting electric power system commercial load using an artificial neural network*. 1994, IEEE.
- [3] M. Kazeminejad, M. Dehghan, M.B. Motamadinejad, H. Rastegar. *A new short term load forecasting using multilayer perceptron*. 2006, IEEE.
- [4] M. Hayati. *Short term forecasting using artificial neural networks for the west of Iran*. 2007, IEEE, pp. 1528-1588.
- [5] W. Dai, P. Wang. *Application of pattern recognition and artificial neural network to load forecasting in electric power system*. Third International Conference on Natural Computation, 2007, IEEE.
- [6] S. Mishra, S. K. Patra. *Short term load forecasting using a neural network trained by a hybrid artificial immune system*. Region 10 Colloquium and the Third International Conference on Industrial and Information Systems, Kharagpur, INDIA December 8 -10, 2008, IEEE.
- [7] A. Jain, B. Satish. *Clustering based short term load forecasting using artificial neural network*. 2009, IEEE.
- [8] A. P. Rewagad, V. L. Soanawane. *Artificial neural network based short term load forecasting*. 1998, IEEE.
- [9] J.K. Mandal, A.K. Sinha, G. Parthasarathy. *Application of recurrent neural network for short term load forecasting in electric power system*, IEEE.
- [10] Yu-Jun He, You-Chan Zhu, Dong-Xing Duan, Wei Sun. *Application of neural network model based on combination of fuzzy classification and input selection in short term load forecasting*. Fifth International Conference on Machine Learning and Cybernetics, Dalian, 13-16 August 2006, IEEE.
- [12] J. Murillos, *Estudio del pronóstico de la demanda de energía eléctrica, utilizando modelos de series de tiempo*. Scientia et Technica Año IX, No 23, Diciembre 2003.

- [13] S.M. Al-Alawi, S.M. Islam. *Principles of electricity demand forecasting, part 1, applications (tutorial)*, Power engineering journal, 1996.
- [14] C. M. González. *Predicción de la demanda eléctrica horaria mediante redes neuronales artificiales*. Universidad de Oviedo.
- [15] C.W. Gellins. *Demand forecasting for electric utilities*. Published by the Fairmont press, inc. 1992.
- [16] S. Makridakis, S. Wheelwright. *Métodos de pronóstico*. Limusa, 2000.
- [17] D. Peña. *Análisis de series temporales*. Alianza, 2005.
- [18] M. Arellano. *Introducción al análisis de series de tiempo*, (en línea) 5campus.com, Estadística,(2001), <http://www.5campus.com/leccion/seriest>.
- [19] S.P. Verma. *Estadística básica para el manejo de datos experimentales: aplicación en geoquímica*. UNAM, 2005.
- [20] J.R. Hiler González. *Redes neuronales artificiales*. Alfaomega, 2000.
- [21] J.A. Anderson. *Redes neurales*, Alfaomega. 2007.
- [22] B.M del Brío, A.S. Molina. *Redes neuronales y sistemas borrosos*. Alfaomega. 2007.
- [23] Lean YU, Shouyang WANG and Kin Keung LAI. *Foreign-exchange-rate forecasting with artificial neural networks*. Stanford university, 2007.
- [24] N. K. Kasabov. *Foundations of neural networks, fuzzy systems, and knowledge engineering*. Massachusetts Institute of Technology, 1998.
- [25] Hwang, H. *Insights Neural Network Forecasting of Time Corresponding to ARMA (p,q) structures*, Omega. The International Journal of management Science 2001; 29: 273-89.
- [26] Balkin SD. *The value of nonlinear models in the M3- competition*. *International Journal of Forecasting* 2001;17:545- 6.

- [27] Widrow B, Rumelhart D, Lehr MA. *Neural Networks: applications in industry, business and science*. Communications of the ACM 1994; 37 (3): 93-105
- [28] Peter Zhang G. *Neural Networks in business forecasting*. Georgia State University, USA. IRM Press, ISBN 1-59140-215-8; 2004.
- [29] Zhang GP. *An Investigation of Neural Networks for Linear Time-Series Forecasting Computers & Operations Research* 2001; 28: 1183-202
- [30] Brockwell Pj. Davis RA. *Introduction to Times Series and Forecasting*. Springer : 2002.
- [31] Tang Z. Fishwick P. *Feedforward Neural Nets as Models for Time Series Forecasting ORSA Journal on Computing* 1993; 5(4): 374-85.
- [32] Hill T, Leonory M. O' Connor M, Remus W. *Artificial Neural Networks Models for Forecasting and Decision Making. International Journal Forecasting* 1994; 10: 5-15.
- [33] Master T. *Practical Neural Network Recipes in C++*. London: Academic Press inc.; 1993.
- [34] D.C. Park, M.A. El- Sharkawi, R.J. Marks II, L.E: Atlas, M.J. Damborg. *Electric load forecasting using an artificial neural network*. Transactions on power systems, Vol. 6. 1991. IEEE.
- [35] Takeshi Haida, *Study on Daily Electric Load Curve Forecasting Method based on Regression Type Hourly Load Modeling with Yearly Load Trends, Day-types and Insolations*. Tokyo Electric Power Company. Pp. 1477-1485.
- [36] D. J. García, A. C. Milanes. *Predicción de la demanda eléctrica de corto plazo mediante redes neuronales artificiales*. SIE. 2003.
- [37] G.A. Adepoju, M.Sc., S.O.A. Ogunjuyigbe, M.Sc., K.O. Alawode, B.Tech. *Application of neural network to load forecasting in nigerian electrical power system*. The Pacific Journal of Science and Technology, Volume 8. Number 1. May 2007.

- [38] J. P. Rothe, A. K. Wadhvani , Mrs. S. Wadhvani. *Short term load forecasting using multi parameter regression*. International Journal of Computer Science and Information Security, Vol. 6, No. 2, 2009.
- [39] S. Haykin. *Neural networks a comprehensive foundation*. Prentice Hall International, Inc. 1999.
- [40] A. Oonsivilai , El-Hawary. *Wavelet Neural Network Based Short Term Load Forecasting of Electric Power System Commercial Load*. Canadian Conference on Electrical and Computer Engineering, 1999.
- [41] S. H. Ling, Frank H. F. Leung, H. K. Lam, Peter K. S. Tam. *Short-term electric load forecasting based on a neural fuzzy network*. Transactions on industrial electronics, vol. 50, no. 6, December 2003. IEEE.
- [42] A. F. Barrientos, J. Olaya, V. M. González. *Un modelo spline para el pronóstico de la demanda de energía eléctrica*. Revista Colombiana de Estadística Volumen 30 No. 2. pp. 187 a 202. Diciembre 2007.