

# **ESCUELA POLITÉCNICA NACIONAL**

## **FACULTAD DE INGENIERÍA ELÉCTRICA Y ELECTRÓNICA**

### **PREDICCIÓN DE LA DEMANDA ELÉCTRICA EN BARRA DE 69 kV DEL SISTEMA ELÉCTRICO DE SUBTRANSMISIÓN DE LA EMPRESA ELÉCTRICA RIOBAMBA S.A.**

PROYECTO PREVIO A LA OBTENCIÓN DEL TÍTULO DE INGENIERO ELÉCTRICO

**LEONARDO DAVID QUISIMALÍN VILLACÍS**

daleoqv@gmail.com

DIRECTOR: FRANKLIN LENIN QUILUMBA GUDIÑO, Ph.D.

Franklin.quilumba@epn.edu.ec

**Quito, junio 2017**

## **DECLARACIÓN**

**Yo, Leonardo David Quisimalín Villacís, declaro bajo juramento que el trabajo aquí descrito es de mi autoría; que no ha sido previamente presentado para ningún grado o calificación profesional; y, que he consultado las referencias bibliográficas que se incluyen en este documento.**

**A través de la presente declaración cedo mis derechos de propiedad intelectual correspondientes a este trabajo, a la Escuela Politécnica Nacional, según lo establecido por la Ley de Propiedad Intelectual, por su Reglamento y por la normatividad institucional vigente.**

---

**Leonardo David Quisimalín Villacís**

## **CERTIFICACIÓN**

**Certifico que el presente trabajo fue desarrollado por Leonardo David Quisimalín Villacís, bajo mi supervisión.**

---

**Franklin L. Quilumba Gudiño, Ph.D.**  
**DIRECTOR DEL PROYECTO**

## **AGRADECIMIENTO**

A mis padres José Luis Quisimalín Miniguano y René Fabiola Villacís Proaño por ser un apoyo incondicional desde el día que decidí emprender un camino lleno de obstáculos, sacrificios, decisiones, tristezas, alegrías, errores, decaimientos, humildad; brindándome tiempo, recursos y por sobre todo amor y confianza en cada instante de mi vida. Gracias a ellos soy quien soy y he podido lograr todo lo que me he propuesto. Me siento afortunado de tenerlos.

## **DEDICATORIA**

Dedico con mucho amor a mi familia José Luis Quisimalín Villacís, René Fabiola Villacís Proaño, Javier Alejandro y Estefanía Fabiola Quisimalín Villacís; y con mucho cariño y afecto a la familia Quisimalín Miniguano y a mi tía Gladis Villacís.

“Primero la Obligación después la Diversión”

# ÍNDICE

CAPÍTULO 1 .....	1
GENERALIDADES .....	1
1.1 INTRODUCCIÓN A LAS REDES NEURONALES .....	1
1.2 OBJETIVO GENERAL .....	2
1.3 OBJETIVOS ESPECÍFICOS .....	2
1.4 JUSTIFICACIÓN DEL ESTUDIO TÉCNICO .....	3
1.5 METODOLOGÍA .....	5
1.6 DESCRIPCIÓN DE LA EMPRESA .....	6
1.6.1 RESEÑA HISTÓRICA .....	6
1.6.2 CENTRALES HIDROELÉCTRICAS .....	8
1.6.3 SISTEMA DE SUBTRANSMISIÓN .....	8
1.6.4 SISTEMA DE DISTRIBUCIÓN .....	9
CAPÍTULO 2 .....	15
MARCO TEÓRICO .....	15
2.1 DEMANDA DE ENERGÍA ELÉCTRICA .....	15
2.1.1 CURVAS DE CARGA DIARIA .....	16
2.1.2 CURVAS DE ENERGÍA ANUAL .....	17
2.1.3 TIPOS DE DEMANDA ELÉCTRICA .....	17
2.2 USO DE LA DEMANDA ELÉCTRICA PARA LA PLANIFICACIÓN .....	19
2.3 PRONÓSTICO DE DEMANDA ELÉCTRICA .....	20
2.3.1 FACTORES GEOGRÁFICOS .....	22
2.3.2 CRECIMIENTO DEMOGRÁFICO .....	22
2.3.3 CRECIMIENTO ECONÓMICO (PIB) .....	23
2.3.4 DATOS HISTÓRICOS .....	23
2.3.5 TARIFAS ELÉCTRICAS .....	24
2.3.6 INFRAESTRUCTURAS PÚBLICAS Y PRIVADAS .....	24
2.3.7 HÁBITOS DE CONSUMO .....	24

2.3.8 DENSIDAD DE CARGA.....	25
2.4 MÉTODOS DE PRONÓSTICO DE LA DEMANDA ELÉCTRICA.....	25
2.4.1 SEGÚN EL HORIZONTE DE TIEMPO.....	25
2.4.2 SEGÚN EL MÉTODO EMPLEADO.....	27
2.4.3 SEGÚN EL SISTEMA ELÉCTRICO OBJETIVO.....	31
2.5 PROCESO PARA REALIZAR UNA PREDICCIÓN.....	31
2.6 SERIE TEMPORAL.....	33
2.6.1 COMPONENTES DE UNA SERIE DE TIEMPO.....	33
2.8 REDES NEURONALES ARTIFICIALES.....	34
2.8.1 MODELO DE NEURONA ARTIFICIAL.....	36
2.8.2 CAPACIDAD NO LINEAL DE LAS RNA.....	38
2.8.3 ARQUITECTURA DE UNA RNA.....	39
2.8.4 EL PERCEPTRÓN.....	39
2.8.5 EL PERCEPTRÓN MULTICAPA.....	40
2.8.6 FUNCIONES DE TRANSFERENCIA.....	42
2.8.7 ENTRENAMIENTO DE LA RED.....	43
2.9 VARIABLES.....	46
2.9.1 VARIABLES CUALITATIVAS.....	46
2.9.2 VARIABLES CUANTITATIVAS.....	47
2.10 CRITERIOS DE VALIDACIÓN Y ANÁLISIS.....	47
2.10.1 ERROR CUADRÁTICO MEDIO.....	47
2.10.2 ERROR RESIDUAL.....	47
2.10.3 DESVIACIÓN MEDIA ABSOLUTA.....	48
2.10.4 ERROR ESTÁNDAR MÚLTIPLE.....	48
2.10.5 PORCENTAJE DE ERROR MEDIO ABSOLUTO.....	49
2.11 VALORES ATÍPICOS Y CRITERIOS PARA ENCONTRARLOS.....	49
CAPÍTULO III.....	51
ANÁLISIS DE DATOS HISTÓRICOS Y METODOLOGÍA.....	51
3.1 PREPARACIÓN DE LA INFORMACIÓN.....	51
3.1.1 MEDICIÓN DE DATOS.....	51
3.1.2 ALMACENAMIENTO Y PROCESAMIENTO DE DATOS.....	53
3.1.3 SELECCIÓN DE ENTRADAS DEL MODELO.....	55

3.1.3 UNIFICACIÓN DE DATOS .....	59
3.2 NORMALIZACIÓN DE DATOS .....	65
3.3 RESOLUCIÓN DEL PROBLEMA .....	67
3.4 HERRAMIENTAS PARA LA PREDICCIÓN DE LA DEMANDA .....	68
3.4.1 SERIES TEMPORALES .....	68
3.4 CASOS DE ESTUDIOS .....	75
3.4.1 CASO 1.....	78
3.4.2 CASO 2.....	79
3.4.3 CASO 3.....	79
3.4.3 CASO 4.....	81
3.4.5 CASO 5.....	82
3.4.6 CASO 6.....	82
3.4.7 CASO 7.....	82
3.4.8 CASO 8.....	83
3.4.9 CASO 9.....	83
CAPÍTULO IV .....	84
RESULTADOS .....	84
4.1 CASO 1 .....	84
4.2 CASO 2.....	85
4.3 CASO 3.....	87
4.4 CASO 4 .....	88
4.5 CASO 5.....	90
4.6 CASO 6.....	91
4.7 CASO 7 .....	93
4.8 CASO 8.....	94
4.9 CASO 9.....	95
CAPÍTULO V .....	97
CONCLUSIONES Y RECOMENDACIONES .....	97

## RESUMEN

En el presente estudio técnico se aplica el análisis mediante Red Neuronal Artificial (RNA) para modelar y realizar la proyección del consumo de energía eléctrica de mediano plazo de la barra de 69 kV del Sistema Eléctrico de Subtransmisión a la Empresa Eléctrica Riobamba S.A. (EERSA) dicho mediano plazo está comprendido entre enero del 2017 y diciembre del 2017.

Se elaboraron diferentes casos para diversas condiciones con el propósito de encontrar la mejor red neuronal artificial posible, para ello, se entrenaron las RNA con los datos históricos, económicos y/o demográficos. Se emplearon redes neuronales tipo perceptrón multicapa con alimentación hacia adelante (*feedforward*) aplicando el algoritmo de retro propagación (*backpropagation*) para el entrenamiento, en el cual se dividieron los conjuntos de datos (*data set*) en un set para el entrenamiento, el cual a su vez se divide en un *subset* de entrenamiento parcial y un *subset* de validación del entrenamiento, así como un set de prueba para comparar las observaciones reales correspondientes al año 2016 y las proyecciones empleando la RNA. Los casos anteriores se analizaron siguiendo diferentes topologías de RNA donde 4 casos cuentan con salidas individuales (consumo del mes próximo) mientras que 5 casos cuentan con 12 salidas simultáneas (consumo de los próximos 12 meses). Se encontró que los modelos con salidas simultaneas mejoran en cierta medida la predicción de los datos que el método actual empleado por la Empresa Eléctrica Riobamba S.A. Los casos que involucran salidas individuales no tuvieron éxito en lograr el modelado de los datos.

## **PRESENTACIÓN**

Actualmente la Empresa Eléctrica Riobamba S.A., es una entidad de renombrado prestigio, por su servicio de calidad y el cumplimiento en su objetivo de generar y comercializar energía.

Sin embargo, la Empresa Eléctrica Riobamba S.A., consciente de los requerimientos de sus usuarios por brindarle una mejor calidad de servicio eléctrico y debido a que la demanda eléctrica crece en forma sostenida, se plantea la necesidad de pronosticar la demanda eléctrica de la empresa mediante una red neuronal artificial como metodología para lograr un gran equilibrio entre la energía generada y la energía consumida, una buena predicción de demanda eléctrica ayudará a que todos los medios productores de energía eléctrica generen exactamente la energía necesaria para confort y satisfacción de los usuarios tanto a nivel residencial, como comercial e industrial; garantizando el cumplimiento de objetivos técnicos y comerciales. Con la aplicación de una red neuronal artificial es posible aplicar un método más flexible que los tradicionales para la proyección de la demanda eléctrica.

Este Proyecto tiene como objetivo predecir la demanda eléctrica en barra de 69 kV del sistema eléctrico de Subtransmisión de la Empresa Eléctrica Riobamba S.A. mediante una red neuronal artificial.

A continuación, se detalla la estructuración y fundamentos que permitirán comprender el desarrollo del presente proyecto.

En el capítulo 1 se plantea la necesidad de pronosticar la demanda eléctrica de la Empresa Eléctrica Riobamba S.A. como metodología para garantizar la planificación adecuada con el objetivo de satisfacer la creciente demanda.

En el capítulo 2 se revisan los conceptos básicos relacionados a la demanda y los tipos de demandas. Se estudian de forma general los métodos de predicción de la demanda, y de forma específica el método por redes neuronales artificiales, su estructura y características de mayor relevancia. Se describen los parámetros estadísticos que permiten verificar modelos de predicción.

En el capítulo 3 se presenta una descripción del procedimiento realizado para la obtención de los datos y su posterior procesamiento mediante la descripción de los pasos involucrados en el proceso, cuyo objetivo es tener los datos con lo menos sesgados posibles. Posteriormente se describe el procedimiento para obtener el modelo de la red neuronal, el cual consiste en la selección del tipo de red, cantidad de capas, tipo de función de transferencia, tipos y cantidades de variables, método y tipo de entrenamiento y casos a evaluar.

En el capítulo 4 se presenta la validación y comparación de los resultados con la proyección de la Empresa Eléctrica Riobamba S.A. mediante la obtención de los parámetros estadísticos descritos en el capítulo 2 y 3.

En el capítulo 5 se exponen las conclusiones más relevantes de los resultados obtenidos del estudio sumado a las recomendaciones previstas para el futuro.

# CAPÍTULO 1

# CAPÍTULO 1

## GENERALIDADES

### 1.1 INTRODUCCIÓN A LAS REDES NEURONALES

Tradicionalmente los problemas se han formulado y solucionado mediante la aplicación de algoritmos. Sin embargo, existen problemas sujetos a muchos factores que no pueden ser representados exactamente por un algoritmo. El ser humano tiene una herramienta capaz de aprender, el “cerebro” que forma el centro de una red neuronal biológica que puede reorganizarse a sí misma durante la vida del individuo y por ende es capaz de aprender al compensar los errores. El estudio de la red neuronal artificial parte del análisis de esa red neuronal biológica capaz de aprender y esa diferencia con la complejidad que puede presentar elaborar un algoritmo, puesto que no hay necesidad de realizar programación, su aprendizaje proviene de un entrenamiento con muestras [1].

La investigación de las redes neuronales artificiales se remonta a 1943, año en el que Warren McCulloch y Walter Pitts presentaron el primer modelo de redes artificiales. Desde ese hito se han realizado propuestas nuevas y más sofisticadas. Las redes neuronales presentan una estructura multicapa que permite que la información se transmita no solo entre las neuronas vecinas sino a unidades más distantes. Una red neuronal artificial no necesita una programación, los parámetros de la red son encontrados adaptativamente. Aunque las redes neuronales artificiales son potencialmente más eficientes que la programación convencional en ciertas áreas de aplicaciones, al momento de su concepción no estaba lista para diversas aplicaciones. La teoría necesaria aún se encuentra en desarrollo y se prevé continúe por muchos años [2].

Las aplicaciones de las redes neuronales artificiales son muy diversas. Google las emplea en el reconocimiento e identificación de imágenes. Microsoft ha desarrollado redes neuronales para ayudar en la traducción de palabras en inglés a chino. Las aplicaciones se expanden no solo a la ingeniería y matemática sino a la medicina, finanzas, negocios y literatura [3].

Particularmente el desempeño de las redes neuronales artificiales ha sido adecuado para tareas de predicción, debido a las características únicas de las mismas, adaptabilidad, no linealidad, habilidad de mapear funciones arbitrariamente. Las redes neuronales artificiales se presentan como alternativas a los métodos tradicionales lineales [4].

En el presente capítulo se describen los fundamentos metodológicos que rigen el uso de las redes neuronales artificiales para la predicción de la demanda eléctrica del sistema eléctrico de subtransmisión de la Empresa Eléctrica Riobamba S.A. considerando los planteamientos anteriores.

## **1.2 OBJETIVO GENERAL**

Predecir la demanda eléctrica en barra de 69 kV del sistema eléctrico de Subtransmisión de la Empresa Eléctrica Riobamba S.A. mediante red neuronal artificial.

## **1.3 OBJETIVOS ESPECÍFICOS**

- Investigar y conocer los métodos de predicción de demanda eléctrica a mediano plazo en sistemas eléctricos.
- Evaluar el comportamiento y las variables que se involucran en la predicción de la demanda eléctrica.
- Recopilar y procesar los datos históricos de los medidores instalados mediante la selección, filtrado, depuración y normalización de las variables a considerar.

- Seleccionar y desarrollar la metodología del tipo de red neuronal artificial a utilizar para la predicción de demanda eléctrica a mediano plazo.
- Desarrollar el modelo de Red Neuronal Artificial propuesto usando una herramienta computacional.
- Realizar la comparación de los resultados obtenidos del método de inteligencia artificial de redes neuronales propuesto con el método estadístico adoptado por la Dirección de Planificación de la Empresa Eléctrica Riobamba S.A. y con los correspondientes datos registrados en los medidores instalados.

#### **1.4 JUSTIFICACIÓN**

La predicción de demanda eléctrica es una de las funciones más importantes en la operación de sistemas eléctricos, ya que es primordial en la planeación de la cantidad de potencia eléctrica que se debe generar para cubrir eficientemente un país o ciudad, sin ella se desperdiciarían los recursos utilizados por las centrales generadoras de energía eléctrica. En el ámbito de las empresas distribuidoras de electricidad, se compran cantidades mensuales de potencia eléctrica, la cual debe cubrir la demanda del total de clientes que las mismas posean; para el caso de no ser suficiente la demanda, se ven obligadas a comprar la demanda faltante al precio que el Mercado Eléctrico Mayorista (MEM) determine, o dejar de atender la demanda faltante. Si la potencia eléctrica adquirida es mayor a la demanda real, el sobrante se desperdicia o se negocia en condiciones propuestas por el MEM, casi siempre siendo desfavorable para las empresas distribuidoras. El mal cálculo de la demanda requerida trae sanciones de diferentes tipos y principalmente ocasiona pérdidas económicas para la empresa distribuidoras.

La predicción de la demanda eléctrica a mediano plazo corresponde al pronóstico mensual con un horizonte de hasta un año, se establece a partir de la demanda histórica y la predicción de algunas variables como el crecimiento de la economía, variaciones en el clima, períodos vacacionales, adición de nuevas cargas y los patrones de demanda de clientes no regulados, además de proporcionar elementos

para llevar a cabo programas de mantenimiento de redes y equipos y coordinación adecuada de recursos hidráulicos de las centrales generadoras, su principal objetivo es el de minimizar los desvíos de los valores previstos con relación a los verificados, optimizando los recursos disponibles, manteniendo la estabilidad del sistema eléctrico, suministrando medios para alcanzar una planificación económica, así como para la toma de futuras decisiones con el fin de acceder a subsidios para el planeamiento de expansión del sistema.

La Empresa Eléctrica Riobamba S.A., es una empresa que tiene como actividad económica principal: la comercialización, distribución y generación de energía eléctrica, en el ámbito de comercialización, compra y vende energía al Mercado Eléctrico Mayorista MEM, en donde las transacciones se realizan según: precios pactados en Contratos o, según precios libres de Mercado Ocasional, de acuerdo al desarrollo de la Provincia de Chimborazo, la demanda eléctrica tiende a aumentar a través del tiempo, por lo cual la empresa elabora planes para suplir la magnitud de esta demanda. El método adoptado por la Dirección de Planificación para realizar estimaciones de demanda eléctrica a mediano plazo es un método estadístico, el cual logra realizar la predicción a partir de series históricas mensuales de al menos cinco años. Los datos históricos de las salidas RIOBAMBA 1, RIOBAMBA 2 Y RIOBAMBA 3 de la barra de 69 kV de Subtransmisión de la subestación Riobamba que se utilizan para realizar las estimaciones se recopilan de los medidores instalados, sin embargo los mismos presentan distorsiones debido principalmente a eventos como traspasos de carga temporales, entre los equipos de medición y paros programados. Estas distorsiones deben ser depuradas para garantizar la calidad y precisión del proceso de estimación de la demanda eléctrica a mediano plazo [5].

Es necesario implementar acciones para solucionar o bien, para atenuar el impacto de este inconveniente sobre la predicción. Por lo tanto se propone evaluar la capacidad de las redes neuronales artificiales para tolerar tales distorsiones presentes en los datos históricos, así como su habilidad para realizar estimaciones de la demanda eléctrica a mediano plazo con precisión.

## **1.5 METODOLOGÍA**

### **1. Fase teórica**

- Se realiza una descripción general del proyecto de titulación, la importancia de la predicción de demanda eléctrica a mediano plazo en sistemas eléctricos, así como también se describe brevemente el Sistema Eléctrico de la Empresa Eléctrica Riobamba S.A.

- Se analiza el método de redes neuronales artificiales para la predicción de demanda eléctrica a mediano plazo evaluando las variables que la involucran.

### **2. Fase planteamiento metodológico**

- Sobre la base de la fase teórica se recopilaron y procesaron los datos históricos de los medidores instalados mediante la selección, filtrado, depuración y normalización de las variables a considerar.

- Se selecciona y desarrolla la metodología del tipo de red neuronal artificial a utilizar para la predicción de demanda eléctrica a mediano plazo.

### **3. Fase Simulación**

- Se desarrolla el modelo de Red Neuronal Artificial propuesta usando una herramienta computacional.

### **4. Fase Validación / Análisis de resultados**

- Se realiza el entrenamiento y validación del modelo de red neuronal artificial propuesto.

- Se realiza la comparación de los resultados obtenidos del método de inteligencia artificial de redes neuronales propuesto con el método estadístico y con los correspondientes datos registrados en los medidores instalados.

- Se analiza los resultados para posteriormente determinar las conclusiones del proyecto.

## 1.6 DESCRIPCIÓN DE LA EMPRESA

### 1.6.1 RESEÑA HISTÓRICA

La historia inicia en 1903 se funda la Sociedad “Alberto Rhor y Cía” la cual se encargó de comercializar energía. Sin embargo, sus operaciones no duraron mucho debido a que quebró en 1907. Posteriormente, en 1911 fue fundada la Sociedad Anónima Riobamba Electric Ligth and Power. Para el año 1953 tiene lugar la conformación de la Empresa de Electrificación Chimborazo S.A., que presenta una variedad de proyectos como la construcción y puesta en servicio de la Central Hidroeléctrica Alao.

El origen de la Empresa Eléctrica Riobamba S.A. se remonta al 3 de abril de 1963, cuando compra los derechos a la Empresa de Electrificación Chimborazo S.A. En el futuro inmediato a su origen se llevaron a cabo inversiones para la construcción de la central Alao, realizando la inauguración de sus dos primeras etapas el 2 de enero de 1967. Posteriormente la empresa se fusiona con la Empresa Eléctrica Alausí para el año 1976, empresa que operaba la central hidroeléctrica Nizag. En el año 1977 se inaugura la tercera etapa de la Central Alao y la cuarta y última etapa tiene su inauguración en 1979; durante el año 1979 pasó a formar parte del Sistema Nacional Interconectado para luego iniciar la construcción de la línea San Juan – Alausí y las subestaciones San Juan, Guamote y Alausí, cada una con 1 MVA [5].



Figura 1.1. Tubería de presión de la Central Hidroeléctrica Alao  
Fuente: Basado en [5]



Figura 1.2. Turbinas de la Central Hidroeléctrica Alao  
Fuente: Basado en [5]

En el año 1997 la Empresa Eléctrica Riobamba culmina e inaugura la central hidroeléctrica Rio Blanco con el propósito de mejorar el servicio mientras se garantiza el suministro a la creciente demanda.



Figura 1.3. Casa de máquinas de la Central Hidroeléctrica Rio Blanco  
Fuente: Basado en [5].

Cuatro años después ingresa a formar parte del sistema eléctrico de la EERSA la subestación TAPI que asume parcialmente la carga del cantón Guano y además el norte de la ciudad de Riobamba. Para el año 2010 se incrementa la capacidad de transformación de la subestación No. 2 al cambiar el transformador de potencia de 10 MVA a 15 MVA mejorando la posibilidad de satisfacer el incremento de la demanda del usuario final. En el mismo año inicia la construcción de la línea de subtransmisión Alao – Guamote que entró en operación a mediados del 2011. Simultáneamente se realizó la contratación para el equipamiento de las salidas en las Subestaciones de Alao y Guamote para interconectar la línea construida [5].

### **1.6.2 CENTRALES HIDROELÉCTRICAS**

Dentro de la provincia Chimborazo, la Empresa Eléctrica Riobamba S.A. opera 3 centrales Hidroeléctricas que proveen de energía eléctrica a la provincia.

#### **Central Alao**

Es una de las 3 centrales con mayor capacidad de generación de energía eléctrica a la provincia, ubicada en la Parroquia Pungalá del Cantón Riobamba la generación tiene lugar gracias al caudal de los ríos Alao y Maguazo, lo que permite generar 10,4 MW. La central Alao cuenta con cuatro grupos de generación.

#### **Central Rio Blanco**

Esta central está ubicada en la parroquia Quimiag del cantón Riobamba, aprovechando las aguas del rio del mismo nombre es capaz de generar 3 Megavatios. La central Rio Blanco cuenta con un grupo de generación.

#### **Central Nizag**

La central Nizag está ubicada en la comunidad de Nizag del cantón Alausí, aprovecha las aguas del Río Zula para generar 0,8 Megavatios. La central Nizag cuenta con un grupo de generación.

### **1.6.3 SISTEMA DE SUBTRANSMISIÓN**

El sistema está conformado por diversas líneas que interconectan 11 subestaciones con una capacidad instalada total de 72 MVA. En total, existen 175,1 km de líneas que operan a un voltaje de 69 kV

En el presente estudio, toma vital importancia la subestación Riobamba perteneciente a TRANSELECTRIC, puesto que esta subestación sirve de enlace con el Sistema Nacional Interconectado, en esta subestación existe la transferencia de energía con el MEM

En la Tabla 1.1 se presenta las diferentes líneas de subtransmisión que conforman el sistema eléctrico de la Empresa Eléctrica Riobamba S.A.

Tabla 1.1. Líneas de subtransmisión del sistema eléctrico de la Empresa Eléctrica Riobamba S.A.

Fuente: EERSA

<b>Líneas de Subtransmisión</b>	<b>Longitud(km)</b>	<b>Cantón</b>
S/E Alao - S/E 1	17,0	Riobamba
S/E 1 - S/E Riobamba	6,4	Riobamba
S/E 1 - S/E 3	4,1	Riobamba
S/E 3 - S/E 2	3,6	Riobamba
S/E 2 - S/E 4	6,7	Riobamba
S/E 4 - S/E Riobamba	2,3	Riobamba
S/E Riobamba - S/E San Juan Chico	7,04	Riobamba
S/E San Juan Chico - S/E Cajabamba	6,0	Riobamba - Colta
S/E Cajabamba - S/E Guamote	28,03	Colta - Guamote
S/E Guamote - S/E Alausí	36,0	Guamote - Alausí
S/E Alausí - S/E Chunchi	14,54	Alausí - Chunchi
S/E Alao - S/E Guamote	19,4	Riobamba - Guamote
S/E Multitud - S/E Alausí	23,9	Alausí

#### 1.6.4 SISTEMA DE DISTRIBUCIÓN

El Sistema de Distribución de energía eléctrica para servir a la provincia de Chimborazo se encuentra constituido por 38 alimentadores primarios de medio voltaje (MV), con 119.382 postes de los cuales el 43 % son de líneas primarias (MV) y el 57% de redes secundarias (BV), con 12.328 centros de transformación de los cuales el 94% son transformadores monofásicos y el 6% transformadores trifásicos, con 3.728 puntos de seccionamiento de los cuales el 62 % son monofásicos, el 24% son bifásicos y el 35 % son trifásicos, con 47.630 luminarias de las cuales el 98,4% son a vapor de sodio y el 1,6% a vapor de mercurio, según información proporcionada por la Dirección de Planificación.

En la Tabla 1.2 se detallan las características de distribución del sistema eléctrico de la Empresa Eléctrica Riobamba S.A.

Tabla 1.2. Características de distribución del sistema eléctrico de la Empresa Eléctrica Riobamba S.A.

Fuente: EERSA

Red de Distribución de MV	13,8 y 4,16 kV
Red de Distribución de BV	127/220 trifásico y 120/240V monofásico
11 Subestaciones AV/MV	69/13,8 kV
1 Subestación MV/MV	13,8/4,16 kV
Centros de Transformación Trifásicos MV/BV	13800/220 V
Centros de Transformación Bifásicos MV/BV	7960/240-120 V

La provincia también cuenta con 11 Subestaciones ubicadas estratégicamente en los distintos cantones, así cada una de ellas está destinada a la sub transmisión, transformación y comercialización de tan importante servicio vital.

Sub Estación #1. (Riobamba –Junto al cementerio).

Sub Estación #2. (Riobamba –Junto al Paseo Shopping).

Sub Estación #3. (Riobamba –Parque Industrial).

Sub Estación #4. (Riobamba –Cdla. 24 de Mayo).

Sub Estación #6. (San Juan Chico -Gatazo).

Sub Estación #7. (Colta –Cajabamba).

Sub Estación #8. (Guamote).

Sub Estación #9. (Alausí).

Sub Estación #10. (Chunchi).

Sub Estación #13. Parroquia Pungalá –Central Alao).

Sub Estación #14. (Pallatanga).

A continuación, se detallan algunas de ellas:

#### **1.6.4.1 Subestación Nro. 01: Chibunga**

Chibunga es la subestación Nro. 01, se encuentra localizada al lado occidental de la ciudad de Riobamba, en la avenida Circunvalación. Esta subestación es una de las más importantes de la ciudad, ya que en ella se realizan varias funcionalidades muy significativas para el sistema eléctrico de la Empresa Eléctrica Riobamba S.A.

Dicha subestación posee una capacidad de transformación de 15 MVA, de un voltaje de 69 kV a 13,8 kV y un transformador adicional de potencia, de 3,125 MVA de un voltaje de 13,8 kV a 4,16 kV.

#### **Alimentadores**

- Alimentador a 4,16kV

Alimentador 0401

- Alimentador a 13,8kV

Alimentador 0101

Alimentador 0201

Alimentador 0301

Alimentador 0501

Alimentador 0601

En las tablas 1.3 y 1.4 se exponen las características de los transformadores de potencia de la subestación Nro. 1: Chibunga.

Tabla 1.3 Datos técnicos del Transformador de 15 MVA de la subestación Chibunga.

Fuente: EERSA

<b>TRANSFORMADOR DE 15 MVA</b>	
<b>Voltaje (V)</b>	<b>Posición</b>
72450	1
70725	2
69000	3
67275	4
65550	5
<b>Tap Actual: 3</b>	
<b>Impedancia: 7,49% Conexión: DY5</b>	

Tabla 1.4 Datos técnicos del Transformador de 3.125 MVA de la subestación Chimbunga.

Fuente: EERSA

<b>TRANSFORMADOR DE 3.125 MVA</b>	
<b>Voltaje (V)</b>	<b>Posición</b>
14400	1
14100	2
13800	3
13500	4
13200	5
<b>Tap Actual: 3</b>	
<b>Impedancia: 5,21% Conexión: DY5</b>	



Figura 1.4. Subestación N°1  
Fuente: Basado en [5].

#### 1.6.4.2 Subestación Nro. 02: Maldonado

La subestación Maldonado se encuentra ubicada al noroeste de la ciudad de Riobamba en el Kilómetro 1 vía a Guano. Además de poseer un alimentador 0402 que puede estar cubierto en su demanda parcialmente con generación de la central hidroeléctrica Río Blanco (alimentador 0403), tiene interconexión con la subestación Nro. 3 llamada Parque Industrial. Su capacidad de transformación es de 15 MVA con una relación de voltaje de 69 kV a 13,8 kV.

#### Alimentadores

Alimentador 0102

Alimentador 0202

Alimentador 0302

Alimentador 0402 - Guano

Alimentador 0602 – Brigada

Alimentador 0702 – Cerámica

En la Tabla 1.5 siguiente se exponen las características de los transformadores de potencia de la subestación Nro, 2: Maldonado.

Tabla 1.5 Datos técnicos del Transformador de 15 MVA de la subestación Maldonado.

Fuente: EERSA

<b>TRANSFORMADOR DE 15 MVA</b>	
<b>Voltaje (V)</b>	<b>Posición</b>
72450	1
70725	2
69000	3
67275	4
65550	5
<b>Tap Actual: 3</b>	
<b>Impedancia: 7,26% Conexión: DY5</b>	

#### 1.6.4.3 Subestación Nro. 03: Parque Industrial

La subestación Nro. 3 conocida como Parque Industrial se encuentra ubicada al sureste de la ciudad de Riobamba, en la parte periférica de la ciudad, en la Av. Circunvalación. Esta subestación, tiene una capacidad de transformación de 10 MVA, de un voltaje de 69 KV a 13,8 kV

### Alimentadores

Alimentador 0103 – San Luis

Alimentador 0203 – San Gerardo

Alimentador 0303 - Tubasec

Alimentador 0403 - Penipe

Alimentador 0503 – Parque Industrial

En la Tabla 1.6 se exponen las características de los transformadores de potencia de la subestación Nro, 2: Maldonado.

Tabla 1.6 Datos técnicos del Transformador de 10 MVA de la subestación Maldonado.

Fuente: EERSA

<b>TRANSFORMADOR DE 10 MVA</b>	
<b>Voltaje (V)</b>	<b>Posición</b>
72450	1
70725	2
69000	3
67275	4
65550	5
<b>Tap Actual: 3</b>	
<b>Impedancia: 7,1% Conexión: DY5</b>	

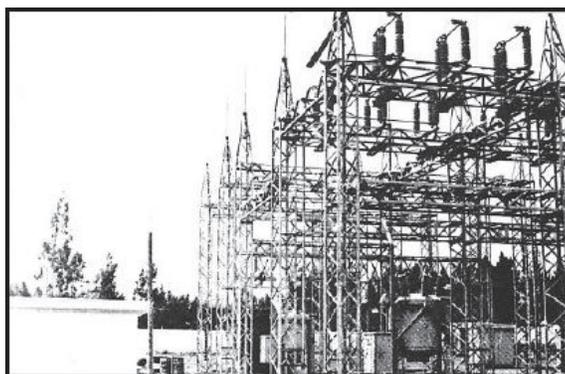


Figura 1.5. Subestación N°3  
Fuente: Basado en [5].

# CAPÍTULO 2

## CAPÍTULO 2

### MARCO TEÓRICO

El marco teórico comprende la exposición y análisis de las teorías, investigaciones y/o antecedentes que sustentan la investigación. Uno de los propósitos es que proporcione ideas nuevas a partir de los descubrimientos recientes de diversos investigadores. El marco teórico permitirá la orientación sobre cómo se han realizado las investigaciones respecto a la proyección de la demanda eléctrica mediante redes neuronales artificiales, particularmente como se han recolectado los datos y como procesarlos para lograr los objetivos teniendo marcos de referencia para interpretar los resultados y compararlos con los obtenidos por otros investigadores [6]. En este capítulo se plantean fundamentos teóricos básicos y necesarios para la comprensión del problema.

#### 2.1 DEMANDA DE ENERGÍA ELÉCTRICA

Representa el promedio de la potencia eléctrica utilizada por un consumidor (carga) en periodo de tiempo determinado, usualmente de 15, 20 o 30 minutos [7]. Generalmente es un valor medio para un intervalo de tiempo, este periodo es denominado intervalo de demanda, su duración variaría en función de la demanda que se desee identificar [8]. Siempre es necesario establecer un intervalo de demanda para poder conocer todas las variables que involucra el consumo del usuario. La demanda se expresa en kVA, kW, etc. La relación entre la demanda y el tiempo se puede representar gráficamente mediante una curva de carga (demanda vs. tiempo) [9]. En la Figura 2.1 se muestra una fracción de una curva de demanda típica, usualmente se emplea un tiempo de 15 minutos para el cálculo de la demanda promedio [10].

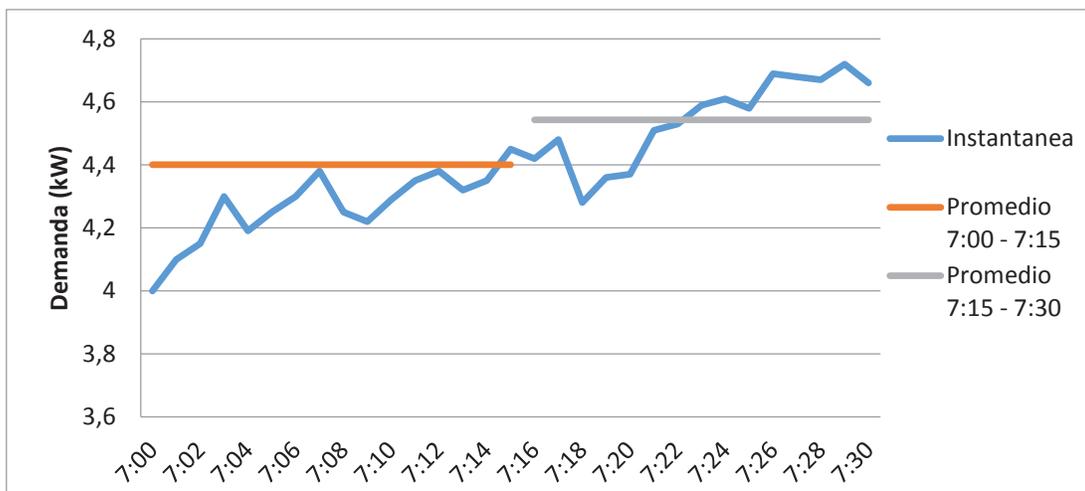


Figura 2.2. Demanda eléctrica instantánea y promedio 15 min.  
Fuente: Basado en [10].

### 2.1.1 CURVAS DE CARGA DIARIA

Es una curva característica de la variación diaria de la demanda como una función de los intervalos de demanda, esta curva es también llamada curva de carga [8]. En la Figura 2.2 se muestra un ejemplo de curva de carga diaria considerando intervalos de demanda de 15 minutos.

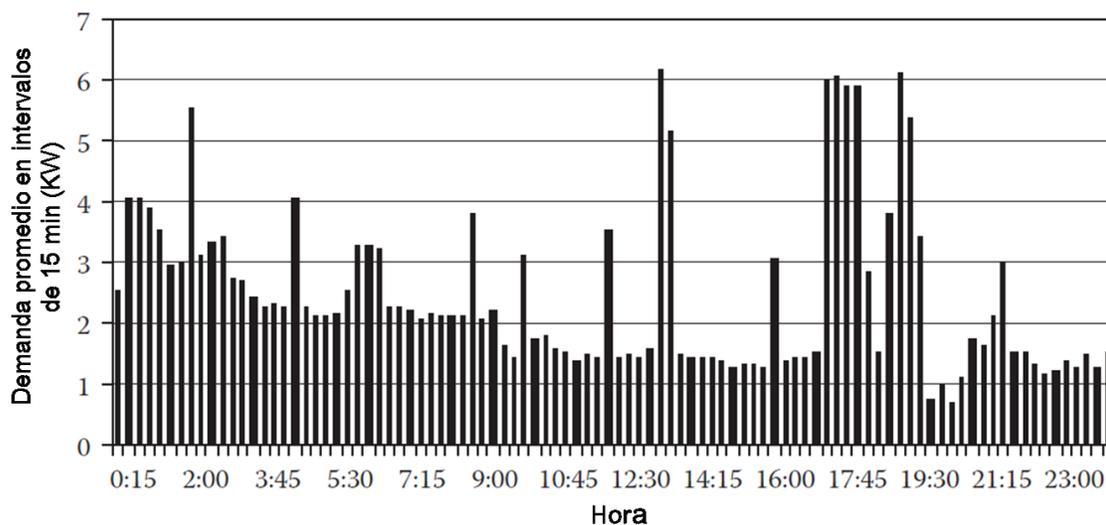


Figura 2.2. Demanda eléctrica diaria con intervalos de demanda de 15 min.  
Fuente: Basado en [10].

### 2.1.2 CURVAS DE ENERGÍA ANUAL

El área debajo de la curva de la demanda representa la energía que se consume o genera en el intervalo de tiempo indicado. Para obtener la generación o consumo de energía anual es posible construir graficas a partir de información diaria considerando que un año típico posee 365 días [11]. Su forma y comportamiento depende de las variables que afectan la demanda a mediano plazo [10]. En la Figura 2.3 se muestra una curva de energía anual para el año 2014, correspondiente a la energía eléctrica consumida en la barra de 69 kV de la EERSA.

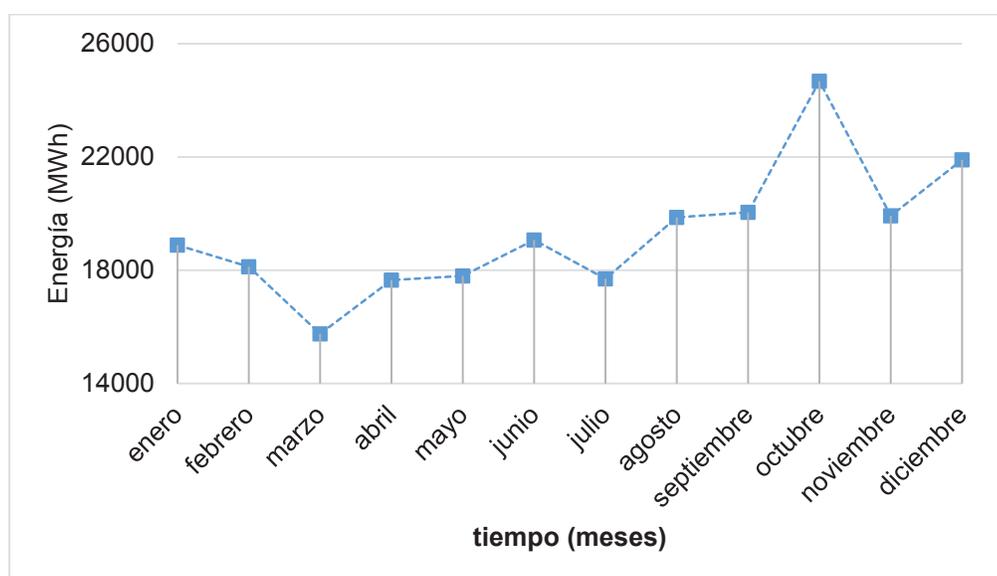


Figura 2.3. Consumo de energía eléctrica mensual (año 2014).  
Fuente: Elaboración propia con datos de la EERSA

### 2.1.3 TIPOS DE DEMANDA ELÉCTRICA

Tradicionalmente se han establecido tres (3) tipos de consumidores: residencial, comercial e industrial.

#### 2.2.3.1 Consumidores residenciales

La carga residencial es la que tiene un comportamiento con una profunda estacionalidad, depende de los hábitos de consumo de la población. El crecimiento o decrecimiento de la demanda residencial está muy ligada a la capacidad económica de la población y al crecimiento demográfico de la misma [12].

La mejora en la eficiencia de los equipos eléctricos y electrónicos ha resultado en la disminución de la demanda individual residencial. Sin embargo, el gran aumento poblacional, exponencial en la mayoría de los casos ha restringido dicha disminución y por ende este tipo de carga aumenta cada día más [12].

Según el pliego tarifario para las empresas eléctricas un consumidor residencial se define como el consumidor constituido por la unidad familiar, independientemente del tamaño de la carga conectada. En esta categoría también se incluyen los consumidores con escasos recursos económicos y bajo consumo que posean en su casa una pequeña actividad comercial o artesanal [13].

#### **2.1.3.2 Consumidores comerciales**

Un consumidor comercial se define como aquella persona natural o jurídica, pública o privada, que emplea los servicios de energía eléctrica para sus negocios, actividades profesionales o demás actividades que tengan fines de lucro [13].

#### **2.1.3.3 Consumidores industriales**

Un consumidor industrial se define como aquella persona natural o jurídica, pública o privada, que emplea los servicios de energía eléctrica para la fabricación o transformación de productos por medio de cualquier proceso industrial. También se considera a los agroindustriales, debido a que transforman productos de la agricultura, ganadería, riqueza forestal y pesca, en productos elaborados [13].

#### **2.1.3.4 Demanda máxima**

La más grande de todas las demandas que ocurren durante un tiempo específico. Debe incluir el intervalo de demanda, período y unidades. Ejemplo: La demanda máxima en kW de 15 min para la semana fue de 150 kW [14].

#### **2.1.3.5 Carga instalada**

Corresponde a la suma de las potencias nominales continuas conectadas a un sistema, usualmente se expresa en kVA, MVA, kW o MW [9].

### 2.1.3.6 Capacidad instalada

Potencia nominal de una unidad generadora o un conjunto de unidades referenciadas como una central de generación, un sistema local o un sistema interconectado [15].

### 2.1.3.7 Factor de potencia

Relación de la demanda media de un cliente individual o de un grupo de clientes durante un período a la demanda máxima durante el mismo período [10]. El factor de potencia relaciona la potencia activa con la potencia aparente mediante la ecuación 2.1.

$$fp = \frac{P}{S} \quad (2.1)$$

Donde:

$P$ : Potencia activa o real

$S$ : Potencia aparente

El factor de potencia depende de la carga conectada y generalmente está ligado al tipo de consumidor. Las cargas residenciales tienen el factor de potencia más cercano a la unidad (0,90 – 0,95), las cargas comerciales tienen un factor de potencia menor (0,80-0,90), mientras que las cargas industriales son las que usualmente tienen el menor factor de potencia (0,75 – 0,80) [16].

## 2.2 USO DE LA DEMANDA ELÉCTRICA PARA LA PLANIFICACIÓN

Las proyecciones de demanda futura son una acción primaria esencial en el proceso de planificación y decisión de las inversiones necesarias para garantizar el futuro abastecimiento de energía eléctrica, constituye un insumo para la elaboración de presupuesto, estudios de pérdidas e inversiones y la realización de cálculos tarifarios [8]. La demanda es esencial para diseñar los circuitos primarios en las redes de distribución [9]. La Figura 2.4 es tomada del plan maestro de electrificación 2013 – 2022 donde se indica el uso de la demanda en la planificación [17].

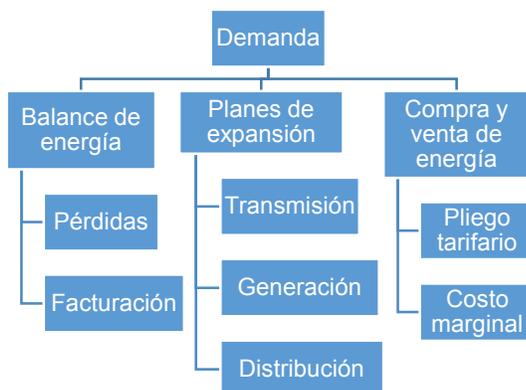


Figura 2.4. Uso de la demanda eléctrica para la planificación  
Fuente: Plan maestro de electrificación 2013-2022

### 2.3 PRONÓSTICO DE DEMANDA ELÉCTRICA

El problema esencial de la predicción de demanda eléctrica es determinar valores de la futura demanda eléctrica, dados valores pasados (datos históricos) de la demanda y algunas variables explicativas (causales). Las entradas deben ser apropiadas para la escala de tiempo de interés [18].

Una predicción precisa es necesaria para que las empresas eléctricas determinen la dinámica y características de la construcción futura necesaria en el sistema. Una proyección precisa necesita tanto datos estadísticos como la pericia de la persona encargada en realizarla [19]. La precisión es el criterio de mayor importancia al momento de seleccionar un modelo de pronóstico, la precisión refleja la proximidad de los valores pronosticados y los valores reales [20]. Las proyecciones pueden clasificarse en corto, mediano y largo plazo con respecto a los periodos de predicción [19].

Las proyecciones realizadas para la demanda y energía eléctrica reflejan sus posibles compromisos futuros en el abastecimiento y requerimientos de capacidad para la cobertura del servicio y mantiene un comportamiento tendencialmente creciente [21]. El pronóstico puede definirse para dos casos, el primer caso es cuando se desea conocer la potencia y el segundo caso cuando se desea conocer la energía. Usualmente se realizan proyecciones considerando la energía (MWh) aunque

teniendo en cuenta que ambas variables no son ajenas, sino vinculadas mediante el periodo de tiempo evaluado [22]. En la Figura 2.5 se detallan los factores que afectan la demanda eléctrica mencionados en [8]. Además de estos factores debe añadirse los hábitos de consumo como otro factor importante en la predicción de la demanda eléctrica [23]. Las variables o factores explicativos también pueden definirse según lo establecido en [18].

- Demanda horaria (corto plazo): factores climatológicos (temperatura, humedad, velocidad del viento, cobertura de nubes, etc.) fechas especiales (eventos deportivos, feriados, fiestas nacionales, etc.)
- Demanda semanal (mediano plazo): factores climatológicos (temperatura diaria mínima, temperatura máxima diaria, etc.) y actividades económicas.
- Demanda anual (mediano y largo plazo): factores económicos (PIB, índice de precios al consumidor, producción industrial, etc.) y factores demográficos (población, emigración, inmigración, etc.)



Figura 2.5. Factores que influyen en el pronóstico de la demanda.  
Fuente: Basado en [8].

### **2.3.1 FACTORES GEOGRÁFICOS**

El consumo eléctrico está íntimamente relacionado con los factores geográficos como lo son: temperatura, vientos, humedad, los cuales depende de la latitud y longitud de la zona estudiada, así como de las características propias de la geografía del lugar. Tradicionalmente la temperatura es la variable que mundialmente afecta más la demanda eléctrica, con énfasis en aquellas zonas donde los cambios o gradientes de temperatura son muy bruscos a lo largo del año [24]. El clima en la provincia Chimborazo es por lo general frío y consta de dos estaciones, una húmeda y una seca. Los vientos en Riobamba pueden producir una sensación térmica de casi 0 °C en algunas épocas del año la máxima temperatura diaria toma valores de 25 °C a 27 °C, es inusual encontrar registros con temperaturas mayores que 27 °C.

Existen además factores geográficos que no afectan directamente la demanda, sino la capacidad de las centrales de generación, particularmente en el Ecuador, los fenómenos del El Niño y La Niña tiene alto impacto en la capacidad de las centrales hidroeléctricas [22].

### **2.3.2 CRECIMIENTO DEMOGRÁFICO**

El aumento en la población de una región conlleva un incremento en el número de suscriptores y consumidores de energía eléctrica. El número de familias rurales ha estado disminuyendo continuamente en este siglo, desde la década de los cincuenta se ha mantenido la migración de la población rural a las áreas urbanas. Esta tendencia continuará y resultará en el incremento de hogares multifamiliares en áreas que hoy en día tienen elevadas densidades poblacionales [8].

La provincia de Chimborazo, se caracteriza por un crecimiento poblacional sostenido. Las nuevas iniciativas que se han tomado para continuar con el desarrollo, algunas en proceso y otras en proyecto impulsarán a mantener el ritmo del crecimiento. Por ende, se espera que la demanda eléctrica en Chimborazo crecerá a medida que aumente su población [25].

Además, un incremento en el número de hogares, cantidad de electrodomésticos y número de habitaciones en un hogar son factores responsables del incremento de la demanda residencial [23].

### **2.3.3 CRECIMIENTO ECONÓMICO (PIB)**

La mayoría de pronósticos de la demanda eléctrica han concluido que esta es una de las variables más influyentes para proyecciones a mediano plazo. El producto interno bruto (PIB) mide el valor de todos los productos y servicios que son producidos en un tiempo determinado (generalmente anual). Un aumento en la producción de bienes y servicios está relacionado positivamente con un crecimiento en la demanda eléctrica. Por el contrario un decrecimiento económico está ligado a una disminución en la demanda eléctrica [26].

La economía de la provincia de Chimborazo se basa en parte en la producción agrícola. Sin embargo, también existe la presencia de industrias de cerámica, cementeras, lácteos, madereros, molineras, elaborados de construcción, piezas automotrices, turismo, ensamblaje de computadores, fabricación de hornos, techos, tuberías, entre otras. Además del potencial turístico que posee, por lo tanto se puede decir que es una ciudad en constante crecimiento económico [27].

### **2.3.4 DATOS HISTÓRICOS**

Estadísticamente se ha comprobado que el comportamiento de la demanda eléctrica tiene componentes estacionarios y no estacionarios, cíclicos, lineales, exponenciales y demás. No obstante, el comportamiento generalizado se mantiene similar a lo largo del tiempo, por lo que el análisis de los datos históricos es de suma importancia y en la mayoría de los casos es la base para aplicar los métodos de predicción en la demanda [22]. El crecimiento de la demanda puede ser obtenido del análisis de los datos históricos considerando como mínimo los últimos 4 años [9].

### **2.3.5 TARIFAS ELÉCTRICAS**

El costo del kilovatio-hora (kWh) es un factor influyente en los hábitos de consumo y por ende en la demanda de un consumidor, el precio referencial es determinado por las instituciones competentes según la ley. No obstante, la demanda no guarda una correlación muy estrecha con esta variable, debido a la necesidad de la electricidad para el desarrollo de las actividades domésticas [21]. Estudios consideran que el precio de la electricidad no tiene efecto apreciable en el uso de la energía a mediano plazo, las políticas gubernamentales implicadas en el precio de la energía es un factor que tiene un efecto nulo o casi nulo en el uso de la electricidad. [23].

En el Ecuador, la Agencia de Regulación y Control de Electricidad (ARCONEL) se encarga de emitir los pliegos tarifarios, puesto que de manera similar a los consumidores, las tarifas se segmentan según el uso y sector de cobertura [21]. En los pliegos anuales emitidos por la ARCONEL se especifican las tarifas segmentadas generalmente según el nivel de voltaje de servicio [13].

### **2.3.6 INFRAESTRUCTURAS PÚBLICAS Y PRIVADAS**

Este factor está relacionado con la puesta en marcha de proyectos públicos y privados que generen cambios en la demanda. Dentro de proyectos públicos se puede encontrar la construcción de plataformas gubernamentales, sistemas de transporte (metro, tranvía, trolebús y otros), construcción de oficinas municipales, entre otros casos. Dentro de proyectos privados se puede encontrar la construcción de hoteles, centros comerciales, oficinas y cualquier infraestructura en general [22].

### **2.3.7 HÁBITOS DE CONSUMO**

Existen carencias a nivel global en la educación de un consumo eficiente de la energía eléctrica. Si bien este factor está ligado a la economía y tecnología de una nación, el hábito de consumo depende exclusivamente de cada ciudadano [22]. El comportamiento de los humanos es usualmente periódico en términos diarios, semanales, mensuales, anuales, etc. Los comportamientos de las cargas incluyen

generalmente periodicidad. Los hábitos en el comportamiento de los ocupantes de unas instalaciones juegan un papel fundamental en el perfil de la demanda [26].

Modelar los hábitos de consumo como una variable independiente que tiene efectos en la demanda eléctrica se presenta como un desafío para los investigadores, en vista de que la población involucra una gran dimensionalidad de variables (edad, sexo, educación, localización, ingresos, salud, tiempo libre) [23].

### 2.3.8 DENSIDAD DE CARGA

Este factor influye principalmente en la planeación de la distribución y las redes eléctricas, ya que indica la razón entre la energía demandada y el área geográfica involucrada, lo que conlleva a una unidad de kVA por metro cuadrado. En las zonas comerciales cada vez es mayor la densidad de carga debido a la construcción de nuevas edificaciones verticales [22].

## 2.4 MÉTODOS DE PRONÓSTICO DE LA DEMANDA ELÉCTRICA

### 2.4.1 SEGÚN EL HORIZONTE DE TIEMPO

Como se mencionó en el apartado 2.3 se pueden distinguir tres subclasificaciones: corto, mediano y largo plazo. En la Figura 2.6 se muestra un diagrama para esta clasificación.

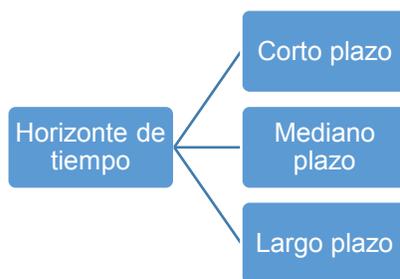


Figura 2.6. Pronóstico según el horizonte del tiempo.  
Fuente: Basado en [18], [22].

#### **2.4.1.1 Pronóstico a corto plazo**

Estudios consideran un pronóstico a corto plazo a aquel donde el periodo de predicción abarca desde una hora hasta una semana. Su principal objetivo es el estudio de la operación de los sistemas de potencia, reducción de pérdidas, regulación de tensión y otros [28]. Los datos climatológicos generalmente son los más importantes como variable independiente para el pronóstico a corto plazo, incluyen: temperatura, humedad, precipitación, velocidad del viento, cobertura de las nieves e intensidad solar [29]. Por lo tanto se caracteriza por tener una demanda máxima en los momentos de temperatura más alta y en las horas en las que llegan las personas a casa luego del trabajo [22].

#### **2.4.1.2 Pronóstico a mediano plazo**

A diferencia del pronóstico a corto plazo, el pronóstico a mediano plazo toma en consideración los datos históricos y datos climatológicos, datos económicos y geográficos y otros factores [30]. Usualmente este pronóstico se establece desde una semana hasta un año [24]. El pronóstico a mediano plazo permite pronosticar los requerimientos energéticos para garantizar el despacho energético a la demanda futura como el consumo de combustible requerido por las unidades generadoras y la disponibilidad hidráulica de las presas [31].

#### **2.4.1.3 Pronóstico a largo plazo**

Se considera un intervalo a partir de un año [24]. Este pronóstico se emplea para la planeación y expansión del sistema eléctrico de potencia en sus 3 etapas, generación, transmisión y distribución. No solo es importante para las empresas eléctricas sino también para las empresas constructoras y fabricantes de equipos eléctricos [22]. Los factores más influyentes son las variables socioeconómicas, debido a que consideran el comportamiento del producto interno bruto (PIB) y el crecimiento demográfico, los factores climatológicos dejan de influir para este pronóstico [31].

## 2.4.2 SEGÚN EL MÉTODO EMPLEADO.

Se pueden distinguir 3 grupos según el método empleado, estos corresponden a métodos estadísticos, de inteligencia artificial y micro áreas. Dentro de los métodos tradicionales se encuentran la regresión, las series de tiempo y la distribución de probabilidad.

### 2.4.2.1 Método econométrico

El modelo econométrico es un modelo económico que tiene especificaciones necesarias para establecer la relación existente entre una o varias variables. Considerando la función de demanda el modelo econométrico sería  $E_t = \alpha + \beta X_t + \varepsilon$ , donde  $E_t$  correspondería a la energía, variable dependiente,  $X_t$  corresponde a la variables o variables explicativas,  $\alpha$  es un término independiente que permite indicar el valor de la variable dependiente cuando las variables explicativas son cero,  $\beta$  son los coeficientes del modelo y  $\varepsilon$  es una variable aleatoria. La variable aleatoria dota al modelo de un mayor realismo permitiendo introducir la incertidumbre existente en cualquier comportamiento social [32].

La forma general de un modelo econométrico se expresa en la ecuación 2.2.

$$\ln Y_t = \ln \alpha + \beta_1 \ln X_{1t} + \beta_2 \ln X_{2t} + \dots + \beta_k \ln X_{kt} + \varepsilon \quad (2.2)$$

Donde:

$\ln$ : Modelo lineal a estimar

$Y_t$ : Variable a explicar

$\varepsilon$ : Término de error aleatorio

$X_t$ : Variables explicativas

$\alpha$ : ordenada al origen

$\beta$ : Elasticidades de  $Y_t$  con respecto a las  $X_t$ .

### **Etapas en la elaboración de un modelo econométrico**

**Especificación del modelo:** en esta etapa se expresa la relación propuesta por la Teoría Económica en un lenguaje matemático, para ello se determinan las variables a introducir y la función que relaciona dichas variables, de igual manera se determinan distintas hipótesis sobre todas las variables del modelo [32].

**Elección y tratamiento de los datos:** estas dos etapas van muy unidas debido a que en el proceso se especifica el modelo y se eligen los datos. La disposición de los datos permite especificar de mejor manera el modelo [32].

**Estimación:** en esta etapa se obtienen estimadores de los parámetros a partir de los datos disponibles [32].

**Evaluación y Contratación:** se comprueba si las hipótesis estadísticas y económicas son coherentes con los datos disponibles [32].

**Predicción:** en esta fase se obtienen valores futuros de la variable dependiente a partir de valores conocidos de las variables explicativas [32].

#### **2.4.2.2 Regresión lineal simple**

La estructura general del modelo de regresión lineal se muestra en la ecuación 2.3.

$$Y = \beta_0 + \beta_1 X + \varepsilon \quad (2.3)$$

En la expresión se muestra que las causas y factores influyentes en la variable dependiente  $Y$  pueden dividirse en dos grupos: uno que contiene una variable explicativa  $X$  y otro que incluye un conjunto de factores no controlados que se pueden englobar bajo el nombre de perturbación o error aleatorio  $\varepsilon$ , este error provoca que la dependencia entre las variables independiente y dependiente no sea perfecta, debido a que está sujeta a incertidumbre [33]. Es deseable que en el modelo de regresión estos errores aleatorios sean en media cero para cualquier valor  $x$  de  $X$ , el parámetro  $\beta_0$  es la ordenada al origen del modelo y  $\beta_1$  la pendiente, pudiéndose interpretar como el incremento de la variable dependiente por cada incremento en una unidad de la variable independiente. Estos parámetros son aquellos que deben ser estimados para poder realizar las predicciones.

### 2.4.2.3 Regresión lineal múltiple

El Análisis de Regresión Lineal Múltiple permite establecer la relación existente entre una variable dependiente y un conjunto de variables independientes  $(X_1, X_2, \dots, X_k)$ . Este modelo, a diferencia del simple, es capaz de aproximarse mejor a situaciones de análisis real, debido a que los fenómenos, hechos y procesos socioeconómicos son complejos, por lo tanto deben ser explicados en la medida de lo posible por todas las variables que participen directa e indirectamente [34].

Es frecuente que la variable dependiente y las independientes sean continuas. Sin embargo, también se puede aplicar el análisis cuando se relacione una variable dependiente continua con un conjunto de variables categóricas; de igual manera se puede aplicar cuando una variable dependiente nominal se relacione con un conjunto de variables continuas [34].

La notación matemática del modelo o ecuación de regresión lineal múltiple se establece en la ecuación 2.4.

$$Y = \beta_0 + \beta_1 X + \beta_2 X + \dots + \beta_n X + \varepsilon \quad (2.4)$$

Donde,  $Y$  es la variable a predecir,  $\beta_0, \beta_1, \dots, \beta_n$  son los parámetros desconocidos a estimar y  $\varepsilon$  es el error aleatorio [34].

### 2.4.2.4 Suavizado exponencial

La razón de que este método es llamado suavizado exponencial es porque cada incremento en el paso se reduce en  $(1 - \alpha)$ , donde  $\alpha$  es llamada constante de suavizamiento. El suavizamiento exponencial es una técnica que aplica la media móvil bajo el siguiente principio [35].

$$\text{Nueva predicción} = \text{vieja predicción} + \alpha(\text{última observación} - \text{vieja predicción})$$

Matemáticamente se expresa según la ecuación 2.5.

$$F_t = F_{t-1} + \alpha(A_{t-1} - F_{t-1}) \quad (2.5)$$

Donde

$F_t$ : Predicción exponencial suavizada para el periodo  $t$

$F_{t-1}$ : Predicción exponencial suavizada realizada para el periodo anterior

$A_{t-1}$ : Demanda actual en el periodo anterior

$\alpha$ : Constante de suavizamiento

#### **2.4.2.5 Métodos estadísticos**

Es la utilización del método científico por la estadística como un método científico de investigación teórica. El fundamento de este método lo constituye la aplicación y el desarrollo de las ideas de la teoría de las probabilidades como una de las disciplinas matemáticas más importantes [36].

El método de predicción probabilística predice la distribución de probabilidad para un intervalo de valores. No predice directamente un valor específico, sino que las probabilidades proyectadas son asignadas a cada valor en particular dentro del intervalo de valores. Existen dos métodos comúnmente empleados en la predicción probabilística, la función de densidad de probabilidad y la función de distribución acumulativa [37].

#### **2.4.2.6 Inteligencia artificial**

La inteligencia artificial (IA) se perfila como una nueva generación de tecnología informática, que se caracteriza no sólo por su arquitectura (hardware), sino también por sus capacidades. Hasta hace algunos años, el énfasis en el desarrollo de la IA estuvo relacionado a aplicaciones científicas o de negocios. La nueva generación de tecnología informática brinda de igual manera manipulación simbólica, con el objetivo de emular el comportamiento inteligente; y, la computación en paralelo que busca conseguir resultados en tiempo real. Dentro de la inteligencia artificial se encuentran los sistemas expertos, las redes neuronales artificiales, las redes Bayesianas y la lógica difusa. Estos métodos han cobrado relevante importancia en la actualidad debido a sus importantes aplicaciones prácticas [22].

#### **2.4.2.7 Micro áreas**

Una micro-área puede definirse como una subdivisión urbana del sistema de distribución que tiene una proyección demanda representando una parte de la demanda máxima prevista para el primario. Al realizar una división en micro áreas se facilita el uso de datos de energía, demanda, etc. Con respecto a los consumidores ubicados en la micro área. Es un método que involucra la división geográfica del área de suministro de una subestación [38].

#### **2.4.3 SEGÚN EL SISTEMA ELÉCTRICO OBJETIVO**

Los pronósticos de demanda eléctrica deben realizarse considerando uno de los siguientes componentes:

- Transformador
- Circuito (primario, secundario, ramal, etc.,)
- Subestación
- Usuario (residencial, comercial y/o industrial)
- Ciudades, zonas, regiones
- Cualquiera que sea de interés para la planificación

### **2.5 PROCESO PARA REALIZAR UNA PREDICCIÓN**

Se plantean 9 pasos o consideraciones al momento de realizar una predicción sin tomar en consideración la naturaleza de los datos. A continuación, se detallan los más relacionados con el problema de la predicción del consumo eléctrico a mediano plazo [39].

#### **Paso 1. Establecer objetivos**

Se deben especificar los objetivos relacionados a la situación. Luego se debe considerar que decisiones tomar para alcanzarlos. Los objetivos tendrán influencia en el tipo de procedimiento empleado en la predicción.

**Paso 2. Estructurar el problema**

El problema debe ser estructurado de tal manera que los planificadores puedan emplear los resultados para la toma de decisiones. Se deben determinar las variables o factores que causan los datos de una serie temporal. Se pueden clasificar como: crecimiento, decrecimiento, regresión, etc. Se debe establecer si las series temporales están marcadas por ciclos y/o tendencias.

**Paso 3. Obtener fuentes e información**

Corresponde a la búsqueda y obtención de fuentes e información que puede ser útil al realizar una predicción. Aunque la teoría es el fundamento, se debe emplear la creatividad para buscar información alternativa sustentada.

**Paso 4. Recolectar los datos**

Luego de haber identificado las fuentes, se debe recolectar datos confiables, válidos y relevantes.

**Paso 5. Preparar los datos**

Consiste en definir cambios, encontrar valores perdidos, ajustar errores. En este paso se limpian los datos.

**Paso 6. Selección de métodos**

Se deben seleccionar los métodos más apropiados, pues aunque existe una gran variedad solo algunos serán adecuados para el problema planteado.

**Paso 7. Implementación del método**

Se debe mantener la predicción lo más simple posible manteniendo una representación realista de la situación.

**Paso 8. Evaluación del método**

Se debe escoger algún criterio para conocer si el método tiene resultados satisfactorios.

## **Paso 9. Empleo del método para la predicción**

Finalmente se presentan las predicciones en conjuntos con los datos que soporten los resultados de una manera entendible.

### **2.6 SERIE TEMPORAL**

Una serie temporal es una secuencia de observaciones ordenadas. Considerando los datos que la integran pueden ser continuos o discretos. La serie continua es una variable cuyos datos se registran continuamente. Una serie discreta es una variable que toma valores específicos en un intervalo regular que se expresa en unidades de tiempo (días, semana, mes, año, etc.). Las series temporales tienen diferentes características según los datos que representa, por ejemplo puede tener comportamientos cíclicos, creciente, decreciente, estacionalidad o no estacionalidad, entre otros, que dependen exclusivamente del comportamiento de las observaciones [22].

Los modelos de series temporales se pueden clasificar como:

**Univariantes:** cuando se analiza una serie temporal en función de su pasado

**Multivariantes:** cuando se analizan más de una serie temporal simultáneamente. Este caso se tiene cuando existe una relación entre diversas variables que afecta la dinámica de ambas de series.

#### **2.6.1 COMPONENTES DE UNA SERIE DE TIEMPO**

##### **2.6.1.1 Tendencia**

Esta componente describe el comportamiento a lo largo del tiempo, se usa para detectar un crecimiento, decrecimiento o estacionalidad [40].

##### **2.6.1.2 Estacionalidad**

Son las fluctuaciones cíclicas relacionadas al calendario. Se define como estacional a una serie temporal que tiene un patrón de cambio a si mismo año tras año [40].

##### **2.6.1.3 Ciclo**

Se define como la fluctuación en forma de onda alrededor de la tendencia [40].

#### 2.6.1.4 Irregular

Son fluctuaciones sistemáticas o aleatorias. Se plantea como todo lo que sobra, es un factor residual [40].

#### 2.6.1.5 Medias Móviles

Consiste en determinar a partir de un conjunto de valores observados, el promedio de los mismos y luego utilizar este promedio para el siguiente pronóstico. La cantidad de observaciones empleadas para obtener el promedio es determinada por la persona que hace el pronóstico y permanecerá constante, por lo tanto, cada observación nueva producirá la eliminación de la observación más antigua, realizando el cálculo de la media con cada nuevo dato, es por ello que se denomina al modelo como “medias móviles”. La representación matemática se expresa en la ecuación 2.6.

$$Y_{t+1} = \frac{1}{n} \sum_{i=t-n+1}^t X_i \quad (2.6)$$

Donde:

$Y_t$ : Valor pronosticado al tiempo t

$X_t$ : Valor actual al tiempo t.

$n$  : Cantidad de datos que abarca el promedio

## 2.8 REDES NEURONALES ARTIFICIALES

Una red neuronal artificial (RNA) es un modelo inspirado en el sistema nervioso, consistiendo en un grupo interconectado de neuronas capaces de procesar información para obtener una salida inteligente. Una RNA aprende a ejecutar una tarea de acuerdo al algoritmo empleado para entrenarla. Un método que emplee RNA es llamado sistema adaptativo ya que es capaz de cambiar su estructura basado en información interna o externa que fluye a través de la red durante el entrenamiento [41]. Una red neuronal artificial es un método de procesamiento de información basado en la manera que el sistema nervioso procesa la información. Un humano tiene en promedio unos 100 billones de neuronas biológicas (se emplea el termino biológica para diferenciar de la neurona artificial) y alrededor de 100.000 conexiones por neurona. Las neuronas están formadas por tres componentes principales denominados dendritas, cuerpo o soma y axón [42].

Una neurona biológica recibe información a través de sinapsis ubicadas en las dendritas (árbol receptor de la neurona). Cuando recibe un impulso eléctrico por encima de un nivel determinado, la neurona se activa y envía una señal al axón. Esta señal viaja hasta otra neurona adyacente [42]. El desarrollo neuronal se caracteriza por ser modificable a través del aprendizaje de cada individuo mediante un proceso que involucra adquisiciones y pérdidas de conexiones neuronales, este desarrollo es crítico en los primeros años de vida de un individuo [43].

Las neuronas artificiales se constituyen en estructuras denominadas capas. Una red RNA está compuesta por un conjunto de capas. Por ende, la información se distribuye a lo largo de las sinapsis de la red, dándole al sistema cierta tolerancia a fallos. Las redes neuronales artificiales son capaces de adaptar su funcionamiento a múltiples entornos al modificar las conexiones entre neuronas, lo anterior se traduce en que puedan aprender de la experiencia y generalizar conceptos [44].

De manera similar, una RNA esta echa de simples conexiones entre elementos denominados neuronodos, celdas, unidad o elemento de procesamiento. Las cuales están arregladas en una estructura para formar una red que permite ejecutar y distribuir cálculos. Al igual que una red neuronal biológica, la RNA es capaz de aprender, reconocer patrones en los datos y adaptarse a los cambios, siendo esta una de las principales razones que ha llevado a su empleo en la modelación de datos [41]. Se puede comparar una neurona biológica con una neurona de un RNA considerando las semejanzas presentadas en la Figura 2.8.

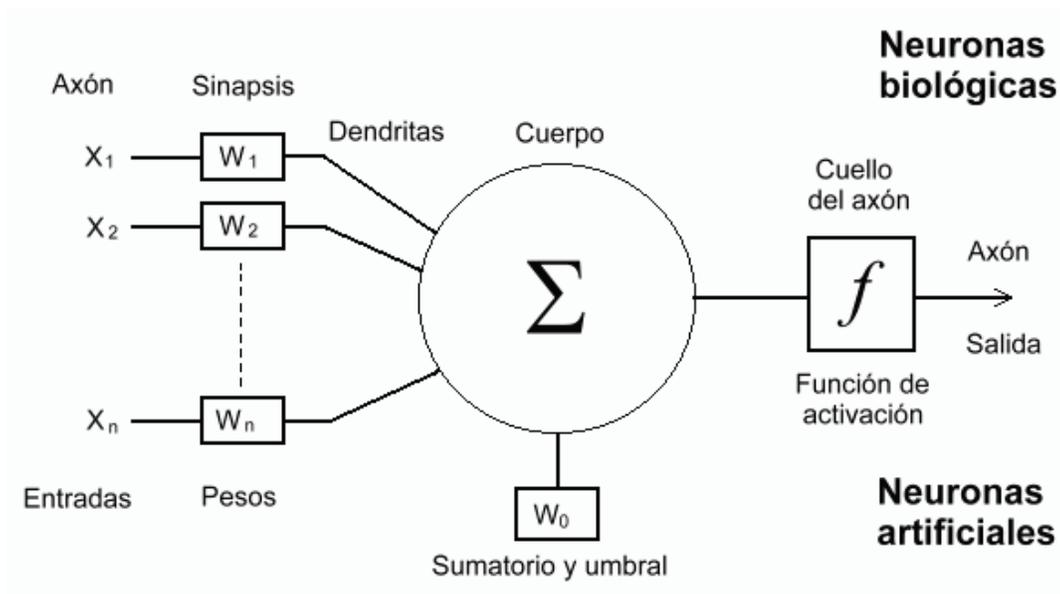


Figura 2.8 Neurona biológica y neurona artificial  
Fuente: Basado en [45]

### 2.8.1 MODELO DE NEURONA ARTIFICIAL

La neurona artificial es un elemento de procesamiento simple que a partir de un vector de entrada produce salidas. En general se encuentran tres tipos de neuronas artificiales, cada una de las cuales cuenta con una neurona similar en el sistema nervioso.

1. Aquellas que reciben información directamente desde el exterior, a estas neuronas se las denomina neuronas de entrada.
2. Aquellas que reciben información desde otras neuronas artificiales, a estas neuronas se las denomina neuronas ocultas. Es en estas neuronas se realiza la representación de la información almacenada.
3. Aquellas que reciben la información procesada para producir una salida, por ello se las denomina neuronas de salida [46].

Los aspectos que involucran una RNA pueden resumirse de acuerdo a los establecido en [47].

- Un conjunto de unidades procesadoras de información llamadas neuronas o nodos.

- Un estado de activación para todas las unidades equivalente a la salida de la unidad.
- Conexión entre las unidades, generalmente en cada conexión se define un peso con el propósito de determinar el efecto que tiene una señal proveniente de una unidad en otra.
- Una función de activación que determina el nivel de activación basado en la entrada a la neurona.
- Alguna entrada externa para las unidades.
- Un método para la obtención de la información o regla de aprendizaje.
- Un ambiente adecuado para que el sistema pueda operar que provea señales de entrada y si es necesario señales de errores.

La Figura 2.7 muestra los elementos que componen una neurona artificial:

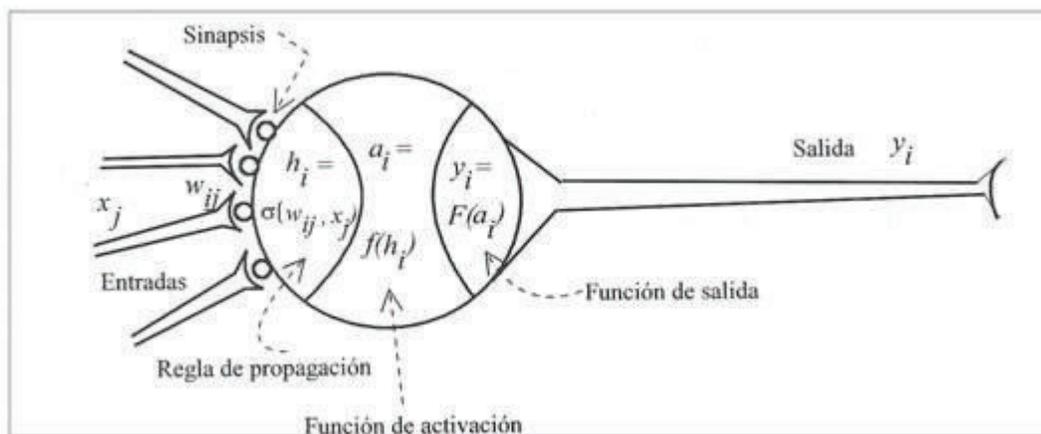


Figura 2.7 Estructura de una Neurona Artificial  
Fuente: Basado en [48]

**Entradas y salidas.** Las entradas  $x_j(t)$  pueden provenir del exterior o de otras neuronas artificiales. Las entradas y salidas de una neurona pueden ser clasificadas en binarias o continuas. Las neuronas binarias o digitales sólo admiten dos valores posibles mientras que las neuronas continuas o analógicas admiten valores dentro de un determinado rango, que usualmente se define como  $[-1, 1]$  [44].

**Peso sinápticos,  $w_{ij}$ .** Representan el grado de comunicación entre la neurona artificial  $j$  y la neurona artificial  $i$ . Pueden ser excitadores o inhibidores [46]. El peso sináptico define la fuerza de una conexión sináptica entre dos neuronas, la neurona  $i$  (presináptica) y la neurona  $j$  (postsináptica). Los pesos sinápticos pueden ser negativos, positivos o incluso cero [44].

**Regla de propagación,  $\sigma_i(w_{ij}, x_j(t))$ .** Integra la información proveniente de las distintas neuronas artificiales y proporciona el valor del potencial postsináptico de la neurona  $i$  [46].

**Función de activación.** La función de activación también es llamada función de transferencia, cumple con la siguiente  $a_i(t) = f_i(a_i(t-1), h_i(t))$ . Transforma la entrada a la red, así como el estado de activación previo  $a_i(t-1)$  en un nuevo estado de activación  $a_i(t)$  considerando el valor límite  $h_i(t)$  [1], [44].

**Función de salida  $F_i(a_i(t))$ .** Representa la salida actual de la neurona  $i$ . De esta forma, la salida producida por una neurona  $i$ , para un determinado instante de tiempo  $t$  puede ser escrita en forma general según la ecuación 2.7 [46].

$$y_i(t) = F_i(f_i[(a_i(t-1), \sigma_i(w_{ij}, x_j(t))]) \quad (2.7)$$

### 2.8.2 CAPACIDAD NO LINEAL DE LAS RNA

Las redes neuronales son un conjunto de circuitos no lineales que han demostrado la capacidad para resolver problemas no lineales. Este método también ha sido empleado para analizar sistemas dinámicos [49].

Las salidas de una RNA son funciones matemáticas lineales o no lineales de las entradas. Las entradas pueden ser la salida de otro elemento de la red, los elementos de la red están arreglados en un conjunto de capas de elementos ubicados entre las

entradas y las salidas. La salida se obtiene al ajustar los pesos de las neuronas en el perceptrón multicapa (cuando se emplea este tipo de neurona) [41].

### 2.8.3 ARQUITECTURA DE UNA RNA

#### 2.8.4 EL PERCEPTRÓN

Este modelo tiene gran importancia histórica ya que fue el primer modelo en poseer un mecanismo de entrenamiento que permite determinar automáticamente los pesos sinápticos que clasifican correctamente a un conjunto de patrones a partir de un conjunto de ejemplos. La arquitectura del perceptrón está compuesta por dos capas de neuronas, una de entrada y una de salida. La capa de entrada es la que recibe la información proveniente del exterior y la transmite a las neuronas sin realizar ningún tipo de operación sobre la señal de entrada. La operación de un perceptrón con  $n$  neuronas de entrada y  $m$  neuronas de salidas puede ser resumida según la ecuación 2.8 [46].

$$y_i = f(\sum_{j=1}^n w_{ij}x_j - \theta_i), \quad \forall i, 1 \leq i \leq m \quad (2.8)$$

Donde,  $y_i$  es la función de salida,  $w_{ij}$  los pesos sinápticos,  $x_j$  las entradas y  $\theta_i$  el valor umbral.

El perceptrón es una neurona básica, cuyo propósito es entender y describir el comportamiento del proceso que involucra la memoria y aprendizaje humano. El modelo matemático del perceptrón se origina de la arquitectura biológica considerando los axones y las sinapsis que son modelados como entradas y pesos [46].

Una vez definida el tipo de neurona que se utilizará en un modelo de redes neuronales artificiales es necesario definir la topología de la misma. La organización y disposición de las neuronas dentro de una red neuronal se denomina topología, y viene dada por el número de capas, la cantidad de neuronas por capa, el grado de conectividad, y el tipo de conexión entre neuronas. Las neuronas suelen agruparse en conjunto de neuronas dentro de unidades funcionales denominadas capas. Se denomina capa de entrada a la capa compuesta por neuronas de entradas que recibe información desde

el exterior. Similarmente, se denomina capa oculta y capa de salida a aquellas compuestas por neuronas ocultas y de salida respectivamente. Una red neuronal artificial está compuesta por una o más capas interconectadas entre sí. Las conexiones que pueden existir entre las neuronas son las sinapsis asociadas con un peso sináptico, siendo direccionales [44].

Cuando una conexión se establece entre dos neuronas de una misma capa se habla de conexiones laterales o conexiones intra-capas. Por otra parte, cuando la conexión se establece entre neuronas de capas diferentes se le llama conexión inter-capas. Cuando la conexión se lleva a cabo inversamente al sentido entrada-salida la conexión se llama recurrente o realimentada. Una red puede estar formada por una única capa de neuronas, denominadas redes monocapas. En este caso las neuronas cumplen la función de neuronas de entrada y salida simultáneamente. Cuando la red está compuesta por dos o más capas se denominan redes multicapas [44].

### 2.8.5 EL PERCEPTRÓN MULTICAPA

El perceptrón multicapa se forma a partir de generalizar el perceptrón simple. Está basado en el uso de varias capas de neuronas artificiales, en vez de usar una sola capa, gracias a modificar una función no diferenciable como la activación logística (empleada en el perceptrón simple) a una función diferenciable y no lineal como la sigmoidea. Esta arquitectura permite a las RNA aprender variaciones lineales o no lineales de diferentes ambientes [50].

Al añadir capas intermedias (ocultas) a un perceptrón simple se obtiene un perceptrón multicapa o MLP (*Multi-Layer Perceptron*). Esta arquitectura generalmente se entrena haciendo uso del algoritmo denominado retropropagación de errores o BP (Back Propagation) de ahí que se conozca también bajo el nombre de red de retropropagación [42]. La topología *feedforward* permite que las señales viajen en una sola dirección, desde la entrada a través de la capa oculta hasta la salida. La capa de entrada representa los datos que han sido procesados y “alimentan” la red. La entrada

de cada capa oculta se define como la suma de los vectores de entrada y los pesos de las conexiones entre la capa de entrada y la capa de salida. La entrada de la salida se define por la suma de las salidas “pesadas” de la capa oculta de neuronas. Las salidas se definen por el tipo de función de transferencia empleada [42].

La arquitectura de una RNA multicapa se muestra en la Figura 2.9

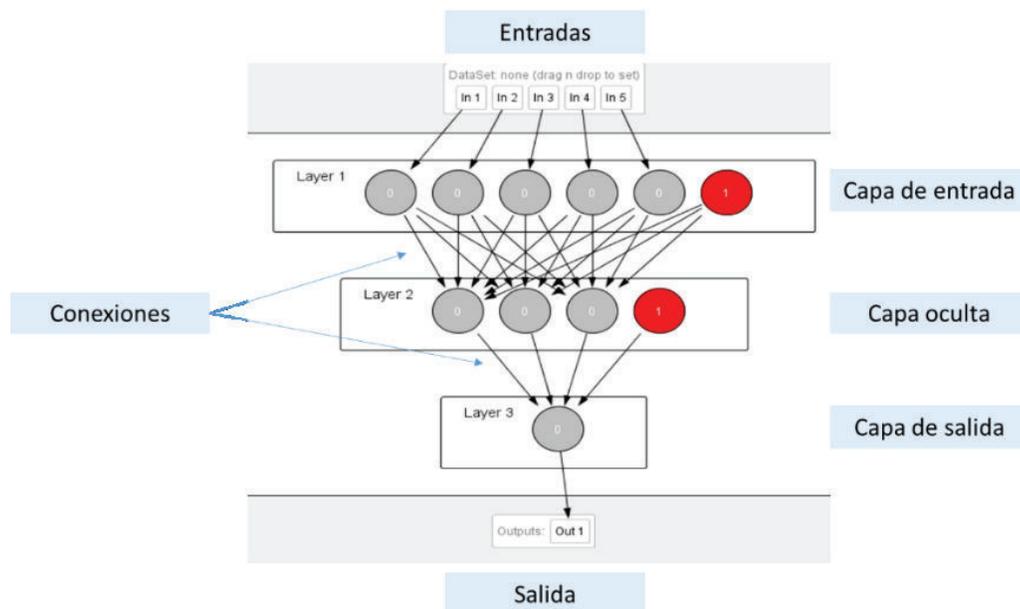


Figura 2.9. Arquitectura de una RNA  
Fuente: Basado en [41].

El espectro de aplicaciones del perceptrón multicapa es muy amplio lo que hace muy difícil enumerar sus aplicaciones más relevantes [51]. Sin embargo, dentro de las áreas de aplicación se puede mencionar:

- Codificación de información
- Traducción de texto en lenguaje hablado
- Predicción de carga eléctrica
- Reconocimiento óptico de caracteres (OCR) [46].

### 2.8.6 FUNCIONES DE TRANSFERENCIA

La función de activación o función de transferencia es la regla que sigue la red para denotar la activación de una neurona. Conecta los pesos de una neurona a la entrada determinando la activación o el estado de la neurona. Existen diversas funciones de transferencia disponibles, incluidas función lineal, senoidal, tangente hiperbólica, sigmoïdal, rampa, escalón y otras [52]. La función de transferencia puede ser lineal o no lineal [2].

#### 2.8.6.1 Función sigmoïdal

El entrenamiento mediante el algoritmo de propagación hacia atrás de una RNA multicapa necesita de funciones de activación suaves [49]. La palabra sigmoïdal significa que la representación tiene forma de S [53]. Esta función está definida en un intervalo con límites superiores e inferiores unitarios. Es particularmente importante para redes que emplean el algoritmo de propagación siendo fácil calcular su derivada [52], [53]. Se expresa según la ecuación 2.9. En la Figura 2.10 está representada la función sigmoïdal.

$$y(x) = \frac{1}{1+e^{-x}} \quad (2.9)$$

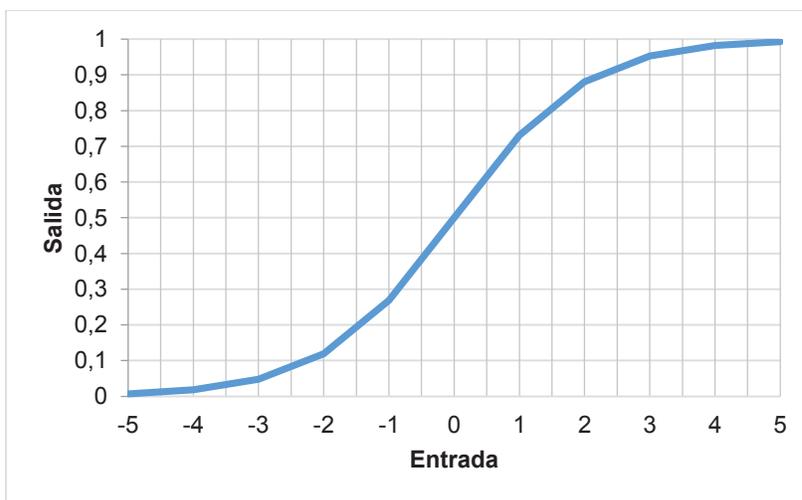


Figura 2.10 Función sigmoïdal

Fuente: Elaboración propia basada en ecuación 2.9

### **2.8.7 ENTRENAMIENTO DE LA RED**

Una de las principales características de las RNA es su capacidad de aprendizaje. El entrenamiento de las RNA muestra algunos paralelismos con el desarrollo intelectual de los seres humanos. No obstante aun cuando parece que se ha conseguido entender el proceso de aprendizaje conviene ser moderado porque el aprendizaje de las RNA está limitado [52].

El objetivo del entrenamiento de una RNA es conseguir que para un conjunto de entradas se obtenga un conjunto de salidas deseadas o consistentes. El entrenamiento consiste en la aplicación de diferentes vectores de entrada con el propósito de ajustar los pesos de las interconexiones según un algoritmo predeterminado. Durante el entrenamiento los pesos convergen eventualmente hacia valores que garanticen que el vector de salida cumpla con los parámetros deseados [52].

#### **2.8.7.1 Errores de la red**

En el proceso de entrenamiento de la red se debe considerar, junto al error cometido respecto al aprendizaje (error de entrenamiento), el denominado error de generalización medible mediante un conjunto de prueba (*test set*) distinto al empleado en el entrenamiento. Generalmente, es más importante una buena generalización de la red (resultados consistentes para cualquier entrada diferente al conjunto de entrenamiento) que un error muy reducido en la submuestra de entrenamiento. La red debe ser capaz de generalizar, para ello las entradas desconocidas son comparadas con las conocidas y la salida se origina mediante una clase de interpolación de los valores aprendidos [54].

Tras una etapa inicial, en la que la tasa de error puede oscilar, el error de aprendizaje disminuye monótonamente, mientras que el error de generalización hasta cierto punto, para incrementarse después, como consecuencia del excesivo ajuste de la red a las particularidades de los patrones empleados en el entrenamiento, que son memorizado en perjuicio de la capacidad de generalización del sistema. Este fenómeno conocido como sobreaprendizaje (*overtraining*), puede evitarse utilizando procesos de

validación cruzada (*cross validation*), esto es, entrenando y validando la red a la vez, a fin de detectar el punto óptimo de aprendizaje. Estos procesos de validación son ampliamente aplicados en la fase de desarrollo de redes supervisadas [42].

En la Figura 2.11 se muestra que el tiempo óptimo de entrenamiento debe estar en el círculo rojo para mantener mínimo el error de generalización. En la práctica puede tener cierta complejidad encontrar esta área idónea [41]. Una manera simple de entrenar la red es hacerlo hasta que el error en la validación comience a crecer, en este punto se espera que el error por generalización crezca de igual manera. No obstante, uno de los problemas potenciales relacionados es que el error de validación puede subir y bajar varias veces durante el entrenamiento [55].

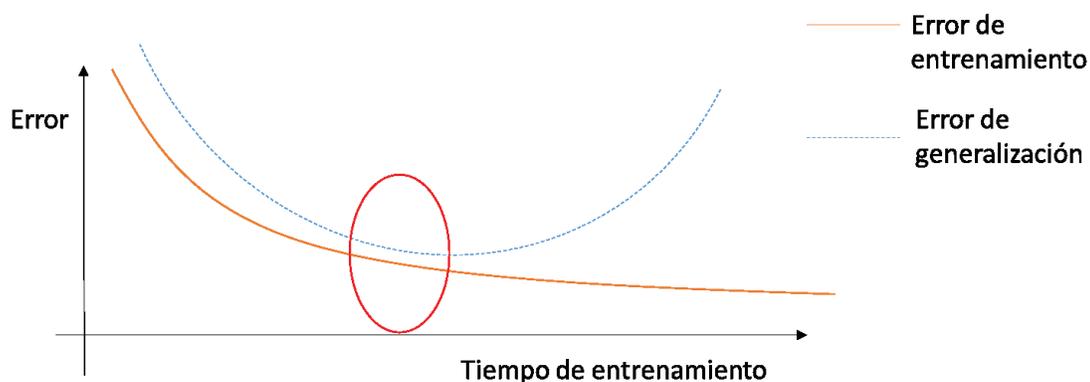


Figura 2.11 Errores de la red  
Fuente: Basado en [41]

### 2.8.7.2 Tipos de entrenamientos

- **Entrenamiento supervisado**

En este entrenamiento se presenta a la red la salida deseada para ejemplo de entrenamiento. Al inicio del entrenamiento, se asignan valores aleatorios a los pesos, a partir de este punto se presentan los valores del set de entrenamiento (training set) a la red [49]. Este tipo de entrenamiento tiene lugar cuando la entrada y la salida son claramente definidas. Por ejemplo, el vector de entrada es representado como

$[x_1, \dots, x_m]$  y el correspondiente vector de salida como  $[y_1, \dots, y_m]$  por lo tanto la salida se relaciona con la entrada de la siguiente manera:

$$[y_1, \dots, y_m] = [x_1, \dots, x_m] + \varepsilon$$

Donde  $\varepsilon$  es el error en la aproximación.

En este caso la RNA busca ajustar los pesos de los enlaces para minimizar el error en la aproximación comparando el resultado obtenido con el deseado. Luego el error retroalimenta la red con el propósito de encontrar el mejor modelo posible [54].

- **Algoritmo de aprendizaje durante el entrenamiento**

El algoritmo de backpropagation es el método de entrenamiento más utilizado en redes con conexión hacia adelante (*feedforward*). Es un método de aprendizaje supervisado de gradiente descendiente, en el que se distinguen claramente dos fases: en la primera fase se aplica un patrón de entrada que propaga por las distintas capas que forman la red hasta generar una salida. Esta salida se compara con la deseada y se calcula el error incurrido por cada neurona. Estos errores se transmiten hacia atrás, desde la capa de salida hacia todas las neuronas de las capas intermedias. Cada neurona recibe un error proporcional a la contribución realizada sobre el error total de la red. Basándose en el error recibido, se ajustan los errores de los pesos sinápticos de cada neurona [46].

En la Figura 2.12 se muestra un diagrama simplificado usando el método backpropagation.

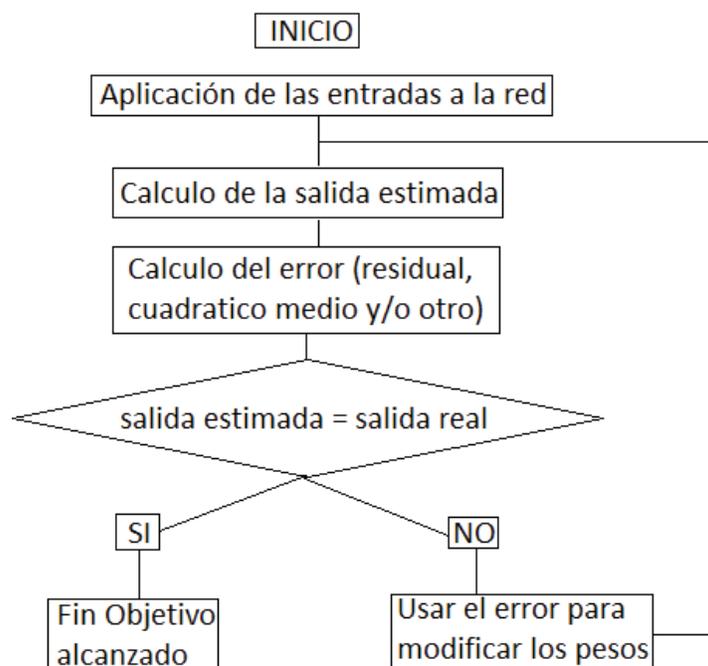


Figura 2.12 Diagrama del algoritmo empleado durante el entrenamiento (backpropagation)

Fuente: Basado en [41]

## 2.9 VARIABLES

Es una característica o atributo que puede asumir diferentes valores y se dividen en cualitativas y cuantitativas.

### 2.9.1 VARIABLES CUALITATIVAS

Son aquellas que no aparecen de forma numérica, sino como categorías o atributos. Este tipo de variables se dividen en:

- **Variables nominales**

Estas variables son consideradas como atributos. Son aquellas en las cuales sus valores no se pueden ordenar, es decir no reflejan ninguna jerarquía. Por ejemplo: presencia o ausencia de energía eléctrica (0/1).

- **Variables ordinales**

Los atributos que se registran pueden mantener entre sí una relación de jerarquía como por ejemplo puede los meses del año, la pertenencia a un grupo socioeconómico.

## 2.9.2 VARIABLES CUANTITATIVAS

Son aquellas que pueden expresarse numéricamente (temperatura, número de habitantes en la zona de estudio). A su vez este tipo de variables se dividen de la siguiente forma:

- **Discretas**

Estas variables pueden tomar un conjunto finito o numerable de valores (generalmente valores enteros). Por ejemplo, número de estaciones en la red eléctrica.

- **Continuas**

Las variables continuas pueden incluir un amplio espectro de valores, incluyendo los intervalos que hay entre números enteros. Por ejemplo, la demanda eléctrica instantánea.

## 2.10 CRITERIOS DE VALIDACIÓN Y ANÁLISIS

### 2.10.1 ERROR CUADRÁTICO MEDIO

El error cuadrático medio (ECM) es una medida de la dispersión entre los datos previstos y los datos históricos, representado según la ecuación 2.10.

$$ECM = \frac{1}{N} \sum_{i=1}^N (Y_i - \hat{Y}_i)^2 \quad (2.10)$$

Donde:

$Y_i$ : Valor observado para la i-ésima observación (datos históricos)

$\hat{Y}_i$ : Valor proyectado para la i-ésima iteración

$N$ : Cantidad de datos

Este error es uno de los usados más comúnmente para la validación de modelos [56].

### 2.10.2 ERROR RESIDUAL

Es la diferencia entre el valor pronóstico y la observación real. Se calcula empleando la ecuación 2.11.

$$e_r = Y_i - \hat{Y}_i \quad (2.11)$$

Donde:

$Y_i$ : Valor real en la i-ésima observación (datos históricos)

$\hat{Y}_i$ : Valor pronosticado para la i-ésima observación

### 2.10.3 DESVIACIÓN MEDIA ABSOLUTA

Permite tener una medida del error global del pronóstico, se calcula al sumar los valores absolutos de los errores residuales y dividirlos entre el número de periodos según la ecuación 2.12.

$$DMA = \frac{1}{N} \sum_{i=1}^N |Y_i - \hat{Y}_i| \quad (2.12)$$

Donde:

$Y_i$ : Valor real para la i-ésima observación (datos históricos)

$\hat{Y}_i$ : Valor pronosticado para la i-ésima iteración

$N$ : Cantidad de datos

### 2.10.4 ERROR ESTÁNDAR MÚLTIPLE

Mide la dispersión de los datos obtenidos de la predicción. El cálculo se realiza mediante la ecuación 2.13.

$$EEM = \sqrt{\frac{\sum_{i=1}^N (Y_i - \hat{Y}_i)^2}{(N-k-1)}} \quad (2.13)$$

Donde:

$Y_i$ : Valor real para la i-ésima observación (datos históricos)

$\hat{Y}_i$ : Valor pronosticado para la i-ésima iteración

$N$ : Cantidad de datos

$K$ : Número de variables independientes

### 2.10.5 PORCENTAJE DE ERROR MEDIO ABSOLUTO

Permite expresar los errores porcentualmente, lo que generalmente resulta más útil que los valores absolutos. Indica el “tamaño” de los pronósticos con respecto a los datos reales de las observaciones. La ecuación que permite su cálculo es la 2.14.

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|Y_i - \hat{Y}_i|}{Y_i} \quad (2.14)$$

Donde:

$Y_i$ : Valor real para la  $i$ -ésima observación (datos históricos)

$\hat{Y}_i$ : Valor pronosticado para la  $i$ -ésima iteración

$N$ : Cantidad de datos

## 2.11 VALORES ATÍPICOS Y CRITERIOS PARA ENCONTRARLOS

Un valor atípico es un dato significativamente diferente al resto de datos. Los datos atípicos son también conocidos como valores anormales, discordantes o anomalías. Un dato atípico contiene información acerca de características anormales de un sistema. La salida de un algoritmo de detección de valores atípicos puede ser uno de dos tipos [57].

- La mayoría del algoritmo de detección de valores atípicos emite una puntuación sobre el nivel de "outlierness" de un punto de datos. Esto puede usarse para determinar una clasificación de los puntos de datos en términos de su tendencia del valor atípico. Esta es una forma muy general de salida, que retiene toda la información proporcionada por un algoritmo particular, pero sí no proporciona un resumen conciso del pequeño número de puntos de datos se deben considerarse valores atípicos.

- Un segundo tipo de salida es una etiqueta binaria que indica si un punto de datos es un *outlier* o no. Aunque algunos algoritmos devuelven las etiquetas binarias, las puntuaciones externas también pueden convertirse en etiquetas binarias. Esto suele hacerse imponiendo puntuaciones *Outlier*, sobre la base de su distribución estadística. Un binario contiene menos información que un mecanismo de puntuación, pero es el resultado final que a menudo se necesita para la toma de decisiones en aplicaciones prácticas.

En la Figura 2.13 se describe el espectro desde los datos normales hasta los atípicos.



Figura 2.13 Espectro de datos, desde normales hasta atípicos  
Fuente: basado en [57]

Considerando observaciones que reflejan cierta cantidad de datos cuantitativos unidimensionales denotados como  $X_1 \dots X_N$  con una media  $\mu$  y desviación estándar  $\sigma$ . Se define el valor  $Z$  como la cantidad de desviaciones estándar por la cual los datos varían de la media [57].

$$Z_i = \frac{|X_i - \mu|}{\sigma} \quad (2.15)$$

Usualmente se usa  $Z_i \geq 3$  como regla general para determinar la anormalidad cuando se tienen datos suficientes y la media se mantiene constante. Existen diferentes pruebas que emplean la ecuación anterior como principio para determinar valores atípicos.

# CAPÍTULO 3

## CAPÍTULO III

### ANÁLISIS DE DATOS HISTÓRICOS Y METODOLOGÍA

En este capítulo se describen los procedimientos empleados para la obtención y procesamiento de datos y la posterior obtención del modelo.

#### 3.1 PREPARACIÓN DE LA INFORMACIÓN

La recolección adecuada de datos es fundamental para llevar a cabo la investigación, datos inadecuados o insuficientes pueden tener impactos en el modelo hasta el punto que los resultados pueden ser inválidos [58]. En este trabajo, los datos empleados para obtener el modelo de predicción de la demanda eléctrica fueron obtenidos de la Empresa Eléctrica Riobamba S.A. A continuación se detalla ciertas consideraciones empleadas en la preparación y posterior uso de los datos. Se consideró la predicción a mediano plazo como objetivo principal, para lograrlo se plantea la preparación de la información considerando datos horarios y su conversión a datos de consumo de energía mensual.

##### 3.1.1 MEDICIÓN DE DATOS

En el Mercado Eléctrico Mayorista (MEM) cada participante es responsable de tener una metodología para la medición de la energía que se inyecta o extrae de un nodo, con el propósito de monitorear constantemente la energía consumida o suministrada. Estas mediciones involucran transformadores de corriente y potencial asociados a medidores comunicados con los centros de despacho de carga. A efectos de este estudio, los datos históricos fueron recolectados mediante medidores instalados en las salidas RIOBAMBA 1, RIOBAMBA 2 y RIOBAMBA 3 de la barra de 69 kV de subtransmisión de la subestación Riobamba, ver diagrama unifilar en el Anexo 1.

En las Figuras 3.1 y 3.2 se muestran ejemplos del consumo energético horario y mensual obtenido a partir de los datos. Inicialmente se muestra el comportamiento de

un día cualquiera para la demanda, usualmente el consumo máximo se da en la noche, alrededor del intervalo 6-8 pm, cuando las familias llegan a casa.

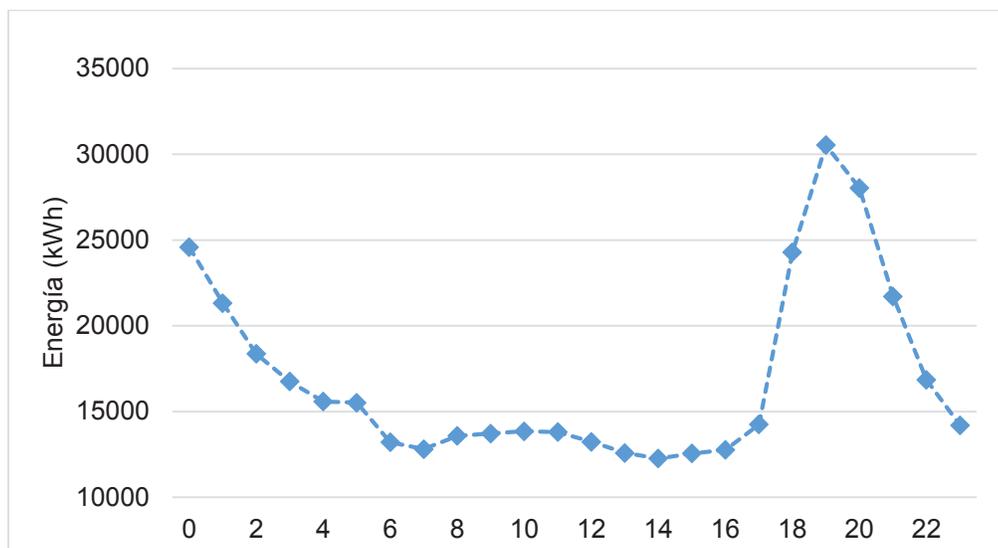


Figura 3.1. Consumo diario (horario) para un día aleatorio  
Fuente: Elaboración propia con datos de la EERSA

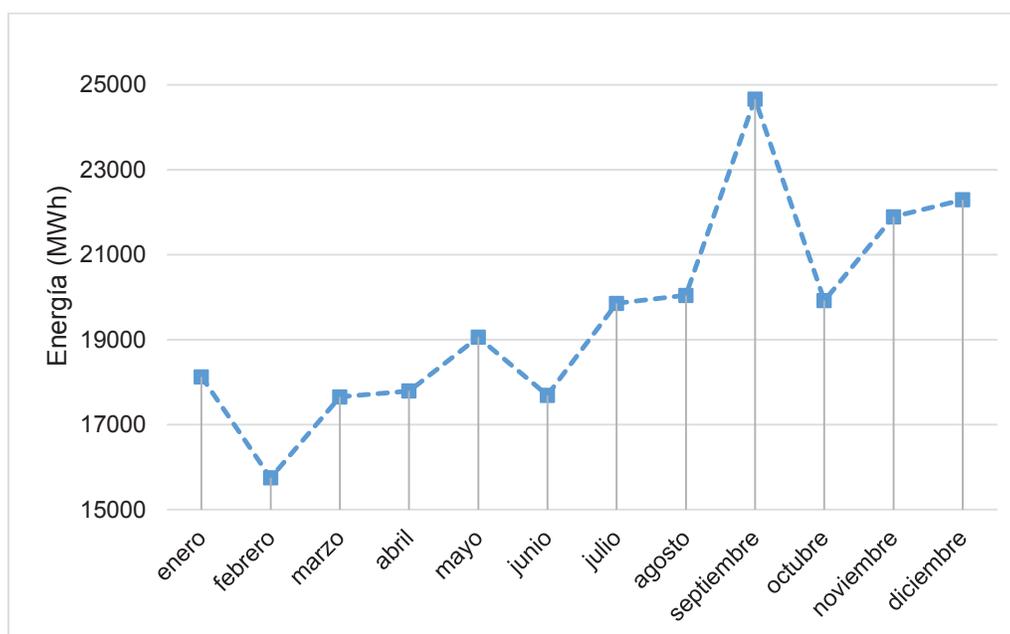


Figura 3.2. Consumo mensual para el año 2015  
Fuente: Elaboración propia con datos de la EERSA

El objetivo principal es encontrar un modelo válido que permita predecir la demanda eléctrica en barra de 69 kV del Sistema Eléctrico de Subtransmisión de la Empresa Eléctrica Riobamba S.A., estos datos corresponden a la energía consumida durante cada mes en MWh, para ello es necesario crear una tabla con datos mensuales (12 meses) desde el 2008 hasta el 2015 (8 años) para un total de 96 datos mensuales que serán empleados como entradas en el entrenamiento de la red neuronal y un total de 12 datos mensuales correspondientes al año 2016 para la validación de la red.

### 3.1.2 ALMACENAMIENTO Y PROCESAMIENTO DE DATOS

Todos los datos registrados por los medidores son almacenados y procesados en archivos Excel, para cada día transcurrido se crea un archivo donde se guardan diferentes variables de importancia para la empresa. Dentro de estos datos, se emplearon los relacionados al consumo de energía activa (kWh) por punto de entrega del distribuidor.

Se obtuvieron datos desde enero del 2008 hasta diciembre del 2016, cada registro contiene la energía medida durante el intervalo de una hora. Se empleó la energía total consumida resultante de la suma de las demandas horarias de energía para entrenar y obtener el modelo de la red neuronal artificial según la ecuación 3.1 basada en [58].

$$Energía_{a,m} = \sum_{d=1}^{d=j} \sum_{h=1}^{h=24} Energía_{a,d,m,h} \quad (3.1)$$

Donde:

$Energía_{a,m}$ : Consumo de energía para el mes  $m$  del año  $a$

$Energía_{a,d,m,h}$ : Consumo de energía para la hora  $h$  del día  $d$  en el mes  $m$  del año  $a$

La suma de las energías se plantea para las 24 horas de cada día para los días desde  $d = 1$  hasta  $d = j$  donde  $j$  depende de la cantidad de días que tenga un mes en particular.

Todos los datos mensuales se encuentran tabulados con detalle en el Anexo 2. Los datos indican un incremento promedio año-año de 7,6% este valor es obtenido al aplicar la media aritmética según lo establecido en las ecuaciones 3.2 y 3.3 basadas en [59].

$$I_i = \frac{\text{Energía}_{a+1,m}}{\text{Energía}_{a,m}} \quad (3.2)$$

Donde:

$\text{Energía}_{a+1,m}$ : Consumo de energía para el año  $a + 1$  en el mes  $m$

$\text{Energía}_{a,m}$ : Consumo de energía para el año  $a$  en el mes  $m$

$I_i$ : Incremento  $i$

$$\bar{x} = \frac{I_1 + I_2 + \dots + I_i}{i} \quad (3.3)$$

Donde:

$I_i$ : Incremento  $i$

$\bar{x}$ : Media aritmética

En la Figura 3.3 se detalla la serie de tiempo para los datos en MWh. El gráfico muestra cierto grado de estacionalidad de 12 meses con una significativa tendencia a incrementarse la media. Con esta tendencia, se espera que los valores de la proyección excedan a los años previos.

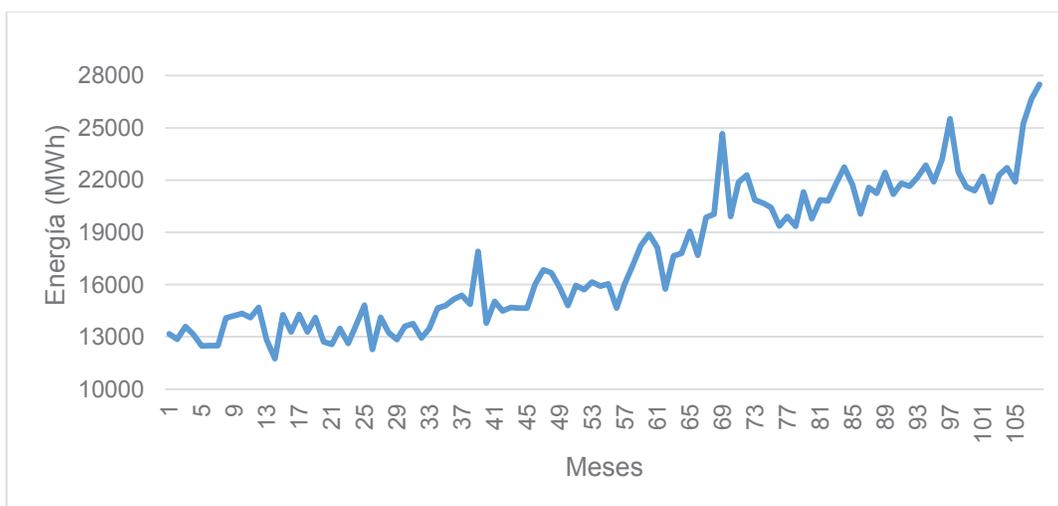


Figura 3.3. Datos mensuales del consumo de energía desde el año 2008 al 2016  
Fuente: Elaboración propia con datos de la EERSA

### 3.1.3 SELECCIÓN DE ENTRADAS DEL MODELO

Considerando los factores que influyen en el pronóstico a mediano plazo mencionados en el Capítulo 2 donde se plantea tomar en cuenta otros factores más allá de los climatológicos y conforme a los resultados de diversas investigaciones [23], [25], [30], se seleccionan los datos económicos (Valor bruto agregado) y el crecimiento demográfico (incremento anual de población) como las variables independientes explicativas más influyentes de la demanda eléctrica a mediano plazo. Esta selección se verifica al encontrar la correlación entre las variables y el comportamiento del consumo de energía eléctrica detallado en la sección 3.5 del Capítulo 3.

Los datos de planes de desarrollo públicos y privados, así como el aumento de la densidad son muy sesgados, pueden tener variaciones abruptas en un periodo corto de tiempo y afectar el modelo.

No se considera la temperatura como variable independiente explicativa puesto que la provincia de Chimborazo se encuentra en una latitud y longitud que no presenta grandes cambios de temperatura durante el año, en la Figura 3.4 se muestra un diagrama de temperatura histórica para la provincia, donde se observa un comportamiento bastante lineal a lo largo del año, manteniendo una temperatura promedio de unos 14° [60], [61].

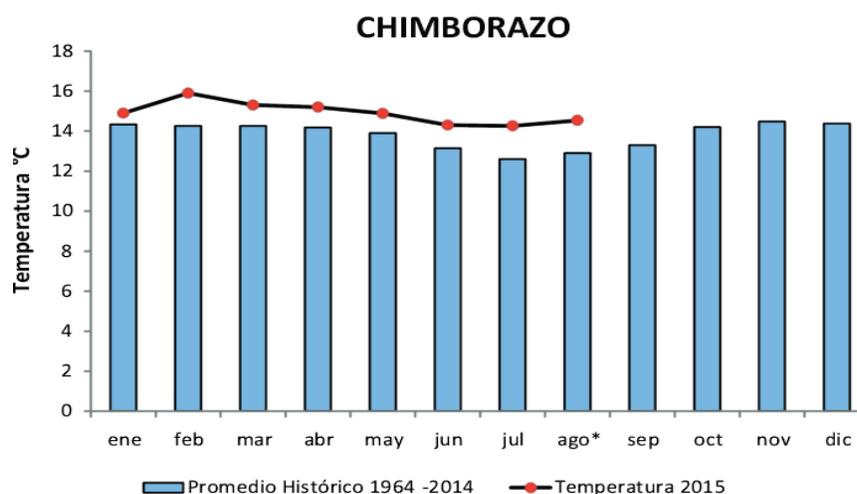


Figura 3.4. Temperatura anual de Chimborazo  
Fuente: Obtenido de [60]

### 3.1.3.1 Datos económicos

Los datos de entrada corresponden a datos de la actividad económica anual de la provincia de Chimborazo, los datos correspondientes al intervalo 2008-2015 fueron obtenidos del Banco Central del Ecuador [62]. Los datos correspondientes a los años 2016 y 2017 se proyectan considerando el crecimiento interanual promedio (3,99%) obtenido mediante la media aritmética obtenida de las ecuaciones 3.4 y 3.5. basadas en [59]. Para los casos donde la salida es mensual se emplea una interpolación de los datos anuales a datos cuatrimestrales. En la tabla 3.1 se muestran los datos anuales y en el Anexo 3 se muestran los datos cuatrimestrales. La representación gráfica se puede observar en la figura 3.5.

$$I_i = \frac{VAB_{a+1}}{VAB_a} \quad (3.4)$$

Donde:

$VAB_{a+1}$ : Valor agregado bruto para el año  $a + 1$

$VAB_a$ : Valor agregado bruto para el año  $a$

$I_i$ : Incremento  $i$

$$\bar{x} = \frac{I_1 + I_2 + \dots + I_i}{i} \quad (3.5)$$

Donde:

$I_i$ : Incremento  $i$

$\bar{x}$ : Media aritmética

Tabla 3.1. Datos de valor agregado bruto (PIB) para la provincia Chimborazo  
Fuente: Banco Central del Ecuador [62]

Año	PIB (miles de \$)
2008	805.360
2009	929.211
2010	1.048.682
2011	1.093.801
2012	1.287.639
2013	1.324.051
2014	1.420.254
2015	1.670.132
2016	1.808.975
2017	2.003.983

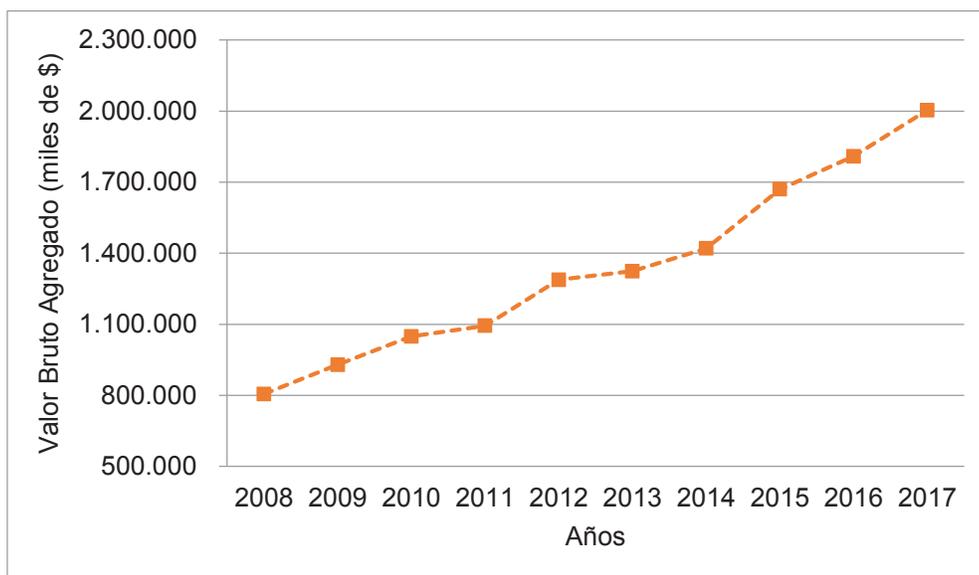


Figura 3.5. Datos económicos para la provincia Chimborazo (PIB).  
Fuente: Banco Central del Ecuador [62].

### 3.1.3.2 Datos demográficos

Los datos de población fueron obtenidos de los censos realizados por el Instituto Nacional de Estadística. No obstante, el último censo realizado fue en el 2010. Debido a esto se carecen de datos anuales correspondientes. Para resolver este planteamiento se consideran las proyecciones desarrolladas por el INEC [25] para la provincia de Chimborazo, estos datos se muestran en la tabla 3.2. Sin embargo, la proyección no presenta datos específicos para los años 2009 y 2008. La proyección

se realiza considerando una extrapolación considerando el mismo planteamiento evaluado en los datos económicos según la tabla 3.3. El resumen de los datos anuales empleados se muestra en la tabla 3.4 y en la figura 3.6 se muestra la representación gráfica de la población.

Tabla 3.2. Proyección poblacional 2011-2017  
Fuente: Elaboración propia con datos del INEC [25]

<b>Proyección población provincia Chimborazo</b>	481.498	486.680	491.753	496.735	501.584	506.325	510.935
<b>Año</b>	2011	2012	2013	2014	2015	2016	2017
<b>Porcentaje de incremento con respecto al año anterior</b>	1,10%	1,08%	1,04%	1,01%	0,98%	0,95%	0,91%

Tabla 3.3. Proyección poblacional 2008-2010  
Fuente: Elaboración propia con datos del INEC [25]

<b>Año</b>	2008	2009	2010	2011
<b>Población según censo</b>				228.226
<b>Extrapolación de población Chimborazo</b>	466.634	471.420	476.255	

Tabla 3.4. Resumen de datos demográficos.  
Fuente: Elaboración propia con datos del INEC [25]

<b>Año</b>	<b>Población (habitantes)</b>
<b>2008</b>	466.634
<b>2009</b>	471.420
<b>2010</b>	476.255
<b>2011</b>	481.498
<b>2012</b>	486.680
<b>2013</b>	491.753
<b>2014</b>	496.735
<b>2015</b>	501.584
<b>2016</b>	506.325
<b>2017</b>	510.935

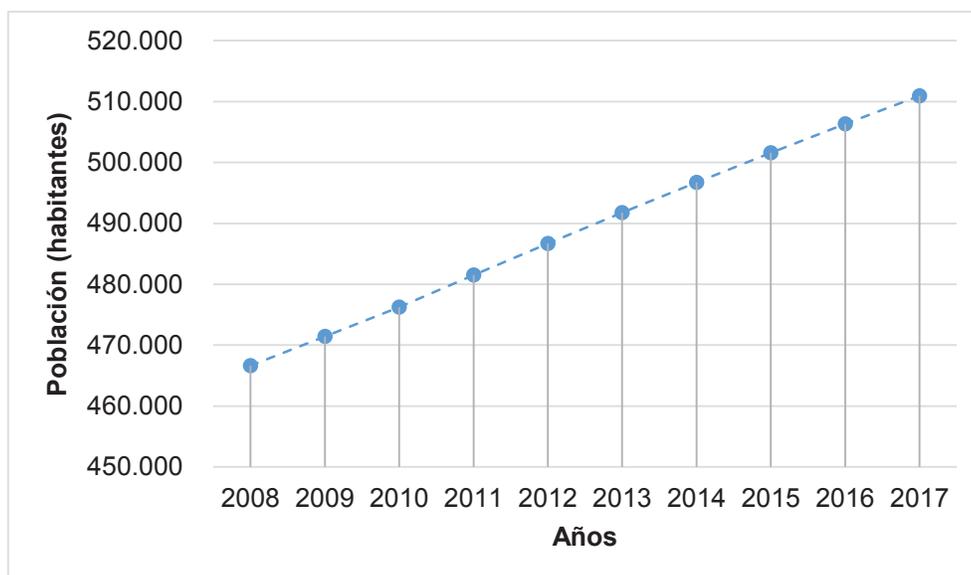


Figura 3.6. Proyección de crecimiento poblacional en Chimborazo.  
Fuente: Elaboración propia con datos del INEC [25]

### 3.1.3 UNIFICACIÓN DE DATOS

Se debe unificar todos bajo un archivo conteniendo las siguientes variables: energía (MWh), fecha, datos económicos (PIB) y población (habitantes). Esta unificación permite un manejo más cómodo y sencillo de la información. Este archivo es la base para la creación de los data set a ser empleados en el entrenamiento y validación de los datos.

La tabla 3.5 muestra un ejemplo del conjunto de datos que sirven de entrada y salida a la RNA. Es importante destacar que este conjunto de datos se obtienen de los valores horarios de energía luego de realizar modificaciones necesarias según varios datos atípicos encontrados.

Tabla 3.5. Datos unificados para el modelo  
Fuente: Elaboración propia con datos de la EERSA

<b>Año</b>	<b>Mes</b>	<b>Energía (MWh)</b>	<b>Población (habitantes)</b>	<b>PIB (miles de \$)</b>
2010	Enero	14818	476.255	1.048.682
2010	Febrero	12276	476.255	1.048.682
2010	Marzo	14120	476.255	1.048.682
2010	Abril	13251	476.255	1.048.682
2010	Mayo	12835	476.255	1.063.722
2010	Junio	13606	476.255	1.063.722
2010	Julio	13759	476.255	1.063.722
2010	Agosto	12931	476.255	1.063.722
2010	Septiembre	13484	476.255	1.078.762
2010	Octubre	14660	476.255	1.078.762
2010	Noviembre	14795	476.255	1.078.762
2010	Diciembre	15156	476.255	1.078.762

#### 3.1.4.1 Valores atípicos

Dado que los datos de una serie temporal proviene de datos recolectados, siempre existe un margen en el cual pueden ocurrir hechos muy diferentes a los demás. Las observaciones pueden ser afectadas por: cortes de energía, grandes trasposos de carga, mantenimientos programados, mediciones incorrectas, falla en los equipos de medición, fallas en las comunicaciones, etc. Es importante destacar que algunos valores atípicos no son datos erróneos sino son consecuencia del manejo de carga, en ciertas ocasiones es necesario hacer botes de carga para garantizar la estabilidad del sistema, lo que trae consecuencia en el comportamiento a mediano plazo de los datos.

Es importante poder identificar estos valores y tomar acciones correctivas, caso contrario puede sesgar el modelo y por ende obtener predicciones incorrectas. En caso de que el valor atípico se encuentre en la parte final de la serie, es usual que se generen predicciones incorrectas aunque los parámetros estén bien estimados. Por otra parte, si los valores tienen un comportamiento no aleatorio, es posible identificarlos e incorporarlos en las predicciones para obtener resultados más realistas [43].

### 3.1.4.2 Prueba de Grubbs

Esta prueba permite determinar valores atípicos de un conjunto de datos. Sin embargo, no se debe tomar la prueba como algo definitivo y verdadero, ya que puede reflejar falsos positivos dependiendo de la distribución de datos, es por ello que luego de ejecutar la prueba se procedió a verificar uno a uno los posibles valores atípicos. Se empleó el programa computacional NCSS 11 para obtener los resultados, los resultados detallan aquellos posibles valores atípicos. En la Figura 3.7 se detallan todos los datos horarios donde pueden observar valores que presentan una gran desviación con respecto a los demás datos.

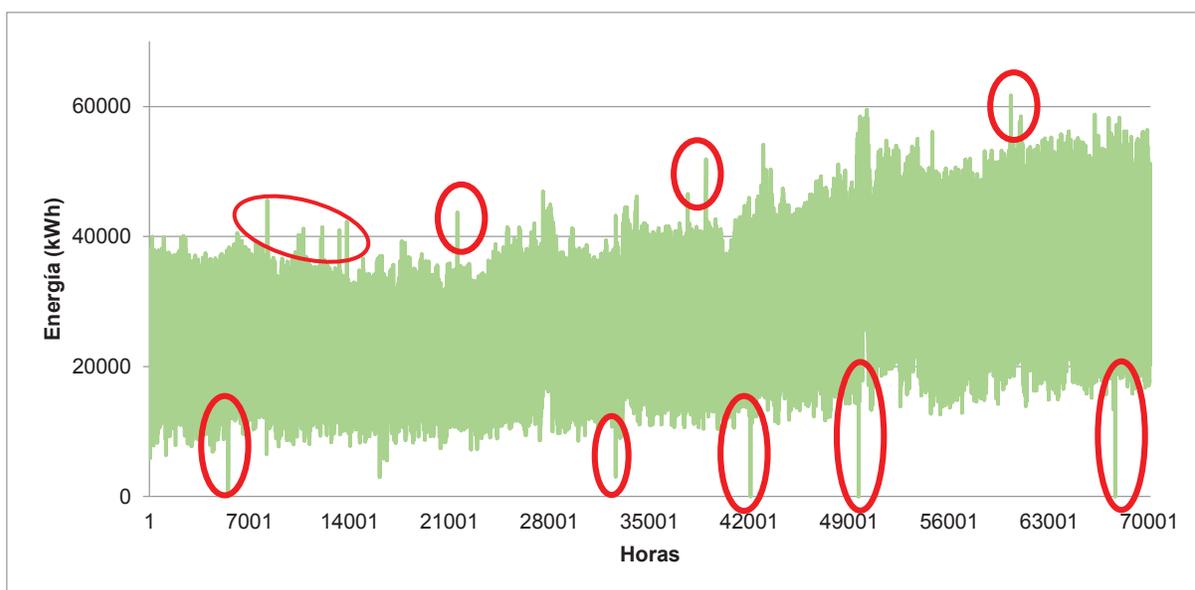


Figura 3.7. Prueba de Grubbs para encontrar valores atípicos.

Fuente: Elaboración propia

### 3.1.4.3 Datos perdidos

Los datos con mayor desviación corresponden a datos perdidos, esto puede ocurrir por las diversas causas mencionadas. Sin embargo, dada la naturaleza de la recolección de datos es difícil obtener una causa específica. En la Tabla 3.6 se detallan datos con valor igual a 0 kWh, lo que representan valores perdidos.

Tabla 3.6. Valores perdidos  
Fuente: Elaboración propia con datos de la EERSA

Energía (kWh)	Hora
0,0	16/08/2008 08:00
0,0	16/08/2008 09:00
0,0	16/08/2008 10:00
0,0	16/08/2008 11:00
0,0	16/08/2008 12:00
0,0	16/08/2008 13:00
0,0	16/08/2008 14:00
0,0	16/08/2008 15:00
0,0	16/08/2008 16:00
0,0	21/10/2012 08:00
0,0	21/10/2012 09:00
0,0	21/10/2012 10:00
0,0	01/09/2013 03:00
0,0	01/09/2013 04:00
0,0	01/09/2013 05:00
0,0	01/09/2013 06:00
0,0	20/09/2015 02:00
0,0	20/09/2015 03:00
0,0	20/09/2015 04:00
0,0	20/09/2015 05:00

Para realizar la sustitución de estos datos se realizó una interpolación considerando lo planteado para casos similares a [58]. Se especifican los años adyacentes para el mismo día y la misma hora según la ecuación 3.6 basada en [58].

$$Energía_{s,a,m,d,h} = \frac{Energía_{h,a-1,m,d,h} + Energía_{h,a+1,m,d,h}}{2} \quad (3.6)$$

Donde:

$Energía_{s,a,m,d,h}$ : Consumo de energía a sustituir para la hora  $h$  del día  $d$  del mes  $m$  del año  $a$ .

$Energía_{h,a-1,m,d,h}$ : Consumo de energía del año anterior para la misma fecha

$Energía_{s,a+1,m,d,h}$ : Consumo de energía del año posterior para la misma fecha

Aplicando la formulación anterior se obtienen los valores presentados en la Tabla 3.7

Tabla 3.7 Valores sustituidos (perdidos)  
Fuente: Elaboración propia

Energía (kWh)	Hora
16304,0	16/08/2008 08:00
17749,0	16/08/2008 09:00
18192,0	16/08/2008 10:00
18877,0	16/08/2008 11:00
16792,0	16/08/2008 12:00
15487,0	16/08/2008 13:00
14848,0	16/08/2008 14:00
14446,0	16/08/2008 15:00
14184,0	16/08/2008 16:00
22748,4	21/10/2012 07:00
23022,7	21/10/2012 08:00
23778,9	21/10/2012 09:00
16392,7	01/09/2013 03:00
17161,0	01/09/2013 04:00
20166,1	01/09/2013 05:00
21707,0	20/09/2014 02:00
21652,1	20/09/2014 03:00
22631,1	20/09/2014 04:00
24613,0	20/09/2014 05:00

#### 3.1.4.4 Otros valores atípicos

Existen otros valores atípicos que corresponden a valores que están muy por debajo y/o muy por encima de la media para una hora determinada. En la tabla 3.8 se muestran estos datos.

Tabla 3.8 Demanda atípicamente baja  
Fuente: Elaboración propia

Energía (kWh)	Fecha
85,0	16/08/2008 07:00
6531	05/12/2008 12:00
2976,1	02/11/2009 10:00
3045,4	22/09/2011 16:00
3665,7	02/11/2009 09:00
4164,7	02/11/2009 11:00
51913,5	12/06/2012 19:00
61696,4	19/11/2014 19:00
56051,1	04/04/2014 19:00

Estos datos fueron encontrados gracias a la prueba de Grubbs. Sin embargo, dada la naturaleza de los datos fue necesario verificar que en realidad fuesen valores atípicos. Para ello, se comparó la energía en horas y fechas cercanas. Se concluyó que ciertamente eran valores atípicos. En la Tabla 3.9 se muestra un ejemplo de la diferencia entre la energía para las 10:00 de los diferentes años y se evidencia que la energía del 2009 es atípica.

Tabla 3.9 Demanda por debajo del promedio  
Fuente: Elaboración propia con datos de la EERSA

Energía (kWh)	Fecha
2976,1	02/11/2009 10:00
18837,5	02/11/2010 10:00
22750,5	02/11/2011 10:00
19766,7	02/11/2012 10:00
28801,1	02/11/2013 10:00
24262,8	02/11/2014 10:00
23516,3	02/11/2015 10:00

De igual manera se compara con valores cercanos y se evidencia que son valores atípicos puesto que la diferencia es elevada (Tabla 3.10). La curva de demanda generalmente no muestra cambios abruptos o semi-abruptos [10]. Por lo tanto, en una hora no cambia tan drásticamente la demanda, se concluye que es un dato atípico.

Tabla 3.10 Demanda por debajo del promedio (2)  
Fuente: Elaboración propia con datos de la EERSA

Energía (kWh)	Fecha
15339,3	02/11/2009 05:00
15952,7	02/11/2009 06:00
15906,6	02/11/2009 07:00
16165,5	02/11/2009 08:00
3665,7	02/11/2009 09:00
2976,1	02/11/2009 10:00
4164,7	02/11/2009 11:00
15614,7	02/11/2009 12:00
14234,1	02/11/2009 13:00
14620,3	02/11/2009 14:00
13975,4	02/11/2009 15:00

Para sustituir estos valores se aplicó la ecuación 3.7 basada en un caso similar en [58].

$$Energía_{s,a,m,d,h} = \frac{Energía_{h-1,a,m,d} + Energía_{h+1,a,m,d}}{2} \quad (3.7)$$

Donde:

$Energía_{s,a,m,d,h}$ : Consumo de energía para la hora  $h$  del día  $d$  del mes  $m$  del año  $a$  que será sustituida

$Energía_{h-1,a,m,d}$ : Consumo de la hora anterior inmediata para la misma fecha

$Energía_{h+1,a,m,d}$ : Consumo de la hora posterior inmediata para la misma fecha

Si bien esta sustitución no garantiza la correcta representación de los datos, si se disminuye en una gran proporción la posibilidad de tener datos sesgados en los resultados [58].

$$\%valores\ atípicos = \frac{36\ observaciones}{70080\ observaciones} \times 100 = 0,051\%$$

El porcentaje de valores atípicos con respecto al total de observaciones es apenas 0,051% por lo tanto la sustitución no tendrá un impacto significativo en la elaboración de los modelos.

### 3.2 NORMALIZACIÓN DE DATOS

Los datos se normalizaron de acuerdo a las ecuaciones 3.8 y 3.9 basadas en [41].

$$Valor\ normalizado = \frac{valor\ actual - valor\ minimo}{valor\ máximo - valor\ minimo} \quad (3.8)$$

$$Valor\ normalizado = \frac{valor\ actual}{valor\ base} \quad (3.9)$$

La ecuación 3.8 se empleó para normalizar los datos de población como se muestra en la Tabla 3.11 y datos económicos según la Tabla 3.12

Tabla 3.11 Datos de población normalizados  
Fuente: Elaboración propia

Año	Población (habitantes)	Población normalizada
<b>2008</b>	466.634	0
<b>2009</b>	471.420	0,108
<b>2010</b>	476.255	0,2172
<b>2011</b>	481.498	0,3355
<b>2012</b>	486.680	0,4525
<b>2013</b>	491.753	0,567
<b>2014</b>	496.735	0,6795
<b>2015</b>	501.584	0,7889
<b>2016</b>	506.325	0,8959
<b>2017</b>	510.935	1

Tabla 3.12 Datos económicos normalizados  
Fuente: Elaboración propia

Año	PIB (miles de \$)	PIB normalizado
<b>2008</b>	805.360	0
<b>2009</b>	929.211	0,1033
<b>2010</b>	1.048.682	0,203
<b>2011</b>	1.093.801	0,2406
<b>2012</b>	1.287.639	0,4024
<b>2013</b>	1.324.051	0,4327
<b>2014</b>	1.420.254	0,513
<b>2015</b>	1.670.132	0,7215
<b>2016</b>	1.808.975	0,8373
<b>2017</b>	2.003.983	1

Mientras que la ecuación 3.4 se empleó para normalizar todos los datos de demanda aplicando un valor base de 40 000 MWh para las demandas de energía mensual.

### 3.3 RESOLUCIÓN DEL PROBLEMA

Las predicciones con RNA involucran datos históricos recientes, lo que puede plantearse como un problema cuando no existan suficientes datos o se duda la validez de los mismos, considerando que el modelo reconoce los patrones de las demandas anteriores y predice extrapolando la demanda futura usando datos recientes de demanda [63]. Entonces, se plantea la necesidad de datos recientes. Por lo tanto, se requiere que la RNA tenga como entrada la demanda previa desde 1 hasta 12 meses según cada caso de estudio de la demanda a mediano plazo [64].

Para la obtención de la demanda a mediano plazo se plantean dos problemas.

En el primer problema, dada una serie de demandas mensuales hasta el año  $a$  en el mes  $m$ , el objetivo es predecir el perfil que tendrá la demanda para el mes  $m + 1$ . Posteriormente, mediante una secuencia de 12 iteraciones el modelo permite determinar la demanda para los próximos 12 meses considerando que  $m + 1$  se convierte en  $m$  para la próxima iteración, es decir, el valor proyectado en una iteración dada alimenta la entrada para la siguiente iteración. Esto tomando en consideración lo propuesto en [64].

En el segundo problema, dada una serie de demandas mensuales hasta el año  $a$  en el mes  $m$ , el objetivo es predecir el perfil que tendrá la demanda para los meses  $m + 1, m + 2, \dots, m + 12$  correspondiente a 12 meses de manera simultánea. Este problema difiere al anterior debido a que se obtienen 12 salidas. Este planteamiento se basa en [64]. Se plantean ambos problemas con el propósito de obtener la red que mejor modele los datos. En la Tabla 3.13 obtenida de [64] donde se muestran diversas consideraciones a tomar en la configuración de las redes.

Tabla 3.13 Configuraciones empleadas en [64]

Entradas	Estructura del modelo
Todas las 12 entradas	12-4-1
	12-5-1
	12-6-1
3 entradas solamente $D(m - 1)$ , $D(m - 5)$ , $D(m - 10)$	3-5-1

Los planteamientos anteriores son generales, para determinar los específicos se emplean funciones de correlación y auto correlación con el propósito de encontrar los rezagos óptimos.

### 3.4 HERRAMIENTAS PARA LA PREDICCIÓN DE LA DEMANDA

#### 3.4.1 SERIES TEMPORALES

Para analizar la serie temporal y obtener su auto correlación y correlación con las variables de población y datos económicos, se empleó el programa computacional NCSS 11, software estadístico de fácil uso para resolución de problemas que impliquen representación, descripción y análisis de una gran cantidad de datos. En la Figura 3.8 se muestra una ventana del NCSS 11 correspondiente al proceso de obtención de la auto correlación de los datos. En el apartado 3.5 se detalla el procedimiento realizado.

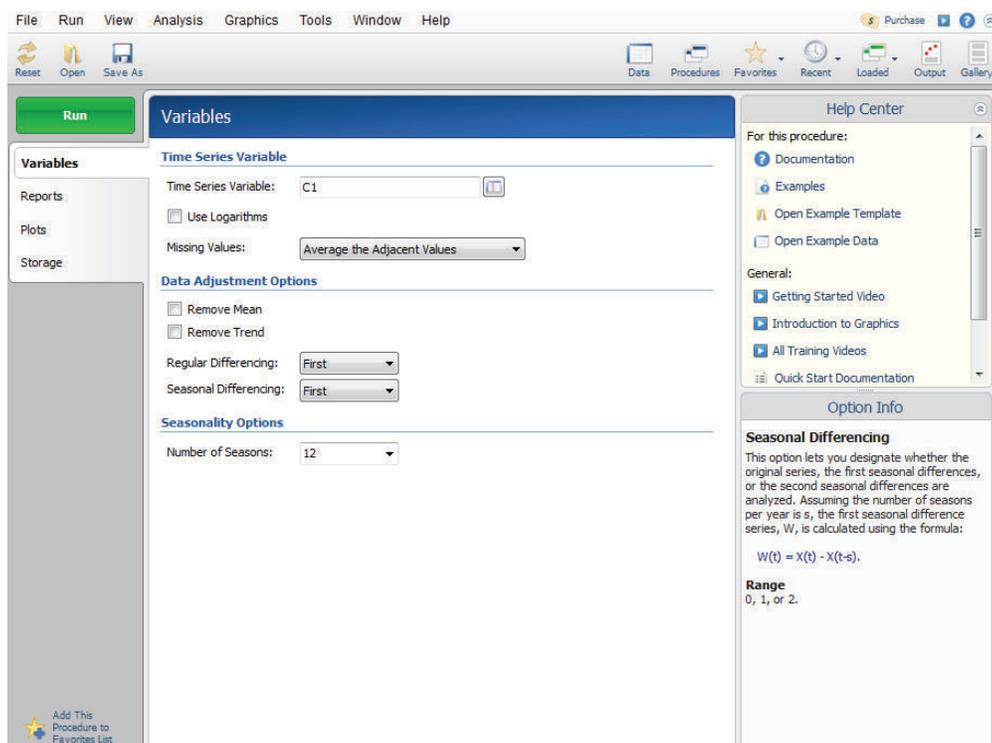


Figura 3.8 Ventana del programa computacional NCSS 11  
Fuente: programa computacional NCSS 11

### 3.4.2 RED NEURONAL ARTIFICIAL

Para el procesamiento de los datos y su aplicación a la RNA se empleó el programa computacional Neuroph Studio, el cual es un framework elaborado en lenguaje JAVA que provee una manera fácil de usar herramientas y asistentes en la creación de redes neuronales, a diferencia de años anteriores en los que se debía desarrollar el código en JAVA y aparte la creación de las redes neuronales. En el anexo IV se detalla el procedimiento para crear y entrenar una RNA.

#### 3.4.2.1 Selección del tipo de red neuronal

Una red neuronal multicapa alimentación hacia adelante (*feed-forward*) se ha empleado con resultados satisfactorios para la proyección de una serie temporal de datos [65]. Adicionalmente, las RNA tipo perceptrón multicapa tiene un desempeño adecuado para tareas de predicción [66]. De igual manera una RNA tipo perceptrón multicapa puede ser una herramienta efectiva para la predicción de datos al realizar

un modelo entre las entradas y las salidas al emplear datos históricos [67]. Asimismo se plantea que una RNA perceptrón multicapa es capaz de obtener mejores resultados que otros métodos de pronóstico si se configura de manera adecuada [68].

#### **3.4.2.2 Selección de la función de transferencia**

El método de propagación hacia atrás (backpropagation) y la función de activación sigmoideal son los métodos más importantes y empleados para el entrenamiento de RNA multicapas de alimentación hacia adelante (feed forward) para minimizar el error cuadrático medio [69]. Similarmente una función sigmoideal es usada en la capa oculta de neuronas para otorgarle a la RNA la capacidad de aprender posibles funciones no lineales [70]. Una RNA tipo perceptrón multicapa que tenga una función de activación no lineal es más robusta a valores atípicos y se adapta mejor en caso de series temporales no estacionarias [71]. El empleo de esta función de transferencia es preferible para tareas de predicción [72]. En vista de los anteriores planteamientos, se decide emplear la función sigmoideal como función de transferencia.

#### **3.4.2.3 Capas ocultas y neuronas**

El número de entradas y salidas de la RNA debe ser igual al set de entrenamiento. El usuario debe decidir el número de capas ocultas y de neuronas en cada capa. Es raro encontrar un problema que requiera el empleo de más de una capa oculta [24]. Por lo tanto, para la investigación se emplea el uso de una sola capa oculta. Siendo más importante el número de neuronas que están en la capa.

La cantidad de neuronas en la capa oculta no debe ser muy baja ni muy alta, ya que de ser muy pequeña con respecto al número de entrada ocasiona una dificultad en detectar las señales, mientras que una cantidad elevada ocasiona que los datos puedan ser insuficientes para entrenar adecuadamente a la red [41].

Deben tomarse las siguientes consideraciones al momento de definir el número de neuronas ocultas [73].

- El número de neuronas ocultas debe encontrarse entre el tamaño de la capa de entrada y la capa de salida.
- El número de las neuronas ocultas debe ser  $2/3$  del tamaño de la capa de entrada más la capa de salida
- El número de las neuronas ocultas debe ser menor al doble de la capa de entrada

#### 3.4.2.4 Numero de neuronas de salida.

Según el caso analizado se emplean doce (12) o una (1) sola neurona como salida de la red, con el propósito de obtener la red con menor error posible.

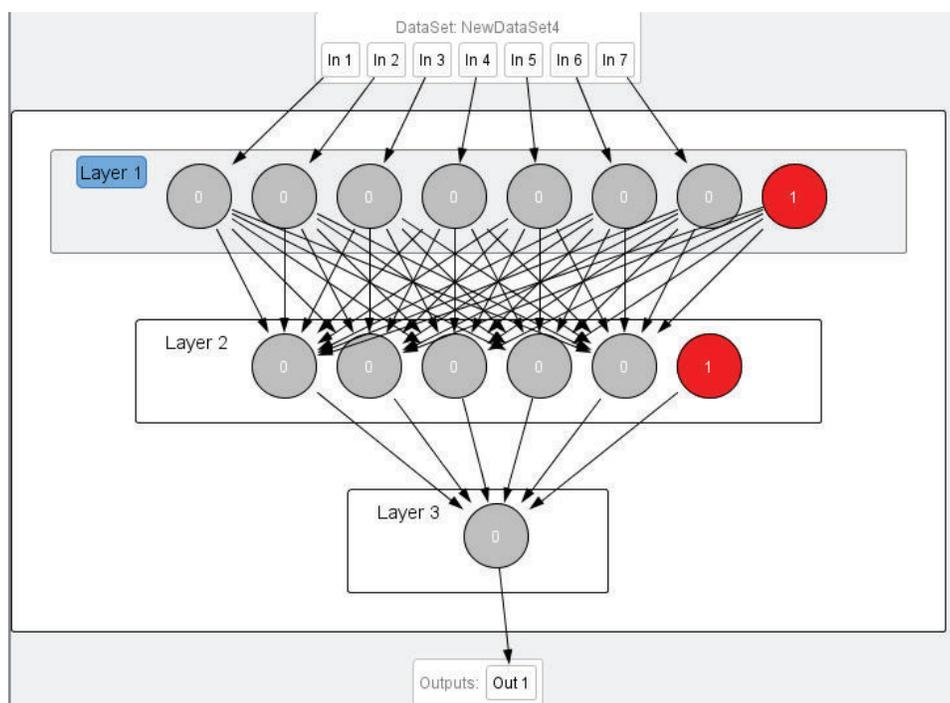


Figura 3.9 Ejemplo de Red Neuronal Artificial creada en el Neuroph Studio  
Fuente: Elaboración propia

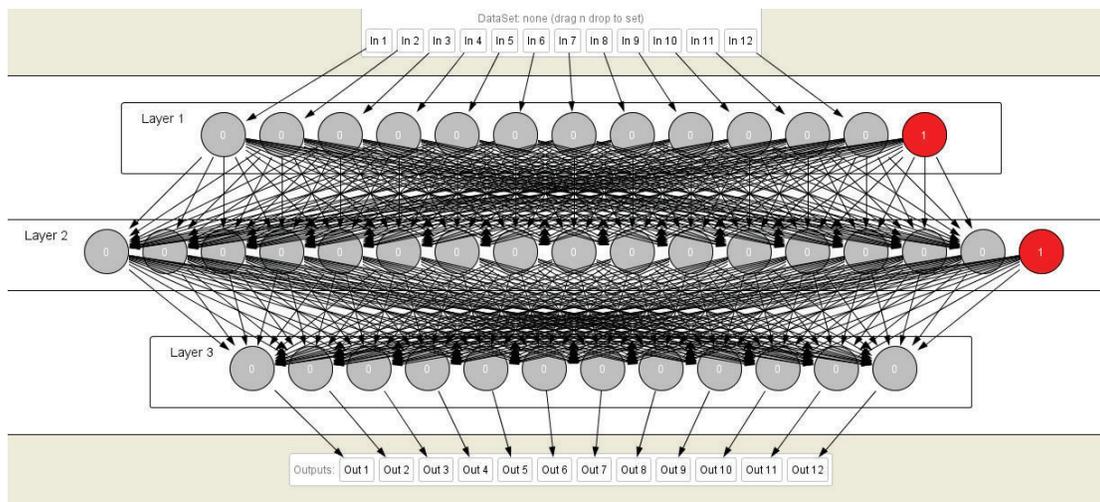


Figura 3.10 Ejemplo de Red Neuronal Artificial con 12 salidas creada en el Neuroph Studio

Fuente: Elaboración propia

### 3.4.2.5 Método de entrenamiento

El Neuroph Studio permite el uso del entrenamiento hacia atrás (*backpropagation*) que es uno de los más comúnmente usados como método de entrenamiento. Emplea el error residual para retroalimentar al sistema y de esta manera iterar hasta obtener el mejor resultado posible bajo unas condiciones preestablecidas.

Para alimentar a la red se debe cargar un archivo delimitado por coma, espacio o tabulación. En este caso se emplearon archivos .csv delimitados por coma. Se debe especificar el número de entrada y el número de salidas, las entradas están asociadas a las primeras columnas del archivo de entrada mientras que las salidas están asociadas a las columnas posteriores a las de entrada.

A medida que aumenta la cantidad de iteraciones se producirá un menor error en el entrenamiento pero hasta llegar a cierto límite a partir del cual el error permanece prácticamente invariante o puede incluso aumentar. El número de iteraciones se establece en 45 000, pero se considera que el error no aumente luego de alcanzar alguna iteración.

La tasa de entrenamiento se establece en 0,1 considerando que este valor ha resultado adecuado para predicciones. El máximo error se establece en 0,0001 como valor objetivo, aunque raramente se puede obtener un error tan reducido.

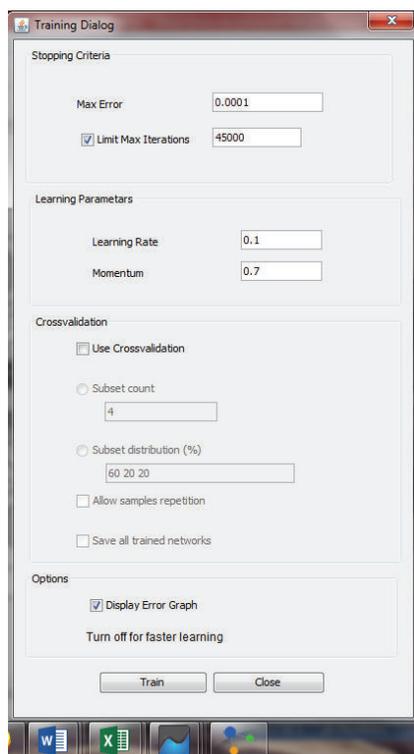


Figura 3.11 Ventana de configuración del entrenamiento  
Fuente: Elaboración propia

Luego de hacer clic en entrenar (Train) comienza el proceso de entrenamiento, en el transcurso se muestra gráficamente el error actual de entrenamiento como se muestra en la Figura 3.12.

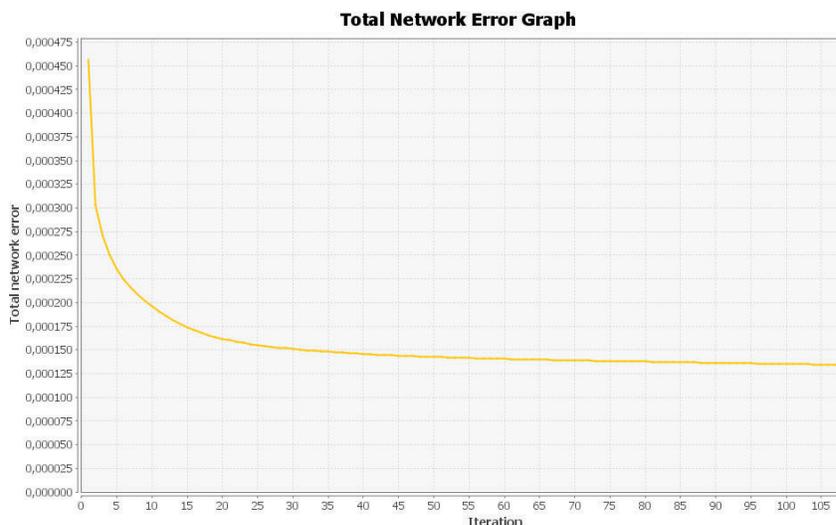


Figura 3.12. Ejemplo del error de entrenamiento

Fuente: Elaboración propia

Posteriormente se ejecuta una prueba o test que tendrá como resultado los datos proyectados y la comparación con los datos reales, se calcula automáticamente el error residual para cada observación y al final se muestra el error cuadrático medio (Figura 3.13).

```

Input: 0,37; 0,409; 0,4928; 0,586; 0,6693; 0,3681; 0,3386; Output: 0,3399; Desired output: 0,3508; Error: -0,0108;
Input: 0,3508; 0,37; 0,409; 0,4928; 0,586; 0,3662; 0,339; Output: 0,3395; Desired output: 0,3379; Error: 0,0016;
Input: 0,3379; 0,3508; 0,37; 0,409; 0,4928; 0,3792; 0,3527; Output: 0,3422; Desired output: 0,3446; Error: -0,0024;
Input: 0,3446; 0,3379; 0,3508; 0,37; 0,409; 0,4246; 0,4; Output: 0,3769; Desired output: 0,3861; Error: -0,0092;
Input: 0,3861; 0,3446; 0,3379; 0,3508; 0,37; 0,3727; 0,3948; Output: 0,3992; Desired output: 0,3778; Error: 0,0214;
Input: 0,3778; 0,3861; 0,3446; 0,3379; 0,3508; 0,4234; 0,4338; Output: 0,3997; Desired output: 0,4129; Error: -0,0132;
Input: 0,4129; 0,3778; 0,3861; 0,3446; 0,3379; 0,4274; 0,4675; Output: 0,4457; Desired output: 0,4393; Error: 0,0064;
Input: 0,4393; 0,4129; 0,3778; 0,3861; 0,3446; 0,4519; 0,4872; Output: 0,4623; Desired output: 0,4658; Error: -0,0035;
Input: 0,4658; 0,4393; 0,4129; 0,3778; 0,3861; 0,4612; 0,4807; Output: 0,4737; Desired output: 0,4713; Error: 0,0024;
Input: 0,4713; 0,4658; 0,4393; 0,4129; 0,3778; 0,4612; 0,4786; Output: 0,4697; Desired output: 0,4692; Error: 0,0005;
Input: 0,4692; 0,4713; 0,4658; 0,4393; 0,4129; 0,4848; 0,4689; Output: 0,4669; Desired output: 0,4699; Error: -0,003;
Input: 0,4699; 0,4692; 0,4713; 0,4658; 0,4393; 0,4943; 0,4456; Output: 0,4609; Desired output: 0,449; Error: 0,0119;
Input: 0,449; 0,4699; 0,4692; 0,4713; 0,4658; 0,4939; 0,4441; Output: 0,4403; Desired output: 0,4362; Error: 0,0041;
Input: 0,4362; 0,449; 0,4699; 0,4692; 0,4713; 0,5023; 0,461; Output: 0,446; Desired output: 0,4338; Error: 0,0121;
Input: 0,4338; 0,4362; 0,449; 0,4699; 0,4692; 0,4458; 0,4622; Output: 0,4419; Desired output: 0,4342; Error: 0,0077;
Input: 0,4342; 0,4338; 0,4362; 0,449; 0,4699; 0,4513; 0,5243; Output: 0,4692; Desired output: 0,4494; Error: 0,0198;
Input: 0,4494; 0,4342; 0,4338; 0,4362; 0,449; 0,6269; 0,7219; Output: 0,6567; Desired output: 0,6403; Error: 0,0164;
Input: 0,6403; 0,4494; 0,4342; 0,4338; 0,4362; 0,8255; 0,8647; Output: 0,818; Desired output: 0,8526; Error: -0,0345;
Input: 0,8526; 0,6403; 0,4494; 0,4342; 0,4338; 0,7684; 0,8078; Output: 0,7884; Desired output: 0,7919; Error: -0,0035;
Input: 0,7919; 0,8526; 0,6403; 0,4494; 0,4342; 0,6693; 0,7102; Output: 0,6779; Desired output: 0,6522; Error: 0,0257;
Input: 0,6522; 0,7919; 0,8526; 0,6403; 0,4494; 0,586; 0,609; Output: 0,5296; Desired output: 0,5479; Error: -0,0183;
Input: 0,5479; 0,6522; 0,7919; 0,8526; 0,6403; 0,4928; 0,5281; Output: 0,4628; Desired output: 0,5274; Error: -0,0646;
Total Mean Square Error: 8.451729932812585E-4

```

Figura 3.13. Ejemplo de prueba del modelo

Fuente: Elaboración propia

### 3.5 CASOS DE ESTUDIOS

Se crearon diversos casos de estudios para encontrar el comportamiento de diversas redes bajo diversas condiciones, este cambio involucra modificación en la dimensionalidad de los vectores de entrada y la cantidad de datos ingresados.

Para conocer las entradas de una RNA se debe conocer el número de variables rezagadas. Las entradas de una RNA pueden especificarse como variables rezagadas que tengan una auto correlación significativa [74]. La auto correlación permite medir la correlación entre los errores residuales en serie temporal en diferentes periodos. Para una serie temporal la correlación existe entre valores de la misma variable en diferentes tiempos  $Y_i$  y  $Y_{i+k}$ , cuando  $k = 1$  se dice que la auto correlación tiene un rezago de 1. Cuando existe una auto correlación, los valores sucesivos no son al azar. Cada valor de la serie para un tiempo  $t$  dependerá en cierta medida de los valores en un tiempo previo [75]. De igual manera estudios que emplean el modelaje de redes tipo perceptrón multicapa afirman que el análisis de la función de auto correlación parcial es suficiente para identificar los rezagos significantes de una serie temporal [76]. En la Tabla 3.14 y 3.15 se establecen los valores obtenidos luego de aplicar la función de auto correlación y auto correlación parcial respectivamente para los datos históricos de demanda. En ambos casos se aplica una diferenciación de primer orden con el propósito de eliminar efectos por tendencia y no estacionalidad. Los resultados muestran que un rezago de 1, 11 y 12 meses según la FAC y 2 meses según la FACP son los meses más significantes, seguidos de rezagos de 3, 4, 5, 7, 8, 9 y 10 meses. Es evidente que la demanda del año anterior para el mismo mes es la que tiene mayor importancia al momento de realizar la proyección. Estos rezagos serán los considerados para plantear los diferentes casos de estudio. No se consideran rezagos mayores a 12 meses porque al emplear estos datos la RNA vería reducida la cantidad de datos para el entrenamiento (por ejemplo si se considera un rezago de 36 meses se necesitarían los datos del 2006 para entrenar las salidas del 2009 pero no se disponen de ellos).

En las Figura 3.14 y 3.15 se muestran las representaciones graficas de las FAC y FACP para los datos.

Tabla 3.14. Resultados de la auto correlación  
Fuente: Elaboración propia

Auto correlación de los datos con una diferenciación regular y estacional de primer orden							
Rezago	Correlación	Rezago	Correlación	Rezago	Correlación	Rezago	Correlación
1	0,536348	11	-0,528073	21	-0,226784	31	0,135463
2	-0,034935	12	-0,956589	22	-0,023433	32	0,224478
3	-0,244418	13	-0,523276	23	0,518742	33	0,226020
4	-0,231191	14	0,032777	24	0,912912	34	0,021141
5	-0,134705	15	0,235651	25	0,509549	35	-0,511572
6	0,002162	16	0,226360	26	-0,029889	36	-0,894516
7	0,132788	17	0,134312	27	-0,228239	37	-0,497041
8	0,224317	18	-0,004568	28	-0,223289	38	0,031867
9	0,234713	19	-0,133914	29	-0,136822	39	0,227837
10	0,027956	20	-0,221974	30	0,003328	40	0,224216
Significante si $ Correlación  > 0,073973$							

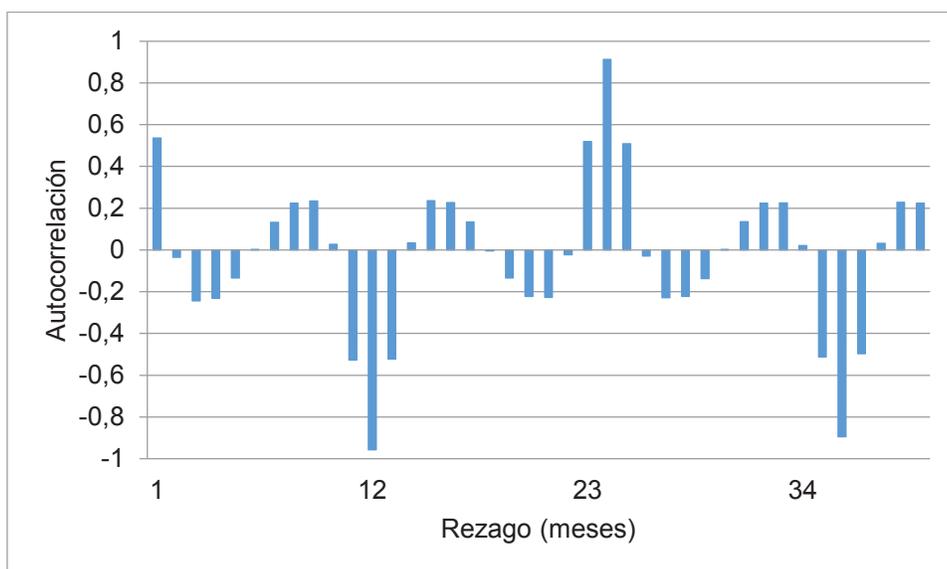


Figura 3.14. Representación gráfica de la auto correlación de los datos  
Fuente: Elaboración propia

Tabla 3.15. Resultados de las auto correlaciones parciales  
Fuente: Elaboración propia

Auto correlación parcial de los datos con una diferenciación regular y estacional de primer orden							
Rezago	Correlación	Rezago	Correlación	Rezago	Correlación	Rezago	Correlación
1	0,536348	11	-0,660507	21	-0,007936	31	-0,011741
2	-0,452884	12	-0,853589	22	-0,067553	32	0,025824
3	0,045403	13	0,191190	23	-0,009701	33	0,027513
4	-0,136464	14	-0,150882	24	-0,074202	34	-0,004126
5	-0,009734	15	-0,157223	25	0,007439	35	-0,036300
6	0,035709	16	-0,053953	26	-0,033646	36	-0,265666
7	0,078134	17	-0,017950	27	-0,035325	37	0,124865
8	0,127080	18	-0,035347	28	-0,032669	38	-0,057062
9	0,096787	19	-0,071648	29	-0,054603	39	-0,014553
10	-0,169278	20	-0,048698	30	-0,045605	40	-0,002306
Significante si $ Correlación  > 0,073973$							

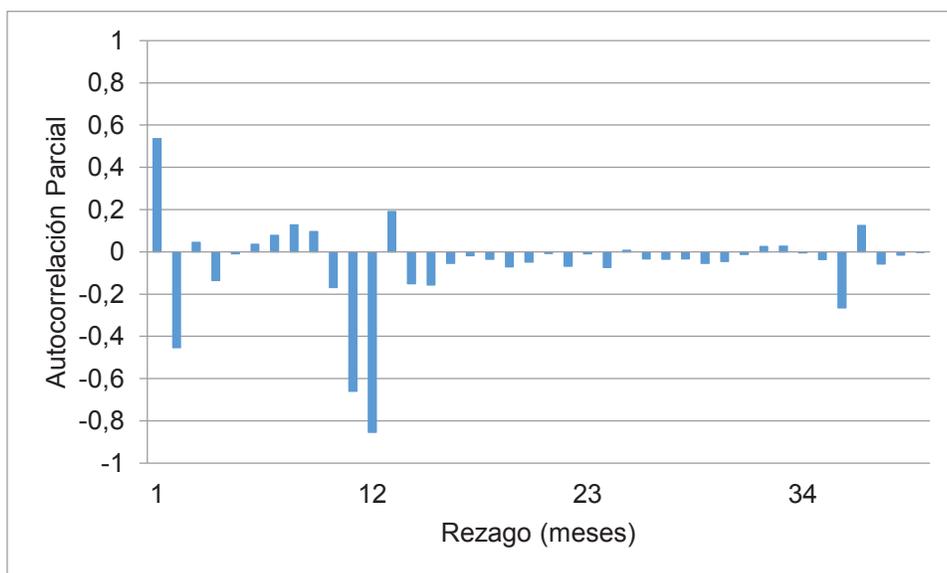


Figura 3.15. Representación gráfica de las auto correlaciones parciales de los datos  
Fuente: Elaboración propia

Para establecer los diversos casos de estudio se plantea la siguiente nomenclatura:

$E(m)$ : Energía normalizada del mes actual (previo a la predicción)

$E(m - i)$ : Energía normalizada considerando el rezago  $i$  con respecto al mes actual

$V_k$ : Vector de entrada  $k$

$V_s$ : Vector de salida  $s$

$P(a)$ : Población normalizada para el año  $a$

$VAB(a)$ : Valor agregado bruto para el año  $a$  (empleado en casos con salida anual)

$VAB(C_a)$ : Valor agregado bruto del cuatrimestre  $C$  del año  $a$  (empleado en casos con salida mensual)

### 3.4.1 CASO 1

Para este primer caso se consideran 4 entradas a la red, consistentes de la demanda para 1, 2, 11 y 12 meses anteriores al mes a pronosticar, estos rezagos fueron obtenidos de los valores más significativos de las auto correlaciones de la serie. El vector de entrada se define como:

$$V_1 = [E(m - 11) \ E(m - 10) \ E(m - 1) \ E(m)] \quad (3.5)$$

El vector de salida se define como:

$$V_{S1} = [E(m + 1)] \quad (3.6)$$

En este caso solo se considera los datos históricos como entradas. En la Tabla 3.16 se muestra un ejemplo del data set (con los valores normalizado) empleado para el entrenamiento de la red del caso 1.

Tabla 3.16. Ejemplo del data set para el caso 1  
Fuente: Elaboración propia

$E(m - 11)$	$E(m - 10)$	$E(m - 1)$	$E(m)$	$E(m + 1)$
0,367308	0,352402	0,321420	0,329127	0,370447
0,352402	0,358317	0,329127	0,370447	0,306892
0,358317	0,355155	0,370447	0,306892	0,352992
0,355155	0,352160	0,306892	0,352992	0,331282
0,352160	0,312290	0,352992	0,331282	0,320872
0,312290	0,312593	0,331282	0,320872	0,340161
0,312593	0,311852	0,320872	0,340161	0,343987
0,311852	0,328494	0,340161	0,343987	0,323267
0,328494	0,339953	0,343987	0,323267	0,337092
0,339953	0,321420	0,323267	0,337092	0,366511
0,321420	0,329127	0,337092	0,366511	0,369877
0,329127	0,370447	0,366511	0,369877	0,378901

### 3.4.2 CASO 2

En este caso se consideran los rezagos con auto correlación total o parcial que cumpla  $|Correlacion| > 0,073973$  como entradas a la red. El vector de entrada se define como:

$$V_2 = [ E(m - 11) E(m - 10) E(m - 9) E(m - 8) E(m - 7) E(m - 6) E(m - 4) E(m - 3) E(m - 2) E(m - 1) E(m) ] \quad (3.7)$$

El vector de salida se define como:

$$V_{S2} = [E(m + 1) ] \quad (3.8)$$

Al igual que el caso anterior solo se considera los datos históricos como entradas. En la Tabla 3.17 se muestra un ejemplo del data set (con los valores normalizado) empleado para el entrenamiento de la red del caso 2.

Tabla 3.17. Ejemplo del data set para el caso 2

Fuente: Elaboración propia

$E$ ( $m - 11$ )	$E$ ( $m - 10$ )	$E$ ( $m - 9$ )	$E$ ( $m - 8$ )	$E$ ( $m - 7$ )	$E$ ( $m - 6$ )	$E$ ( $m - 4$ )	$E$ ( $m - 3$ )	$E$ ( $m - 2$ )	$E$ ( $m - 1$ )	$E$ ( $m$ )	$E$ ( $m + 1$ )
0,36731	0,35240	0,35832	0,35515	0,35216	0,31229	0,31185	0,32849	0,33995	0,32142	0,32913	0,37045
0,35240	0,35832	0,35515	0,35216	0,31229	0,31259	0,32849	0,33995	0,32142	0,32913	0,37045	0,30689
0,35832	0,35515	0,35216	0,31229	0,31259	0,31185	0,33995	0,32142	0,32913	0,37045	0,30689	0,35299
0,35515	0,35216	0,31229	0,31259	0,31185	0,32849	0,32142	0,32913	0,37045	0,30689	0,35299	0,33128
0,35216	0,31229	0,31259	0,31185	0,32849	0,33995	0,32913	0,37045	0,30689	0,35299	0,33128	0,32087
0,31229	0,31259	0,31185	0,32849	0,33995	0,32142	0,37045	0,30689	0,35299	0,33128	0,32087	0,34016
0,31259	0,31185	0,32849	0,33995	0,32142	0,32913	0,30689	0,35299	0,33128	0,32087	0,34016	0,34399
0,31185	0,32849	0,33995	0,32142	0,32913	0,37045	0,35299	0,33128	0,32087	0,34016	0,34399	0,32327
0,32849	0,33995	0,32142	0,32913	0,37045	0,30689	0,33128	0,32087	0,34016	0,34399	0,32327	0,33709
0,33995	0,32142	0,32913	0,37045	0,30689	0,35299	0,32087	0,34016	0,34399	0,32327	0,33709	0,36651
0,32142	0,32913	0,37045	0,30689	0,35299	0,33128	0,34016	0,34399	0,32327	0,33709	0,36651	0,36988
0,32913	0,37045	0,30689	0,35299	0,33128	0,32087	0,34399	0,32327	0,33709	0,36651	0,36988	0,37890

### 3.4.3 CASO 3

En este caso se plantea la variable crecimiento demográfico como entrada a la red. El coeficiente de correlación es una medida de la asociación entre dos variables. Para verificar la relación existente entre la serie de datos de demanda y la población se realiza una prueba de correlación de Pearson. La prueba tiene como resultado 0,8925

lo que supone una alta correlación entre las variables. Considerando que el intervalo de correlación es -1 a 1, donde un valor cercano a 1 indica una fuerte relación positiva, un valor elevado de un variable indica un valor elevado en la otra variable [77]. La correlación de Pearson permite conocer si la asociación entre variables es creciente o decreciente. Cuando unas variables tienen una correlación cercana a 0 se dice que las variables no están correlacionadas [78]. En la Figura 3.15 se puede mostrar la comparación entre las variables. Estos resultados muestran la conveniencia de emplear el crecimiento demográfico como una variable de entrada a la red. El vector de entrada se define como:

$$[V_3 = P(a) E(m - 11) E(m - 10) E(m - 9) E(m - 8) E(m - 7) E(m - 6) E(m - 5) E(m - 4) E(m - 3) E(m - 2) E(m - 1) E(m)] \quad (3.9)$$

Donde:

$P(a)$ : Población para el año  $a$

El vector de salida se define como:

$$V_{S3} = [E(m + 1)] \quad (3.10)$$

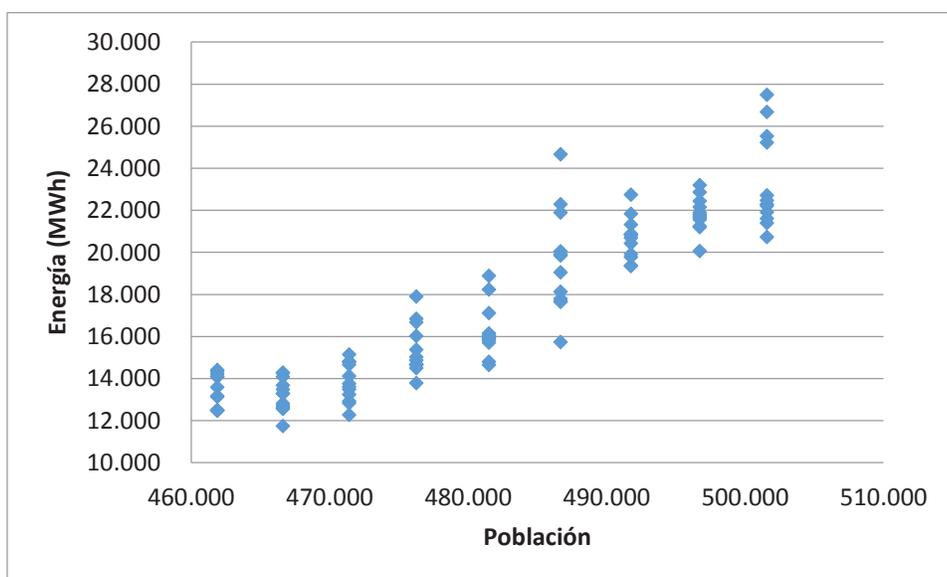


Figura 3.15. Correlación entre la población y la energía  
Fuente: Elaboración propia

### 3.4.3 CASO 4

En este caso se plantea la variable datos económicos (valor bruto agregado) como entrada a la red. Para verificar la relación existente entre la serie de datos de demanda y la población se realiza una prueba de correlación de Pearson. La prueba tiene como resultado 0,9359 lo que supone una alta correlación entre las variables. En la Figura 3.16 se puede mostrar la comparación entre las variables. El vector de entrada se define a continuación:

$$[V_4 = VAB(c) E(m - 11) E(m - 10) E(m - 9) E(m - 8) E(m - 7) E(m - 6) E(m - 5) E(m - 4) E(m - 3) E(m - 2) E(m - 1) E(m)] \quad (3.11)$$

El vector de salida se define como:

$$V_{S4} = [E(m + 1)] \quad (3.12)$$

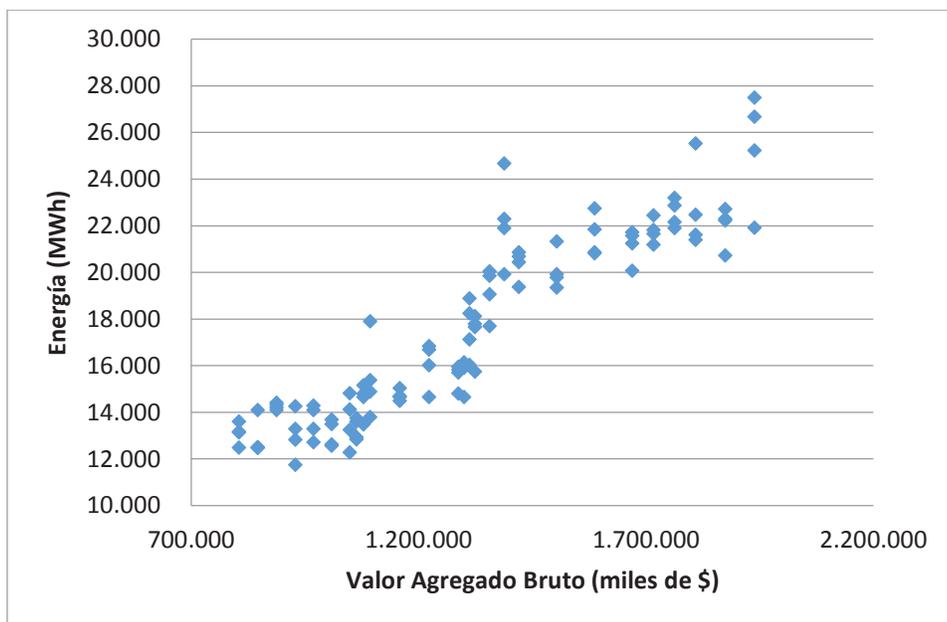


Figura 3.16. Correlación entre los datos económicos y la energía  
Fuente: Elaboración propia

### 3.4.5 CASO 5

Caso similar al primero, se diferencia porque se plantean 12 salidas. El vector de entrada se define como:

$$V_5 = [E(m - 11) E(m - 10) E(m - 1) E(m)] \quad (3.13)$$

El vector de salida se define como:

$$V_{S5} = [E(m + 1) E(m + 2) E(m + 3) E(m + 4) E(m + 5) E(m + 6) E(m + 7) E(m + 8) E(m + 9) E(m + 10) E(m + 11) E(m + 12)] \quad (3.14)$$

En las Tablas 3.11 y 3.12 se muestran las variables normalizadas de entrada y de salida. En este caso la RNA tendrá simultáneamente la predicción de los 12 meses del año 2017.

### 3.4.6 CASO 6

Este caso se plantea similar al 2, se consideran los rezagos con auto correlación total o parcial que cumpla  $|Correlación| > 0,073973$  como entradas a la red. Se emplean 12 salidas. El vector de entrada se define como:

$$V_6 = [E(m - 11) E(m - 10) E(m - 9) E(m - 8) E(m - 7) E(m - 6) E(m - 4) E(m - 3) E(m - 2) E(m - 1) E(m)] \quad (3.15)$$

El vector de salida se define como:

$$V_{S6} = [E(m + 1) E(m + 2) E(m + 3) E(m + 4) E(m + 5) E(m + 6) E(m + 7) E(m + 8) E(m + 9) E(m + 10) E(m + 11) E(m + 12)] \quad (3.16)$$

### 3.4.7 CASO 7

Similar al caso 3. La única modificación es que se presentan las 12 proyecciones simultáneamente. En vector de entrada permanece igual.

$$[V_7 = P(\alpha) E(m - 11) E(m - 10) E(m - 9) E(m - 8) E(m - 7) E(m - 6) E(m - 4) E(m - 3) E(m - 2) E(m - 1) E(m)] \quad (3.17)$$

Donde:

$P(a)$ : Población para el año  $a$

El vector de salida se define como:

$$V_{S7} = [E(m+1) E(m+2) E(m+3) E(m+4) E(m+5) E(m+6) E(m+7) E(m+8) E(m+9) E(m+10) E(m+11) E(m+12)] \quad (3.18)$$

### 3.4.8 CASO 8

El vector de entrada se define como:

$$[V_8 = VAB(c) E(m-11) E(m-10) E(m-9) E(m-8) E(m-7) E(m-6) E(m-4) E(m-3) E(m-2) E(m-1) E(m)] \quad (3.19)$$

El vector de salida se define como:

$$V_{S8} = [E(m+1) E(m+2) E(m+3) E(m+4) E(m+5) E(m+6) E(m+7) E(m+8) E(m+9) E(m+10) E(m+11) E(m+12)] \quad (3.20)$$

### 3.4.9 CASO 9

El vector de entrada se define como:

$$[V_9 = P(a) VAB(c) E(m-11) E(m-10) E(m-9) E(m-8) E(m-7) E(m-6) E(m-4) E(m-3) E(m-2) E(m-1) E(m)] \quad (3.21)$$

El vector de salida se define como:

$$V_{S9} = [E(m+1) E(m+2) E(m+3) E(m+4) E(m+5) E(m+6) E(m+7) E(m+8) E(m+9) E(m+10) E(m+11) E(m+12)] \quad (3.22)$$

# CAPÍTULO 4

## CAPÍTULO IV

### RESULTADOS

En este capítulo se muestran los resultados de las validaciones de los modelos. Luego del entrenamiento con datos desde el año 2008 hasta el 2015 se realiza una proyección para el 2016 y se comparan con los valores reales. Se emplea el MAPE para identificar el error de cada modelo y compararlo con la proyección de la Empresa Eléctrica Riobamba S.A.

#### 4.1 CASO 1

En la Tabla 4.1 se muestran los detalles de errores residuales y MAPE, para este caso la proyección de la EERSA fue mejor que la red neuronal artificial.

Tabla 4.1 Resultados del caso 1 en MWh  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual RNA	Proyección EERSA ( $\hat{Y}_i$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
Enero	25522	22520	3002	25553	-31	0,1176	0,0012
Febrero	22467	22800	-333	22588	-121	0,0148	0,0054
Marzo	21614	22804	-1190	22508	-894	0,0551	0,0414
Abril	21396	23040	-1644	21802	-406	0,0768	0,0190
Mayo	22220	23040	-820	22779	-559	0,0369	0,0251
Junio	20732	23040	-2308	22783	-2051	0,1113	0,0990
Julio	22296	23120	-824	23091	-795	0,0370	0,0357
Agosto	22716	23200	-484	22933	-217	0,0213	0,0096
Septiembre	21913	23400	-1487	24268	-2355	0,0679	0,1075
Octubre	25231	24440	791	23679	1552	0,0314	0,0615
Noviembre	26674	23760	2914	24462	2212	0,1092	0,0829
Diciembre	27498	23920	3578	24206	3292	0,1301	0,1197
					MAPE(%)	6,7450	5,0666

En la Figura 4.1 se representan los resultados de manera gráfica donde se muestra una debilidad de parte de la RNA en adaptarse a la no linealidad de los datos, es descartado el caso 1 como una red válida para pronosticar los datos.

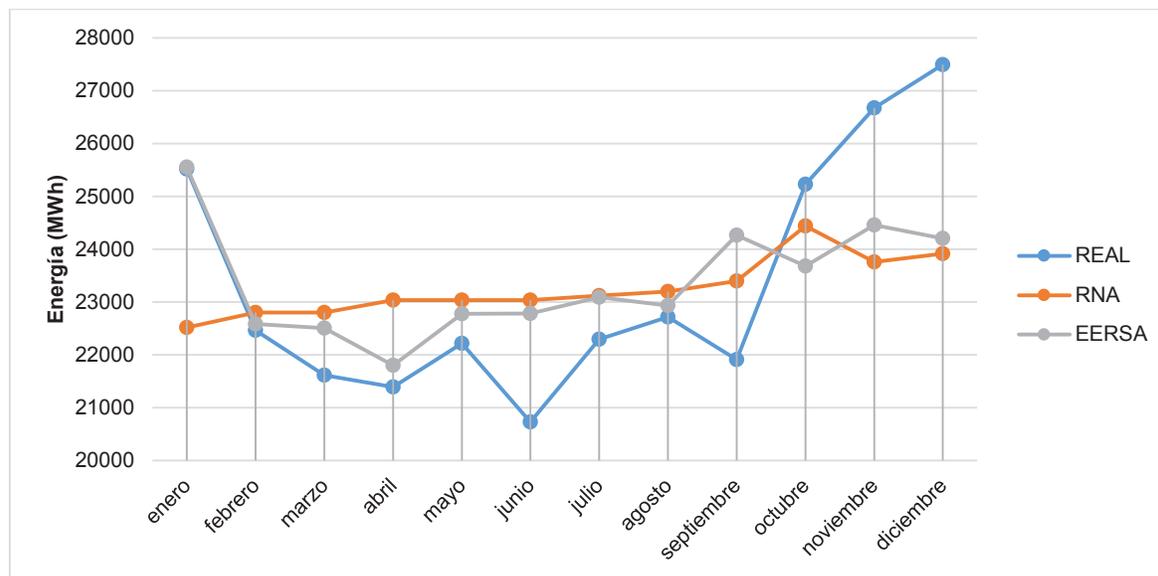


Figura 4.1. Gráficas de los resultados del caso 1  
Fuente: Elaboración propia

## 4.2 CASO 2

Los resultados de este caso fueron similares al caso 1, no se presentan grandes variaciones, los resultados presentan un comportamiento bastante lineal, particularmente las deficiencias son más profundas en los meses de enero, noviembre y diciembre debido al aumento drástico en el consumo para meses. El MAPE se encuentra por encima del valor de las proyecciones de la EERSA por lo que resulta evidente la debilidad de esta configuración particular de RNA para el pronóstico de datos. En la Tabla 4.2 se detallan los resultados y en la Figura 4.2 la comparación gráfica. Se evidencia que emplear un vector de entrada con mayores rezagos no es sinónimo de mejoría en los resultados.

Tabla 4.2. Resultados del caso 2 en MWh  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual RNA	Proyección EERSA ( $\hat{Y}_i$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
Enero	25522	23080	2442	25553	-31	0,0957	0,0012
Febrero	22467	23200	-733	22588	-121	0,0326	0,0054
Marzo	21614	23400	-1786	22508	-894	0,0826	0,0414
Abril	21396	23440	-2044	21802	-406	0,0955	0,0190
Mayo	22220	23800	-1580	22779	-559	0,0711	0,0251
Junio	20732	23760	-3028	22783	-2051	0,1461	0,0990
Julio	22296	24000	-1704	23091	-795	0,0764	0,0357
Agosto	22716	24120	-1404	22933	-217	0,0618	0,0096
Septiembre	21913	24280	-2367	24268	-2355	0,1080	0,1075
Octubre	25231	24520	711	23679	1552	0,0282	0,0615
Noviembre	26674	24520	2154	24462	2212	0,0808	0,0829
Diciembre	27498	24760	2738	24206	3292	0,0996	0,1197
					MAPE(%)	8,1533	5,0666

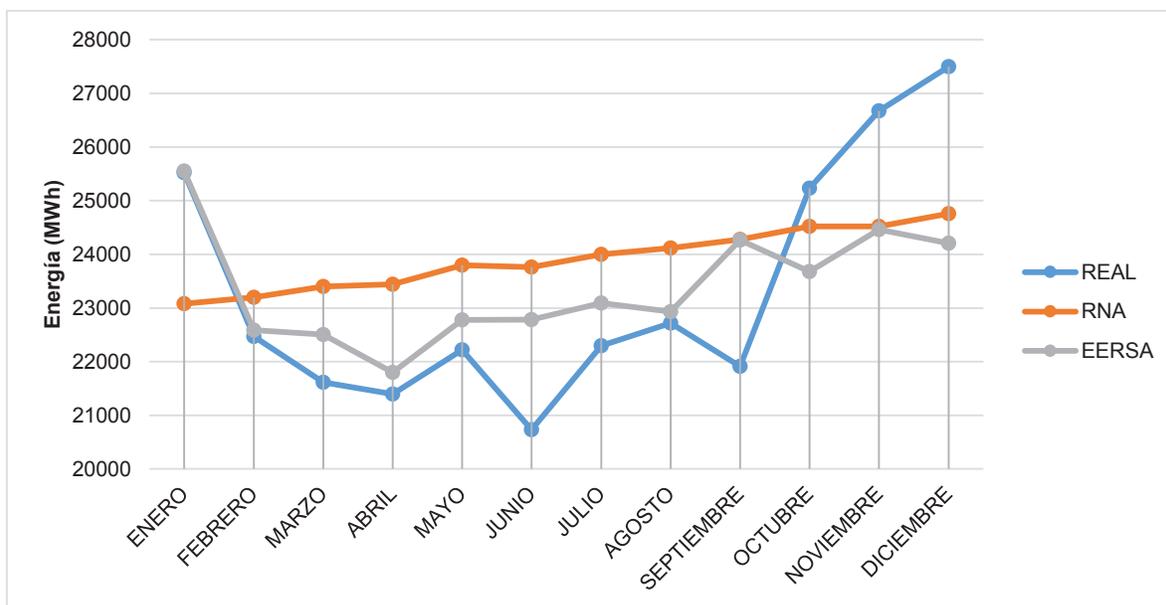


Figura 4.2. Gráficas de los resultados del caso 2  
Fuente: Elaboración propia

### 4.3 CASO 3

Los resultados son muy similares al caso anterior, mantiene un comportamiento lineal, no se presentan grandes diferencias salvo por la disminución del MAPE. No obstante, el error continúa siendo mayor que el obtenido de las proyecciones de la EERSA, en la Tabla 4.3 se muestra el detalle de los resultados mientras que en la Figura 4.3 se comparan gráficamente.

Tabla 4.3. Resultados del caso 3 en MWh  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual RNA	Proyección EERSA ( $\hat{Y}_j$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
enero	25522	22520	3002	25553	-31	0,1176	0,0012
febrero	22467	22400	67	22588	-121	0,0030	0,0054
marzo	21614	22520	-906	22508	-894	0,0419	0,0414
abril	21396	22480	-1084	21802	-406	0,0507	0,0190
mayo	22220	22640	-420	22779	-559	0,0189	0,0251
junio	20732	22560	-1828	22783	-2051	0,0882	0,0990
julio	22296	22912	-616	23091	-795	0,0276	0,0357
agosto	22716	22988	-272	22933	-217	0,0120	0,0096
septiembre	21913	22936	-1023	24268	-2355	0,0467	0,1075
octubre	25231	23036	2195	23679	1552	0,0870	0,0615
noviembre	26674	22760	3914	24462	2212	0,1467	0,0829
diciembre	27498	22880	4618	24206	3292	0,1679	0,1197
					MAPE(%)	6,7350	5,0666

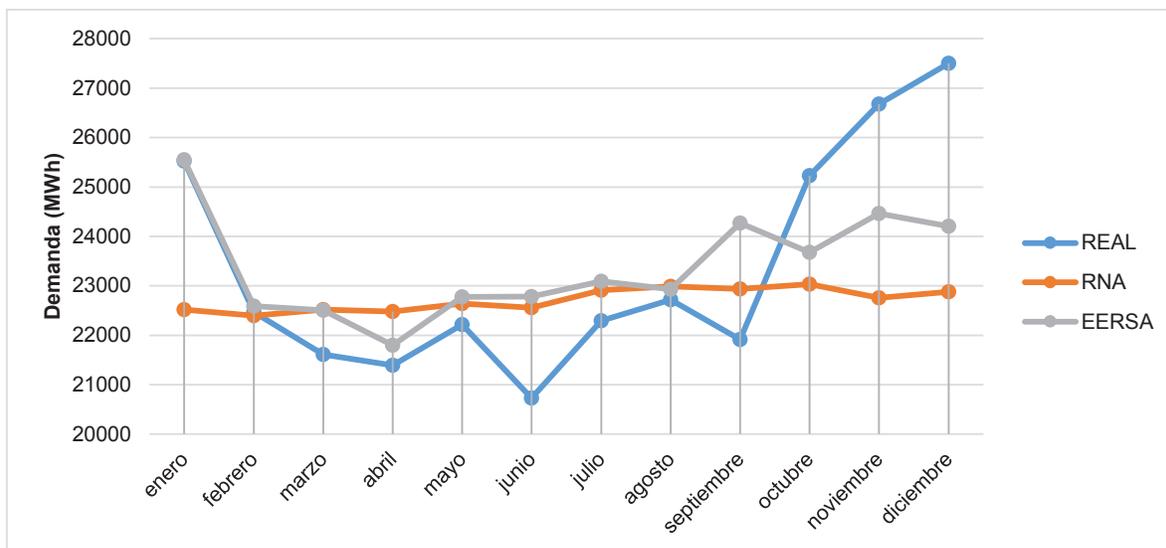


Figura 4.3. Gráficas de los resultados del caso 3  
Fuente: Elaboración propia

#### 4.4 CASO 4

Los resultados son similares a los presentados en los casos anteriores con un comportamiento muy lineal de las proyecciones. Se puede mencionar que una red iterativa como las planteadas en los casos 1-4 no logra adaptarse de manera correcta a los datos involucrados, no pude afirmarse que la red no es adecuada para la tarea de emitir pronósticos, ya que el desempeño de las RNA está muy ligado al tipo de datos y la configuración de la misma. En las Tablas 4.4 y Figura 4.4 se detallan los resultados.

Tabla 4.4. Resultados del caso 4 en MWh  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual al RNA	Proyección EERSA ( $\hat{Y}_j$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
enero	25522	22700	2822	25553	-31	0,1106	0,0012
febrero	22467	22944	-477	22588	-121	0,0212	0,0054
marzo	21614	22836	-1222	22508	-894	0,0565	0,0414
abril	21396	22936	-1540	21802	-406	0,0720	0,0190
mayo	22220	23140	-920	22779	-559	0,0414	0,0251
junio	20732	23196	-2464	22783	-2051	0,1189	0,0990
julio	22296	23192	-896	23091	-795	0,0402	0,0357
agosto	22716	23304	-588	22933	-217	0,0259	0,0096
septiembre	21913	23572	-1659	24268	-2355	0,0757	0,1075
octubre	25231	23664	1567	23679	1552	0,0621	0,0615
noviembre	26674	23704	2970	24462	2212	0,1113	0,0829
diciembre	27498	23724	3774	24206	3292	0,1372	0,1197
					MAPE(%)	7,2750	5,0666

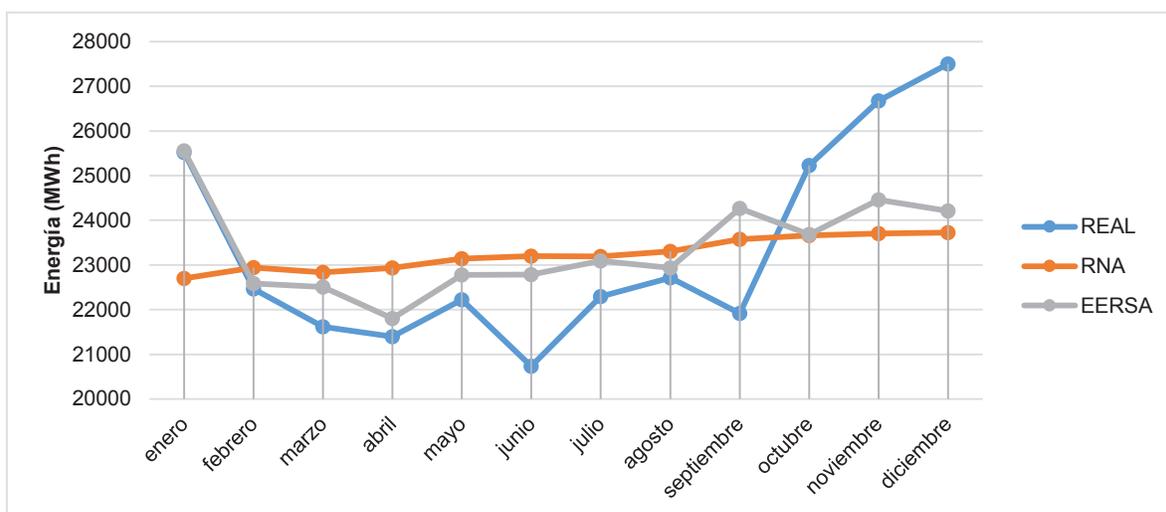


Figura 4.4. Gráficas de los resultados del caso 4  
Fuente: Elaboración propia

#### 4.5 CASO 5

Aunque el MAPE de la proyección mediante RNA es mayor al de la EERSA se nota una clara mejoría en el comportamiento de los datos. Particularmente en el intervalo marzo-septiembre los datos generados por la RNA son prácticamente iguales a los reales. La deficiencia radica en que los meses de enero, octubre, noviembre y diciembre presentan valores medianamente atípicos, al incrementarse por encima del promedio. Cabe destacar que el error es muy sensible al tiempo de entrenamiento, cumpliéndose una de las premisas planteadas en el Capítulo 2, la configuración y el tiempo de entrenamiento son entonces factores decisivos para obtener un modelo adecuado de red neuronal artificial. Por lo anterior planteado no se puede eliminar la posibilidad de disminuir el MAPE en cierto grado. En la Tabla 4.5 y Figura 4.5 se detallan los resultados.

Tabla 4.5. Resultados del caso 5 en MWh  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual al RNA	Proyección EERSA ( $\hat{Y}_j$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
enero	25522	21628	3894	25553	-31	0,1526	0,0012
febrero	22467	20652	1815	22588	-121	0,0808	0,0054
marzo	21614	21312	302	22508	-894	0,0140	0,0414
abril	21396	20900	496	21802	-406	0,0232	0,0190
mayo	22220	22088	132	22779	-559	0,0059	0,0251
junio	20732	20856	-124	22783	-2051	0,0060	0,0990
julio	22296	22316	-20	23091	-795	0,0009	0,0357
agosto	22716	21708	1008	22933	-217	0,0444	0,0096
septiembre	21913	23444	-1531	24268	-2355	0,0699	0,1075
octubre	25231	22732	2499	23679	1552	0,0990	0,0615
noviembre	26674	23216	3458	24462	2212	0,1296	0,0829
diciembre	27498	24196	3302	24206	3292	0,1201	0,1197
					MAPE (%)	6,2200	5,0666

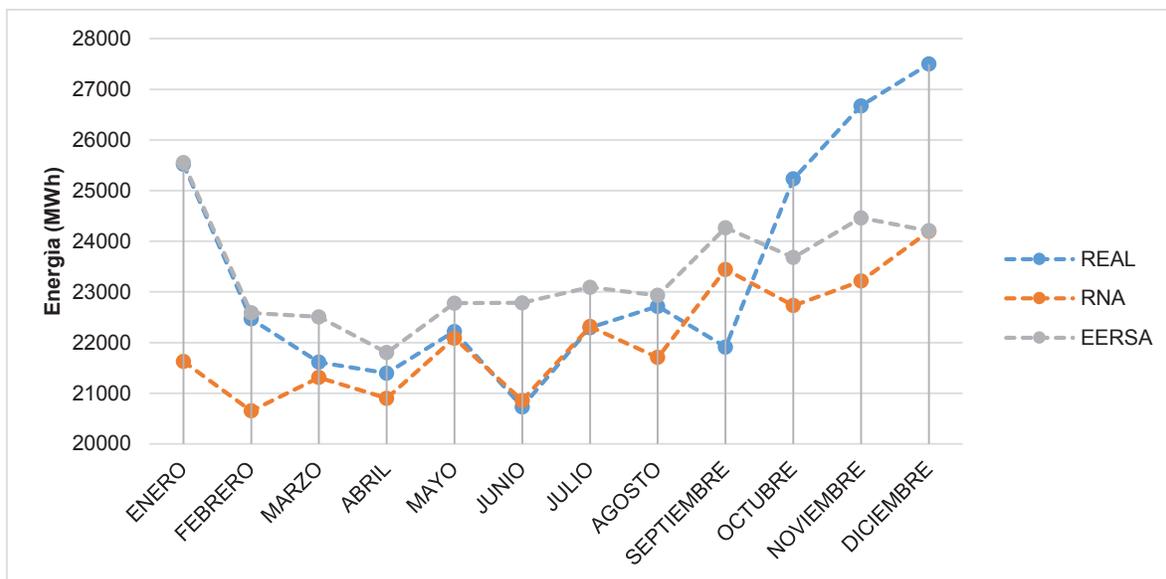


Figura 4.5. Gráficas de los resultados del caso 5  
Fuente: Elaboración propia

#### 4.6 CASO 6

El caso 6 presenta marginalmente un MAPE menor a la proyección de la EERSA. Aunque no es clara la relación entre la dimensión del vector de entrada y el mejoramiento del pronóstico. Mantiene un mejor pronóstico de marzo a agosto con respecto al de la EERSA. Resulta evidente que la red neuronal es muy exacta cuando los datos se encuentran en cierto margen respecto a los promedios históricos pero presenta deficiencia cuando los datos se alejan excesivamente de los promedios. En la Tabla 4.6 y Figura 4.6 se detallan los resultados.

Tabla 4.6. Resultados del caso 6 en MWh  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual RNA	Proyección EERSA ( $\hat{Y}_i$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
enero	25522	22784	2738	25553	-31	0,1073	0,0012
febrero	22467	21512	955	22588	-121	0,0425	0,0054
marzo	21614	21712	-98	22508	-894	0,0045	0,0414
abril	21396	21612	-216	21802	-406	0,0101	0,0190
mayo	22220	22568	-348	22779	-559	0,0157	0,0251
junio	20732	21472	-740	22783	-2051	0,0357	0,0990
julio	22296	22836	-540	23091	-795	0,0242	0,0357
agosto	22716	22420	296	22933	-217	0,0130	0,0096
septiembre	21913	23908	-1995	24268	-2355	0,0910	0,1075
octubre	25231	23536	1695	23679	1552	0,0672	0,0615
noviembre	26674	24032	2642	24462	2212	0,0990	0,0829
diciembre	27498	24836	2662	24206	3292	0,0968	0,1197
					MAPE(%)	5,0583	5,0666

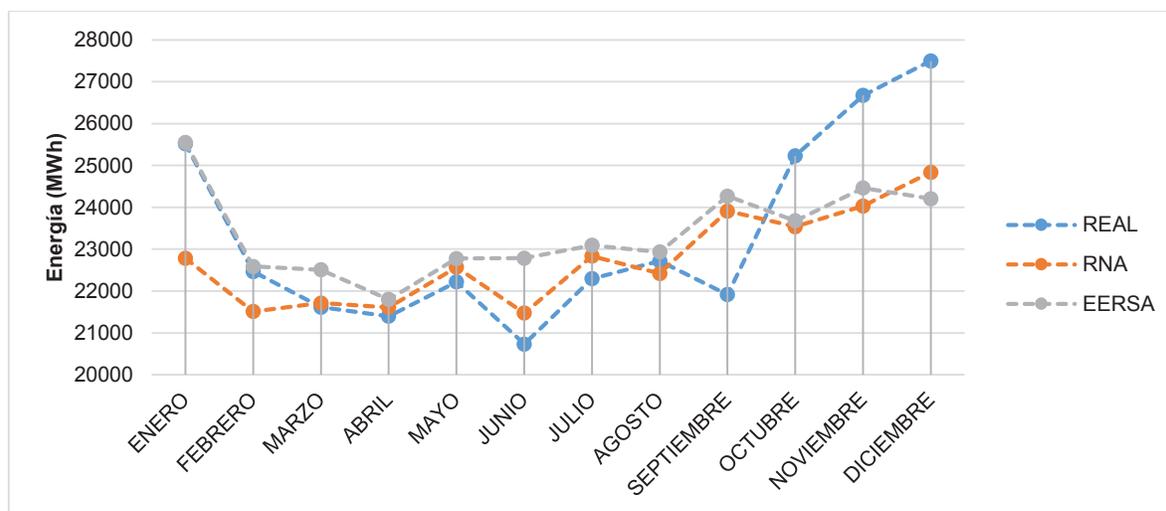


Figura 4.6. Gráficas de los resultados del caso 6  
Fuente: Elaboración propia

#### 4.7 CASO 7

Los resultados tienen una ligera mejoría con respecto al caso 6 manteniendo un MAPE por debajo de las proyecciones de la EERSA, considerando además los resultados del caso 6 se puede afirmar que los modelos que involucran las 12 proyecciones simultáneamente muestran mejores resultados que los modelos donde se aplicó un método iterativo. En la Tabla 4.7 y Figura 4.7 se detallan los resultados.

Tabla 4.7. Resultados del caso 7  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual RNA	Proyección EERSA ( $\hat{Y}_j$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
enero	25522	22796	2726	25553	-31	0,10681	0,00122
febrero	22467	21200	1267	22588	-121	0,05639	0,00540
marzo	21614	21508	106	22508	-894	0,00490	0,04136
abril	21396	21464	-68	21802	-406	0,00318	0,01897
mayo	22220	22504	-284	22779	-559	0,01278	0,02515
junio	20732	21292	-560	22783	-2051	0,02701	0,09895
julio	22296	22748	-452	23091	-795	0,02027	0,03568
agosto	22716	22300	416	22933	-217	0,01831	0,00955
septiembre	21913	23512	-1599	24268	-2355	0,07297	0,10748
octubre	25231	23504	1727	23679	1552	0,06845	0,06150
noviembre	26674	24020	2654	24462	2212	0,09950	0,08293
diciembre	27498	25164	2334	24206	3292	0,08488	0,11971
					MAPE(%)	4,79541	5,0666

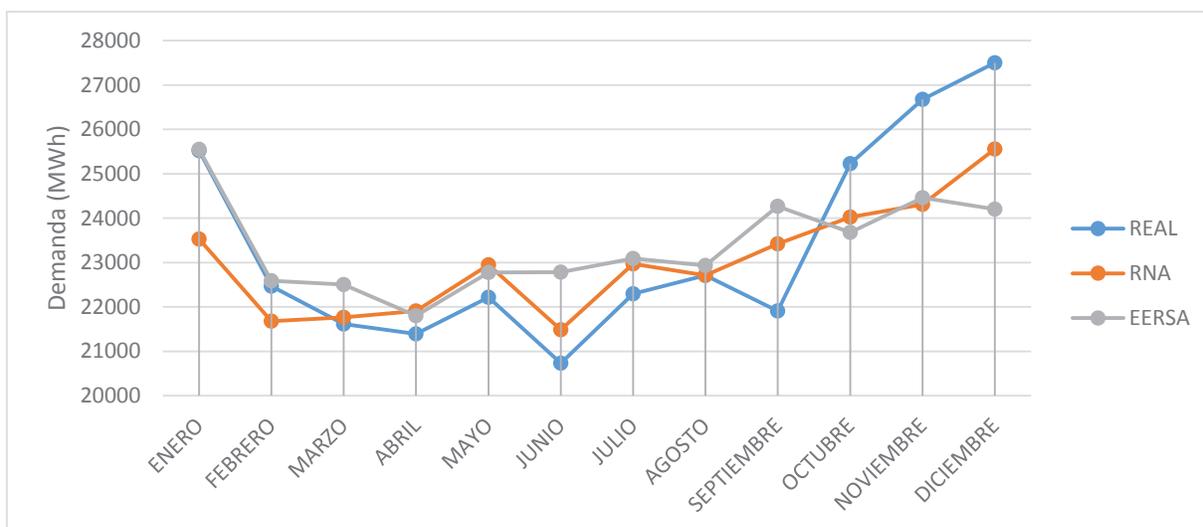


Figura 4.7. Gráficas de los resultados del caso 7

Fuente: Elaboración propia

#### 4.8 CASO 8

El caso 8 presenta una ligera mejoría con respecto a las proyecciones de la EERSA, pero se mantiene la debilidad con datos por encima de los promedios. En la Tabla 4.8 y Figura 4.8 se detallan los resultados.

Tabla 4.8. Resultados del caso 8 en MWh

Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual al RNA	Proyección EERSA ( $\hat{Y}_j$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
enero	25522	23344	2178	25553	-31	0,0853	0,0012
febrero	22467	21708	759	22588	-121	0,0338	0,0054
marzo	21614	22072	-458	22508	-894	0,0212	0,0414
abril	21396	21780	-384	21802	-406	0,0179	0,0190
mayo	22220	23028	-808	22779	-559	0,0364	0,0251
junio	20732	21456	-724	22783	-2051	0,0349	0,0990
julio	22296	22996	-700	23091	-795	0,0314	0,0357
agosto	22716	22532	184	22933	-217	0,0081	0,0096
septiembre	21913	24156	-2243	24268	-2355	0,1024	0,1075
octubre	25231	23972	1259	23679	1552	0,0499	0,0615
noviembre	26674	24428	2246	24462	2212	0,0842	0,0829
diciembre	27498	25372	2126	24206	3292	0,0773	0,1197
					MAPE(%)	4,8566	5,0666

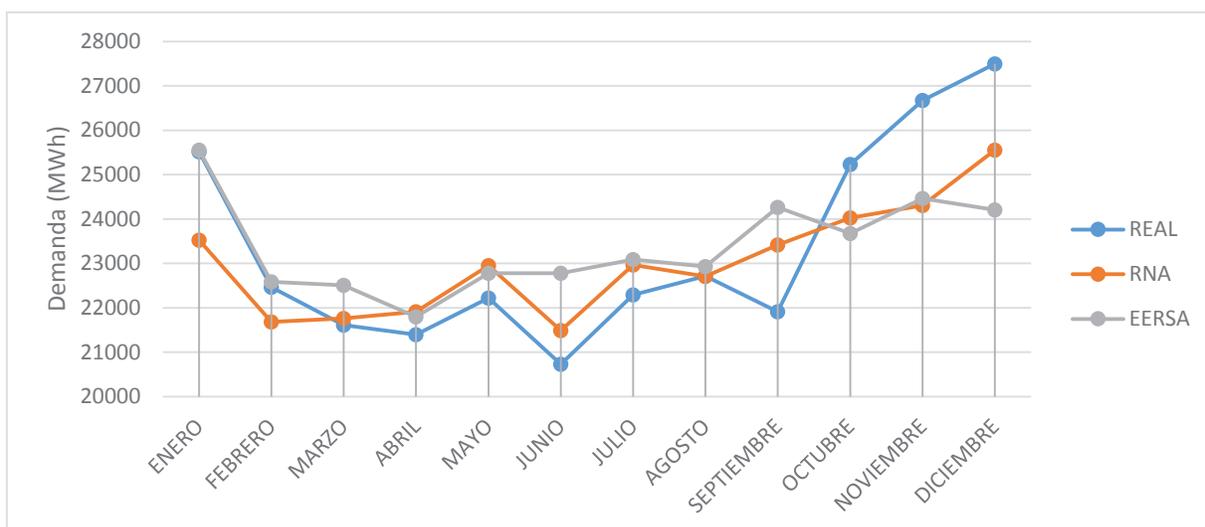


Figura 4.8. Gráficas de los resultados del caso 8  
Fuente: Elaboración propia

En la Tabla 4.9 se muestra como el incremento del 2015 al 2016 fue el mayor en el intervalo de datos analizados, como se ha mencionado anteriormente, es muy factible que este hecho haya causado un pronóstico por debajo de los datos de consumo real.

Tabla 4.9. Incremento anual porcentual del consumo  
Fuente: Elaboración propia

	2008-2009	2009-2010	2010-2011	2011-2012	2012-2013	2013-2014	2014-2015	2015-2016
enero	97%	116%	104%	103%	114%	115%	104%	117%
noviembre	90%	117%	114%	108%	120%	100%	100%	122%
diciembre	95%	111%	110%	113%	118%	102%	102%	119%

## 4.9 CASO 9

El caso 9 presenta el menor MAPE de todos los casos estudiados con respecto a las proyecciones de la EERSA por lo que se puede afirmar que un uso correcto de datos y de configuración de RNA puede mejorar las proyecciones por los métodos empleados actualmente. No se descarta que mediante una profundización de este caso mediante otros tipos de redes y/o algoritmos de entrenamiento puedan mejorarse los resultados hasta disminuir significativamente el MAPE de la proyección. En la Tabla 4.10 y Figura 4.9 se detallan los resultados.

Tabla 4.10. Resultados del caso 9 en MWh  
Fuente: Elaboración propia

Mes	Energía real ( $Y_i$ )	Proyección RNA ( $\hat{Y}_i$ )	Error residual al RNA	Proyección EERSA ( $\hat{Y}_j$ )	Error residual EERSA	$\frac{ Y_i - \hat{Y}_i }{Y_i}$	$\frac{ Y_i - \hat{Y}_j }{Y_i}$
enero	25522	23356	2166	25553	-31	0,0849	0,0012
febrero	22467	21476	991	22588	-121	0,0441	0,0054
marzo	21614	21612	2	22508	-894	0,0001	0,0414
abril	21396	21800	-404	21802	-406	0,0189	0,0190
mayo	22220	22840	-620	22779	-559	0,0279	0,0251
junio	20732	21320	-588	22783	-2051	0,0284	0,0990
julio	22296	23000	-704	23091	-795	0,0316	0,0357
agosto	22716	22860	-144	22933	-217	0,0063	0,0096
septiembre	21913	24352	-2439	24268	-2355	0,1113	0,1075
octubre	25231	23864	1367	23679	1552	0,0542	0,0615
noviembre	26674	24548	2126	24462	2212	0,0797	0,0829
diciembre	27498	25564	1934	24206	3292	0,0703	0,1197
					MAPE(%)	4,6475	5,0666

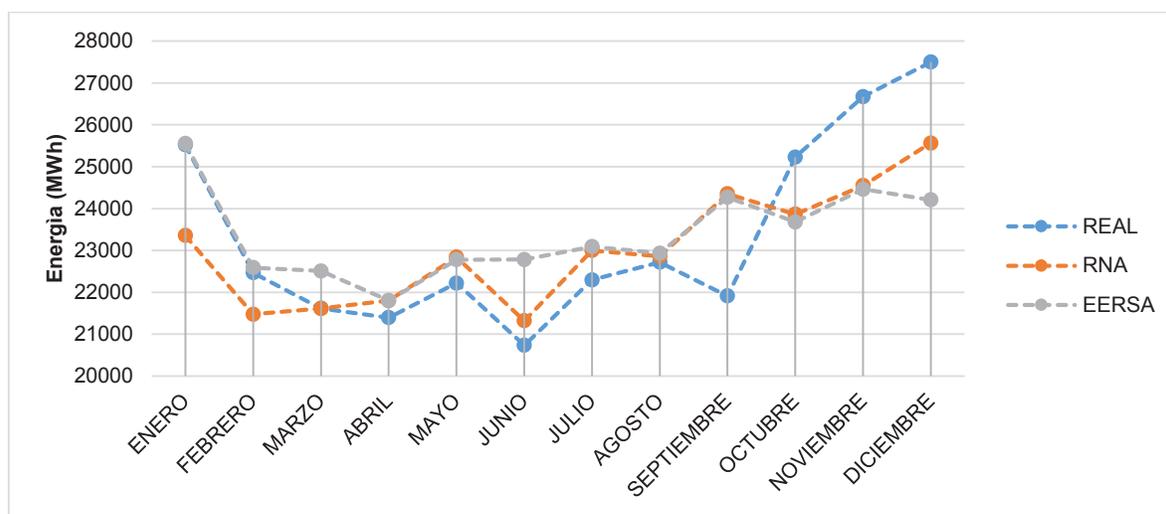


Figura 4.9. Gráficas de los resultados del caso 9  
Fuente: Elaboración propia

# CAPÍTULO 5

## CAPÍTULO V

### CONCLUSIONES Y RECOMENDACIONES

#### 5.1 CONCLUSIONES

- El caso de estudio 9 presenta los mejores resultados con un error global por debajo del pronóstico de la EERSA, lo que verifica que es posible obtener una proyección adecuada con una red neuronal artificial.
- Los datos históricos de energía presentan una alta auto correlación así como una fuerte correlación con los datos de crecimiento demográfico y los datos económicos (PIB), siendo estos últimos los más correlacionados. Sin embargo, no son concluyentes los resultados en cuanto a la mejora que aportan estos datos a los modelos.
- Un vector con mayor dimensión de entrada no necesariamente mejora el pronóstico, si bien en los casos de estudios se encontró una mejoría parcial pero no puede concluirse que exista una correlación fuerte entre ambos hechos. La dificultad de tratar con tantas variables puede no ser necesaria en vista de la mejoría que se obtiene.
- Los modelos de predicción son únicos, independientemente del método empleado, por tanto las redes neuronales del presente trabajo la energía adecuada para la Empresa Eléctrica Riobamba S.A. de necesitarse la predicción para otro caso de estudio relacionado debe obtenerse un *data set* acorde y entrenar una red conforme al caso de estudio.
- Los modelos de predicción permiten lograr una planificación eficiente y con un significativo ahorro en el presupuesto anual predestinado para la compra de

energía eléctrica del MEM mismo que puede ser direccionado en otras inversiones que beneficien a la Empresa Eléctrica Riobamba S.A.

- Es importante limpiar los datos históricos de esta forma evitar acumulación de datos atípicos que puedan impactar en la precisión de los modelos de predicción.
- La exactitud y precisión del modelo de red neuronal artificial está muy ligado al tiempo de entrenamiento y configuración de la red, actualmente no hay algoritmos exactos para encontrar el mejor tiempo de entrenamiento y configuración, lo que trae como consecuencia se deba iterar para conseguir el menor error posible. No siempre el error es menor a mayor tiempo de entrenamiento. Con el acceso a una supercomputadora que permita crear miles de modelos es indudable que los resultados serían óptimos y se podría lograr una disminución considerable en el error con respecto a los métodos tradicionales de predicción.

## **5.2 RECOMENDACIONES**

- Se recomienda aplicar otros tipos de redes neuronales artificiales con otros tipos de entrenamientos con el propósito de encontrar el mejor modelo posible para la predicción.
- Aunque el Neuroph Studio es capaz de modelar de manera adecuada redes neuronales de tipo perceptron multicapa se recomienda modelar redes no lineales autoregresivas, para ello es posible el uso del MATLAB y su función narnet, en este ambiente es posible definir la topología de la red, así como las variables de entrada y de salida, las redes no lineales auto regresivas se perfilan con un gran potencial para tareas de predicción debido a su complejidad.

- Se recomienda otro programa computacional (framework y/o librerías) como el Theano, que es una librería muy flexible empleada con Python, capaz de usar el CPU y el GPU en caso de disponibilidad para potenciar los resultados y la rapidez. De igual manera, se puede emplear TensorFlow que posee librerías de redes neuronales. También está disponible MXNet, que usa una forma modular sencilla para crear redes neuronales y entrenarlas. Suele ubicarse dentro de las librerías más rápidas, aunque presenta falta de flexibilidad.
- Se recomienda que con los resultados encontrados de predicción utilizando Redes Neuronales Artificiales se realice un análisis del método propuesto con el utilizado por la Dirección de Planificación de la EERSA.

## REFERENCIAS BIBLIOGRÁFICAS

- [1] D. Kriesel, *A Brief Introduction to Neural Networks*, 2005.
- [2] R. Rojas, *Neural Networks A Systematic Introduction*, 1996.
- [3] D. J. Beale y H. Demuth, segunda ed., pp. 23-27.
- [4] G. Zhang, B. E. Patuwo y M. Y. Hu, «Forecasting with artificial neural networks: The state of the art,» *International Journal of Forecasting*, nº 14, pp. 35-62, 1998.
- [5] R. J. Sifontes, *ESTIMACIÓN DE LA DEMANDA A MEDIANO PLAZO DE LA SUBESTACIÓN LAGUNITA UTILIZANDO REDES NEURONALES*, Sartanejas, 2010.
- [6] R. Hernández Sampieria, C. Fernández Collado y P. Baptista Lucio, de *Metodología de la investigación*, cuarta ed., México D.F., 2006, pp. 64-67.
- [7] T. A. Short, «Electric Power Distribution Handbook,» segunda ed., CRC Press, 2014, pp. 26-31.
- [8] T. Gonen, «Electric Power Distribution System Engineering,» segunda ed., CRC Press, 2008, pp. 1-57.
- [9] S. E. Castaño, «Redes de Distribución de Energía,» tercera ed., Manizales, Universidad Nacional de Colombia, 2004, pp. 17-32.
- [10] W. H. Kersting, *Distribution System Modeling and Analysis*, tercera ed., Las Cruces, Nuevo Mexico: CRC Press, 2012.
- [11] J. D. Juárez Cervantes, de *Sistemas de distribución de energía eléctrica*, México, Universidad Autónoma Metropolitana, 1995, pp. 39-41.
- [12] J. D. R. Llanos Proaño, *Método para la generación de perfiles de demanda en comunidades aisladas y predicción de demanda de corto plazo, para micro-redes basadas en energías renovables*, Santiago de Chile, 2012.
- [13] Agencia de Regulación y Control de Electricidad - ARCONEL, *Pliego tarifario para las empresas eléctricas*, 2016.
- [14] W. H. Kersting, *Distribution System Modeling and Analysis*, New Mwxico: State University, 2012.
- [15] N. E. Vásquez, M. Sánchez y E. Henao, *Estudio de capacidad instalada*, Antioquia, 2014.
- [16] C. V. Quintero García, *Estimación de la demanda para la localización de pérdidas no-técnicas por punto de transformación*, 2012.
- [17] Consejo Nacional de Electricidad - CONELEC, *Plan maestro de electrificación 2013 - 2022*, 2013.
- [18] J. Ringwood, *Intelligent Forecasting of Electricity Demand*, Maynooth.

- [19] S. Avdakovic, E. Becirovic, N. Hasanpahic, M. Music, A. Merzic, A. Tuchic, J. Karadza, D. Pesut y A. Kinderman Loncarevic, «Long-term forecasting of energy, electricity and active power demanda - Bosnida and Herzegovina case study,» *Balkan Journal of Electrical & Computer Engineering*, vol. 3, nº 1, 2015.
- [20] L. P. H. Orbea, Pronóstico de la demanda utilizando inteligencia artificial, Quito: Escuela Politecnica Nacional, Julio 2007.
- [21] J. R. Berrezueta Suárez y A. P. Encalda Serrano, *Análisis de los factores que afectan la demanda de energía eléctrica y su estimación. sector residencial del área de concesión de la Empresa Eléctrica Regional Centro Sur C.A. período 2002-2012*, Cuenca, 2014.
- [22] A. M. Ariza Ramírez, *Métodos utilizados para el pronóstico de demanda de energía eléctrica en sistemas de distribución*, Pereira, 2013.
- [23] A.-B. Mostafa Imbarek Ihbal, *Investigation of Energy Demand Modeling and Managment for Local Communities*, 2012.
- [24] T. H. Fong, «Case study of short-term electricity load forecasting with temperature dependency,» octubre 2009.
- [25] Instituto Nacional de Estadística y Censos, «Proyección provincias, Sexos y Áreas 2010-2020,» [En línea]. Available: <http://www.ecuadorencifras.gob.ec/proyecciones-poblacionales/>.
- [26] M. Matijas, *Electric load forecasting using multivariate meta-learning*, Zagreb, 2013.
- [27] INEC, Censo Económico 2010, Quito: INEC, 2010.
- [28] B. U. Islam, «Comparison of Conventional and Modern Load Forecasting Techniques Based on Artificial Intelligencia and Expert Systems,» *IJCSI/ International Journal of Computer Science Issues*, vol. 8, nº 3, septiembre 2011.
- [29] M. Usman Fahad y N. Arbab, «Factor Affecting Short Term Load Forecasting,» *Journal of Clean Energy Technologies*, vol. 2, nº 4, octubre 2014.
- [30] F. Naimi, «Java Based Load Forecasting for a Small Number of Public Properties,» 20 marzo 2008.
- [31] E. E. Ramírez Jiménez, «Pronostico de demanda de energía eléctrica utilizando regresión lineal,» 12 diciembre 2008.
- [32] Profesores, Material Docente de Econometría, Universidad de Valladolid, 2011-2012.
- [33] N. A., Análisis de Regresión, Departamento de Economía Cuantitativa, Septiembre de 2010.
- [34] U. Ezequiel, Regresión Lineal Múltiple, Universidad de Valencia, 2013.
- [35] R. Mahendra Gor, «Forecasting Techniques,» de *Industrial Statistics and Operational Managment*.
- [36] R. O. J., El método Estadístico.

- [37] K. Werner, D. Brandon y S. Shumate, Probabilistic Forecast Distribution Verification Primer, 2003.
- [38] L. F. Arias Roman, Programa digital para el pronóstico de carga eléctrica espacial, 2000.
- [39] J. S. Armostrong, Principles of forecasting, Kluwer Academic Publishers, 2001.
- [40] J. Murillos S, A. Trejos y P. Carvajal Olaya, «Estudio del pronóstico de la demanda de energía eléctrica, utilizando modelos de series de tiempo,» *Scientia et Technica*, diciembre 2003.
- [41] S. Amakali, *Development of models for shor-term load forecasting using artificial neural networks*, 2008.
- [42] B. M. d. Brio y A. Sanz Molina, Redes Neuronales y Sistemas Difusos, Alfaomega, 1997.
- [43] D. I. Cabrera Calopiña, *Diseño de una red neuronal artificial para la predicción de la demanda eléctrica*, Loja, 2014.
- [44] M. D., Redes Neuronales: Conceptos Básicos y Aplicaciones, GIAIQ, Marzo 2001.
- [45] G. T. Cerna, «Sistemas Inteligentes,,» Domingo Julio 2016. [En línea]. Available: <http://sistemasinteligente-redesneuronales.blogspot.com/>.
- [46] B. F, Entrenamiento de Redes Neuronales basado en algoritmos evolutivos, Buenos Aires: Universidad de Buenos Aires, Noviembre 2005.
- [47] B. Krose y P. Van der Smagt, An introduction to Neural Networks, octava ed., 1996.
- [48] C. B. s. R. Neuronales, «Conceptos Básicos sobre Redes Neuronales,» [En línea]. Available: <http://grupo.us.es/gtocom/pid/pid10/RedesNeuronales.htm>.
- [49] D. Costarelli, *Sigmoidal Functions Approximation and Applications*, 2014.
- [50] L. P., I. Inza y A. Moujahib, Departamento de Ciencias de la Computación e inteligencia Artificial-Redes Neuronales, Universidad del País Vasco.
- [51] X. B. Olabe, Redes Neuronales y sus Aplicaciones, Dpto. Ingeniería de Sistemas y Automática Escuela Superior de Ingeniería de Bilbao.
- [52] L. Marulanda, *CONTROL POR MEDIO DE REDES NEURONALES SISTEMAS AVANZADOS DE CONTROL*, Santiago de Cali, 2010.
- [53] W. Duch y N. Jankowski, *Survey of Neural Transfer Functions*.
- [54] R. Rojas, *Statiscs and Neural Networks*, Berlín, 1996.
- [55] J. Bullinaria, *Improving Generalization*, 2004.
- [56] S. Khaled Safi, «Artificial neural networks approach to time series forecasting for electricity consumption in Gaza Strip,» *IUG Journal of Natural and Engineering Studies*, vol. 21, n° 2, pp. 1-22, 2013.
- [57] C. Aggarwal, *Outlier analysis*, Nueva York.

- [58] J. L. Orellana Romero, *Modelación y pronóstico de la demanda de energía eléctrica de mediano plazo de El Salvador*, Ciudad Universitaria, 2012.
- [59] D. R. Anderson, D. J. Sweeney y T. A. Williams, *Estadística para administración y economía*, décima ed., México, D.F.: CENGAGE Learning.
- [60] Coordinación General del Sistema de Información Nacional, *Boletín de Precipitación y Temperatura*, Riobamba, 2015.
- [61] Comisión Interamericana para el Control del Abuso de Drogas, *Cantón de Riobamba*.
- [62] Subgerencia de Regulación y Programación-Banco Central del Ecuador, [En línea]. Available: <https://www.bce.fin.ec/index.php/component/k2/item/293-cuentas-provinciales>.
- [63] «Electric load forecasting using an artificial neural network,» *IEEE Transactions on Power Systems*, vol. 6, n° 2, pp. 442-449, 1991.
- [64] A.-A. R. E., *Univariate modeling and forecasting of monthly energy demand time series using abductive and neural networks*.
- [65] D. C. Tamal y G. Indranil, «Artificial Neural Network and Time Series Modeling Based Approach to Forecasting the Exchange Rate in a Multivariate Framework,» *Journal of Insurance and Financial Management*, vol. 1, n° 5, pp. 92-123, 2016.
- [66] T. Koskela, M. Lehtokangas, J. Saarinen y K. Kaski, *Time Series Prediction with Multilayer Perceptron FIR and Elman Neural Networks*, Tampere.
- [67] K. Neeraj y K. J. Govind, «A Time Series ANN Approach for Weather Forecasting,» *International Journal of Control Theory and Computer Modeling*, vol. 3, n° 1, enero 2013.
- [68] S. Benkachcha, J. Benhra y H. El Hassani, «Seasonal Time Series Forecasting Models based on Artificial Neural Network,» *International Journal of Computer Applications*, vol. 116, n° 20, abril 2015.
- [69] N. Morariu, E. Iancu y S. Vlad, «A Neural Network Model for Time-series Forecasting,» *Romanian Journal of Economic Forecasting*, vol. 4, 2009.
- [70] J. J. Montaña Moreno, A. Palmer Pol y P. Muñoz Gracia, «Artificial neural networks applied to forecasting time series,» *Psicithema*, vol. 23, n° 2, pp. 322-329, 2011.
- [71] A. Raudys y J. Mockus, «Comparison of ARMA and Multilayer Perceptron Bases Methods for Economic Time Series Forecasting,» *Informatika*, vol. 10, n° 2, pp. 231-244, 1999.
- [72] I. Kaastra y M. Boyd, «Designing a neural network for forecasting financial and economic time series,» *Neurocomputing*, vol. 10, pp. 215-236, 1996.
- [73] S. Param, *Electricity demand prediction using artificial neural network framework*, Fargo, 2015.

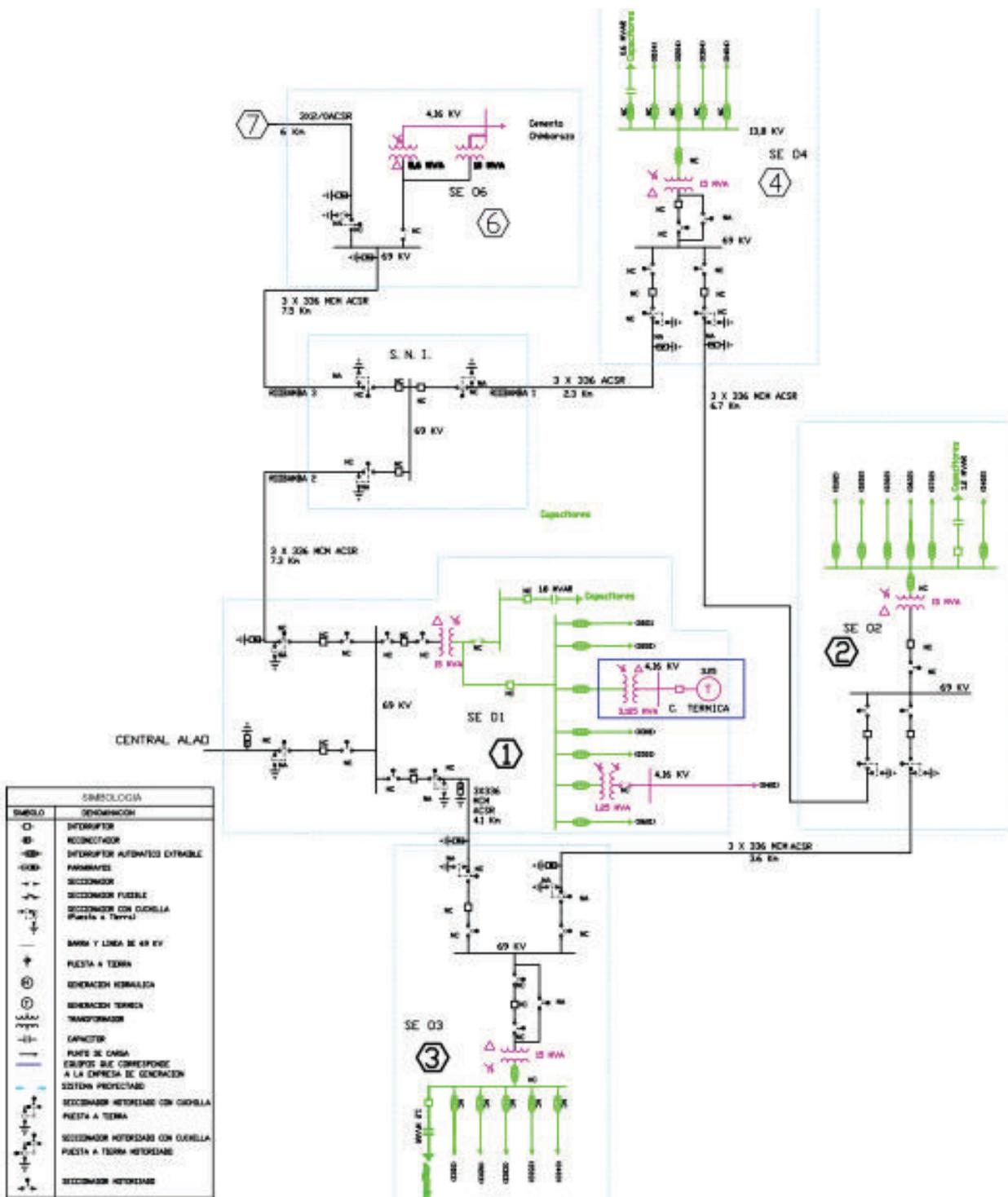
- [74] H. Abolfazli, S. Masoud, N. Nazari-Shirkouhi, S. Asadzadeh y K. Rezaie, «Forecasting Rail Transport Petroleum Consumption Using an Integrated Model of Autocorrelation Functions-Artificial Neural Network,» *Acta Polytechnica Hungarica*, vol. 11, nº 2, 2014.
- [75] D. Gerbing, *Autocorrelation*, 2016.
- [76] Z. Tang y P. A. Fishwick, *Feed-forward Neural Nets as Models for Time Series Forecasting*.
- [77] NCSS Statistical Software, *Correlation*.
- [78] R. Beaumont, *An Introduction to statistics Correlation*, 2012.
- [79] J. A. Alvarez Jiménez, *Estudio e implementación del sistema de protecciones eléctricas de los alimentadores primarios de la subestación La Carolina de EMELNORTE S.A.*, Quito, 2013.
- [80] J. A. Mauricio, *Análisis de series temporales*.
- [81] F. Rodrigues, C. Cardeira y J. Calado, «The daily and hourly energy consumption and load forecasting using artificial neural network method: a case study using a set of 93 households in Portugal,» de *The Mediterranean Green Energy Forum*, 2014.
- [82] C. J. Cruz Tenempaguay, *Diseño de la coordinación del sistema de protecciones eléctricas en los alimentadores interconectados urbanos de la EERRSA*, Quito, 2015.
- [83] Unidad de planeación minero energética, «Proyección de demanda de energía eléctrica y potencia,» Bogotá D.C., 2008.
- [84] F. Gunther y S. Fritsch, *neuralnet: Training of Neural Networks*.
- [85] M. J. N. Sánchez, Auditoría administrativa a la jefatura de recursos humanos de la Empresa Eléctrica Riobamba S.A. año 2010, Riobamba: Escuela Superior Politécnica de Chimborazo, 2011.
- [86] A. B. S. Fonseca y P. A. Brasales Panchi, Diseño de un modelo de comunicación organizacional y marketing para fortalecer la imagen de la Empresa Eléctrica Riobamba S.A., periodo 2012-2013, Riobamba: Escuela Superior Politécnica de Chimborazo, 2013.
- [87] B. S. O. Topón, Modelo de gestión Financiera y Reducción de morosidad en el Departamento Financiero de la Empresa Eléctrica de Riobamba S.A., Riobamba: Escuela Superior Politécnica de Chimborazo, 2011.
- [88] C. E. V. Guerra, Medición y Evaluación en Comunicación Estratégica en la Empresa Eléctrica Riobamba S.A., Loja: Universidad Técnica Particular de Loja, febrero 2015.
- [89] J. Palacios, Análisis climático del Ecuador Continental durante el año 2013 y su Impacto en el sector agrícola, Quito: Estudios e Investigaciones, Enero 2014.
- [90] B. S. Y., *Discrete Multivariate Analysis: Theory and practice*.

[91] R. Hyndman, Forecasting: Principles & Practice, 2014.

[92] A. Callegaro, Forecasting methods for spare parts demand, 2010.

ANEXO I

Diagrama Unifilar del Sistema Eléctrico de EERSA



## ANEXO II

	<b>2008</b>	<b>2009</b>	<b>2010</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016</b>
<b>enero</b>	12474	12821	14818	15378	15831	18125	20868	21725	25522
<b>febrero</b>	12492	11741	12276	14880	14800	15747	20682	20067	22467
<b>marzo</b>	12504	14260	14120	17903	15955	17650	20433	21569	21614
<b>abril</b>	12857	13288	13251	13787	15704	17793	19369	21241	21396
<b>mayo</b>	13140	14284	12835	15029	16146	19059	19918	22443	22220
<b>junio</b>	13165	13290	13606	14495	15913	17693	19353	21196	20732
<b>julio</b>	13598	14096	13759	14688	16040	19857	21329	21812	22296
<b>agosto</b>	14086	12716	12931	14650	14650	20045	19785	21656	22716
<b>septiembre</b>	14096	12567	13484	14650	16030	24669	20863	22160	21913
<b>octubre</b>	14206	13492	14660	16023	17118	19920	20815	22862	25231
<b>noviembre</b>	14333	12626	14795	16841	18236	21897	21840	21902	26674
<b>diciembre</b>	14692	13690	15156	16681	18889	22296	22748	23193	27498

## ANEXO III

		PIB (Miles de \$)	Valor Normalizado
2008	C 1	805360	0,0000
	C 2	846643	0,0307
	C 3	887927	0,0615
2009	C 1	929211	0,0922
	C 2	969035	0,1219
	C 3	1008858	0,1516
2010	C 1	1048682	0,1812
	C 2	1063722	0,1924
	C 3	1078762	0,2036
2011	C 1	1093801	0,2148
	C 2	1158414	0,263
	C 3	1223026	0,3111
2012	C 1	1287639	0,3592
	C 2	1299776	0,3682
	C 3	1311914	0,3773
2013	C 1	1324051	0,3863
	C 2	1356119	0,4102
	C 3	1388186	0,4341
2014	C 1	1420254	0,458
	C 2	1503547	0,52
	C 3	1586839	0,582
2015	C 1	1670132	0,6441
	C 2	1716413	0,6786
	C 3	1762694	0,713
2016	C 1	1808975	0,7475
	C 2	1873978	0,7959
	C 3	1938980	0,8443
2017	C 1	2003983	0,8927
	C 2	2075992	0,9464
	C 3	2148002	1,0000

## Anexo IV

### Neuroph Studio

El Neuroph Studio es un framework basado en Java para el desarrollo de redes neuronales artificiales.

#### **Procedimiento para emplear el Neuroph Studio**

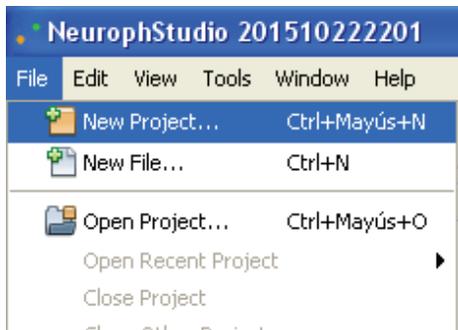
El programa computacional cuenta con un instalador guiado que facilita una instalación rápida y sencilla bajo ambiente Windows o Linux. Posterior a la instalación existen 5 pasos básicos para el desarrollo de una RNA.

1. Crear el proyecto
2. Crear la red neuronal artificial
3. Crear los conjuntos de entrenamiento (training set)
4. Entrenar la red
5. Probar la red ya entrenada

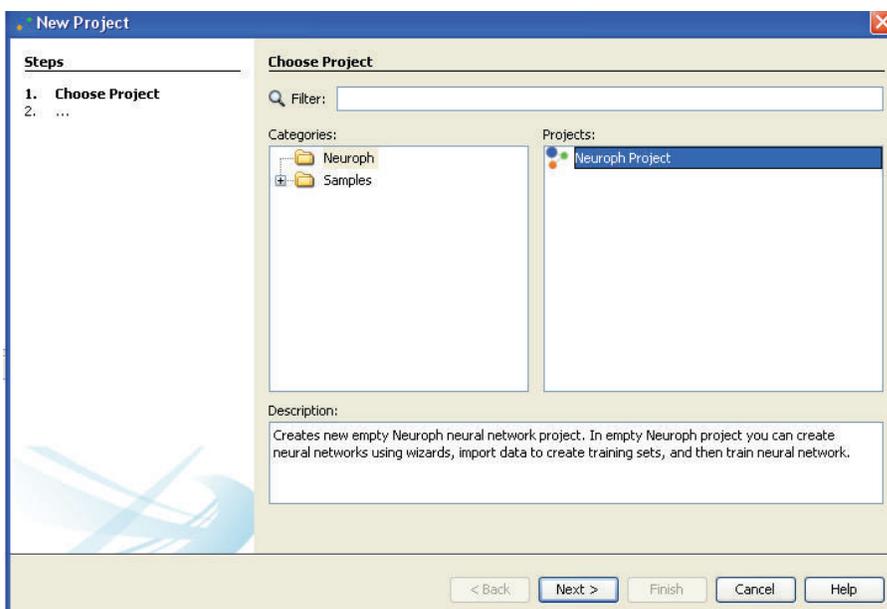
A continuación se detallan los pasos mencionados

1. Crear el proyecto

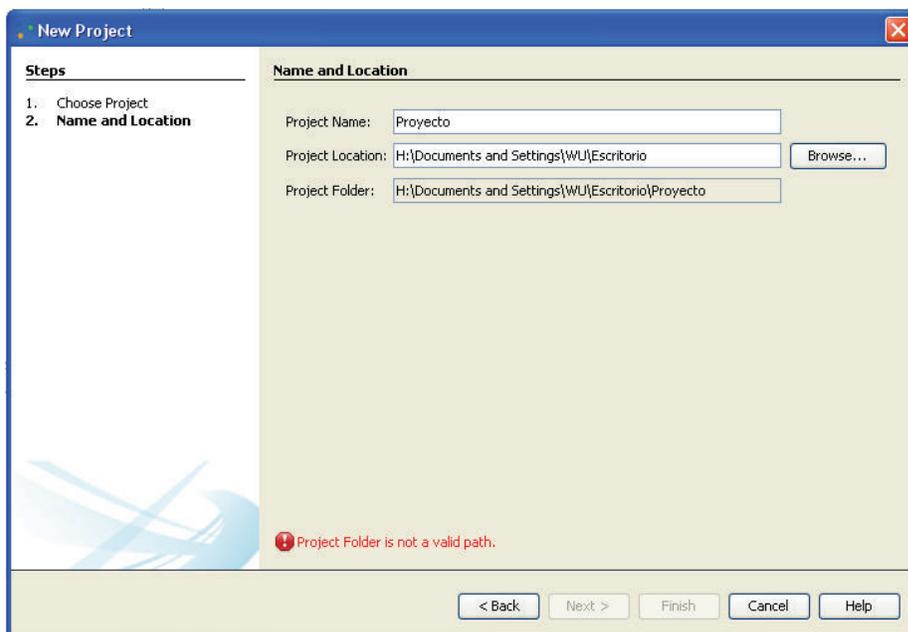
Click Archivo (*File*) > Nuevo Proyecto (*New Project*).



Posteriormente se debe seleccionar Neuroph Project para crear un proyecto nuevo y hacer clic en Siguiente (*Next*)

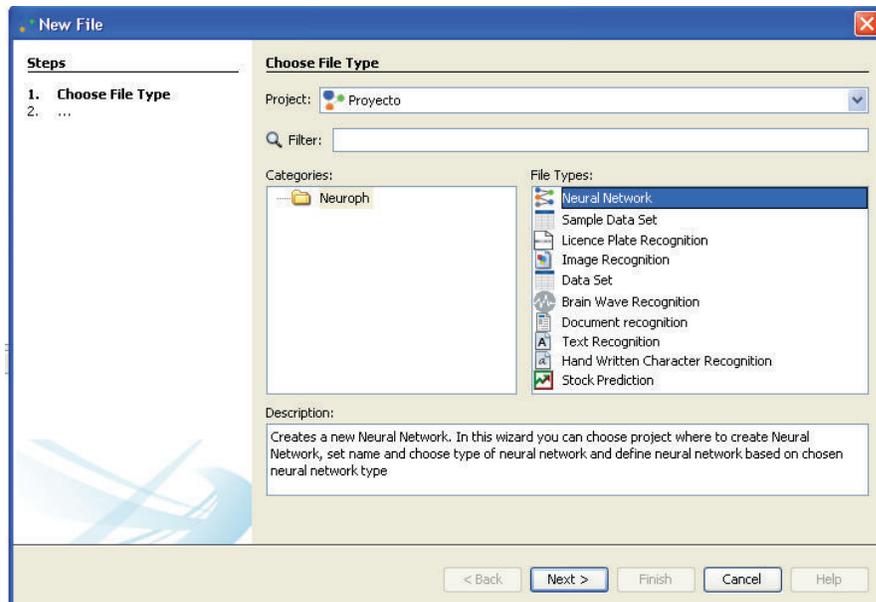


Luego se debe seleccionar un nombre para el proyecto y escoger la ubicación donde se guardará el mismo.

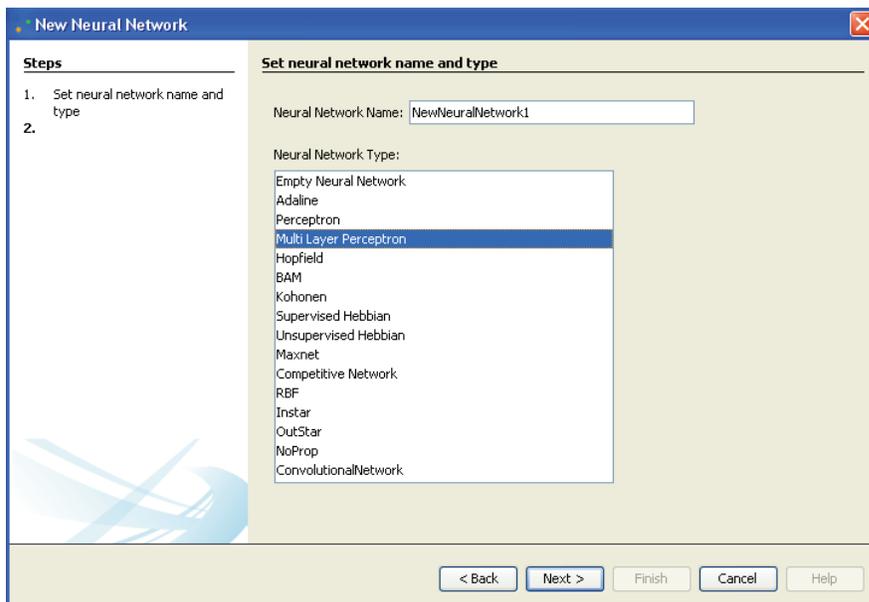


## 2. Crear la red neuronal artificial

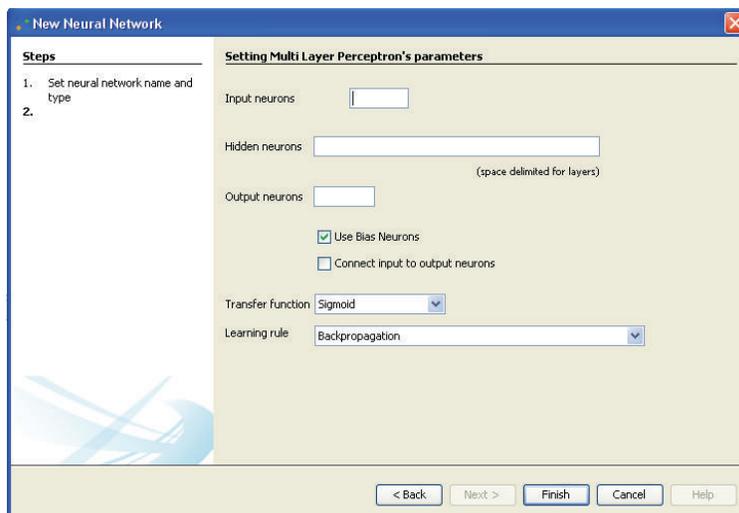
Hacer clic Archivo (File) > Nuevo archivo (New File) y seleccionar Red Neuronal (Neural Network) como tipo de archivo



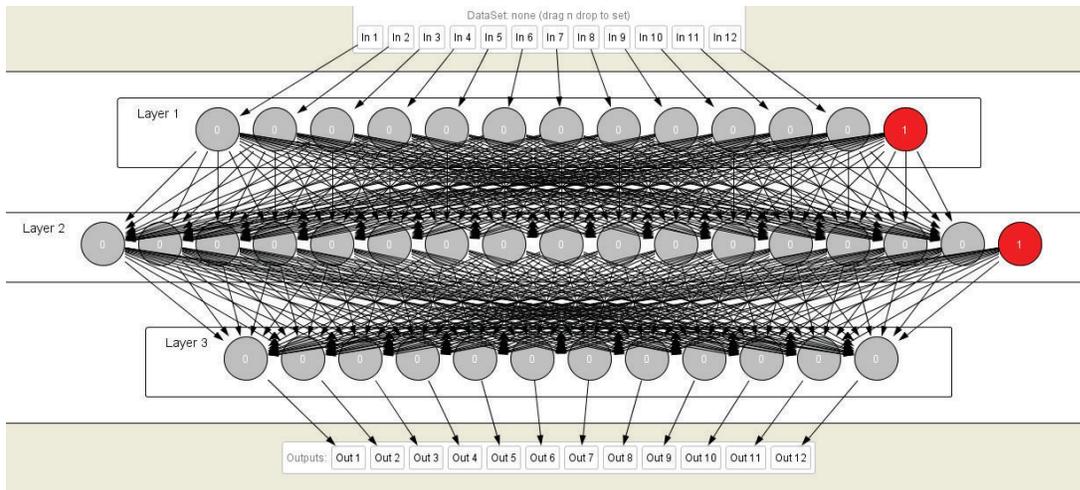
Se debe introducir un nombre para la RNA y seleccionar el tipo de red, en el caso de esta investigación se empleó el Perceptron Multicapa (Multi Layer Perceptron).



En la siguiente ventana se debe establecer la topología de la red (neuronas de entrada, ocultas y de salida así como la función de transferencia y regla de aprendizaje)

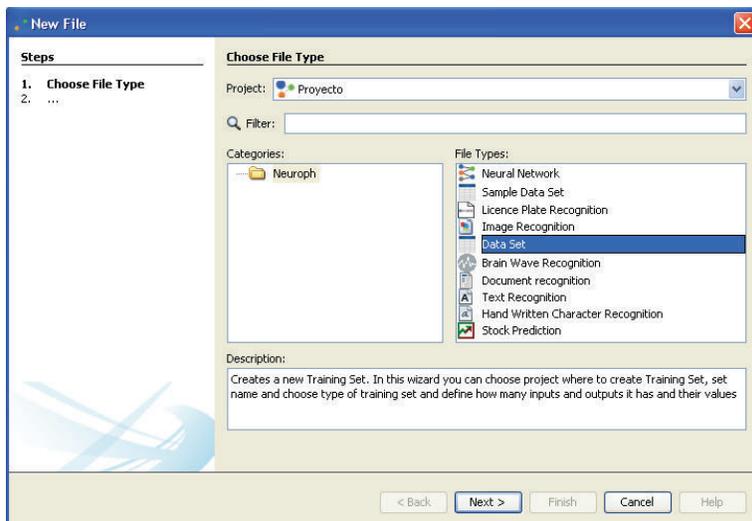


Al hacer clic en finalizar (Finish) se crea la RNA de manera similar a la siguiente.



### 3. Crear los conjuntos de entrenamiento (training set)

Hacer clic Archivo (File) > Nuevo archivo (New File) y seleccionar Conjunto de datos (Data Set) como tipo de archivo. Hacer clic en siguiente (Next).



Posteriormente se debe seleccionar el nombre del data set, tipo (supervisado o no supervisado), número de entradas, salidas y la ubicación del archivo origen donde se encuentran los datos delimitados. Hacer clic en finalizar (Finish).

**New Data Set**

**Steps**

1. Choose File Type
2. **Set data set name, type and number of inputs and outputs**

**Set data set name, type and number of inputs and outputs**

Data set name:

Type:

Number of inputs:

Number of outputs:

Load from file

File:

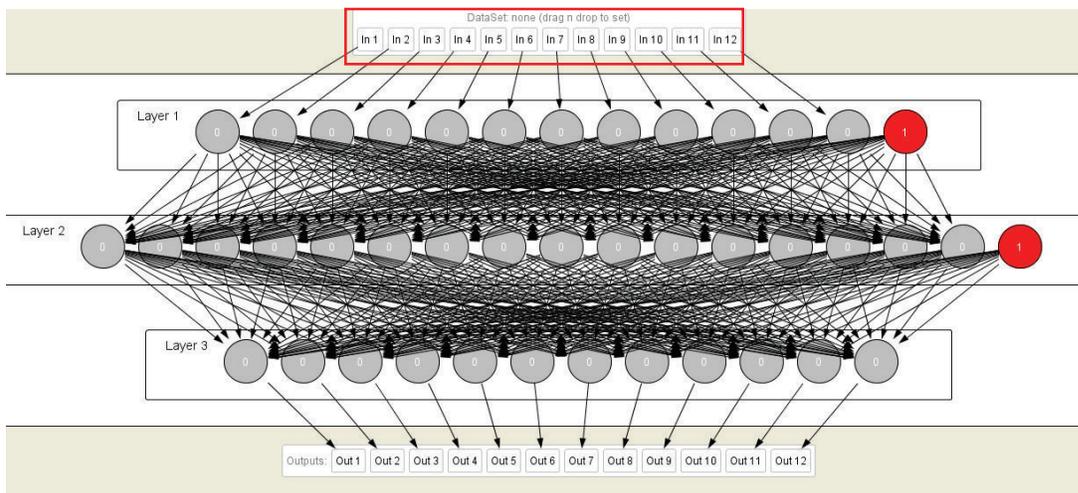
Delimiter:

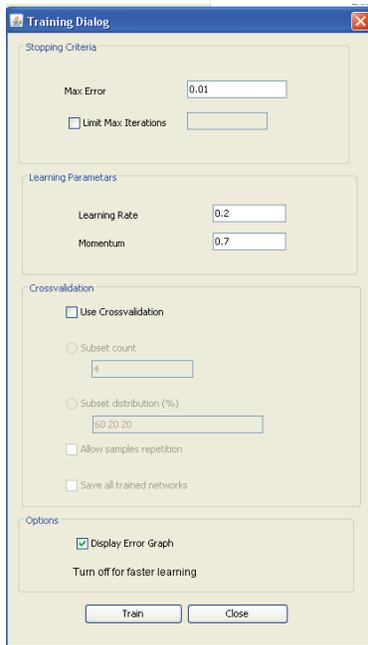
First Line Contains Column Names

< Back   Next >   **Finish**   Cancel   Help

#### 4. Entrenar la red

Se debe seleccionar el data set creado hasta el área marcada en la siguiente imagen (data set de la capa de entrada).





The image shows a 'Training Dialog' window with the following sections and settings:

- Stopping Criteria:**
  - Max Error: 0.01
  - Limit Max. Iterations
- Learning Parameters:**
  - Learning Rate: 0.2
  - Momentum: 0.7
- Crossvalidation:**
  - Use Crossvalidation
  - Subset count: 4
  - Subset distribution (%): 60 20 20
  - Allow samples repetition
  - Save all trained networks
- Options:**
  - Display Error Graph
  - Turn off for faster learning

Buttons: Train, Close

El ingreso de los datos se realiza en una hoja dividida por columnas y filas (hoja de cálculo) considerándose que cada columna representará una variable en los casos de análisis.

## Anexo V

### NCSS 11

The screenshot displays the NCSS 11 software interface. At the top, there is a menu bar with options: File, Edit, View, Data, Analysis, Graphics, Tools, Window, Help. Below the menu bar is a toolbar with various icons for file operations (New, Open, Last, Save), data manipulation (Sort, Fill, Trans, Filter), and other functions (Filter Active, Hide Filtered Rows, Procedures, Favorites, Recent, Loaded, Output, Gallery). A 'Purchase' button is also visible in the top right corner.

The main window shows a data table with 18 columns (C1 to C18) and 33 rows (1 to 33). The 'Column Info' pane is open on the left, showing the following information for each column:

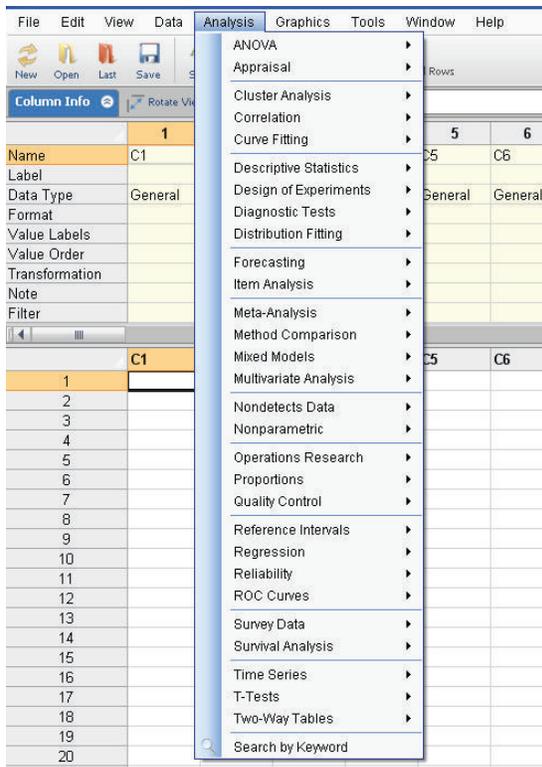
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
Name	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18
Label																		
Data Type	General																	
Format																		
Value Labels																		
Value Order																		
Transformation																		
Note																		
Filter																		

The data table below shows the following structure:

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16	C17	C18
1																		
2																		
3																		
4																		
5																		
6																		
7																		
8																		
9																		
10																		
11																		
12																		
13																		
14																		
15																		
16																		
17																		
18																		
19																		
20																		
21																		
22																		
23																		
24																		
25																		
26																		
27																		
28																		
29																		
30																		
31																		
32																		
33																		

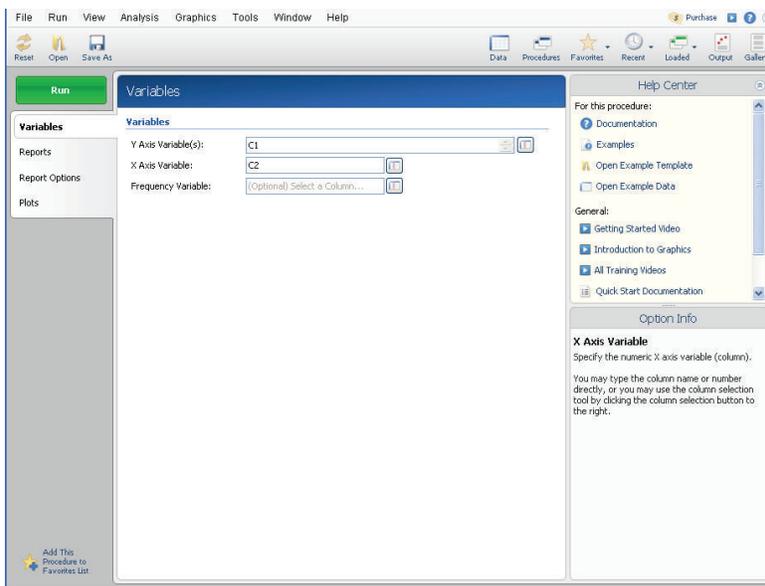
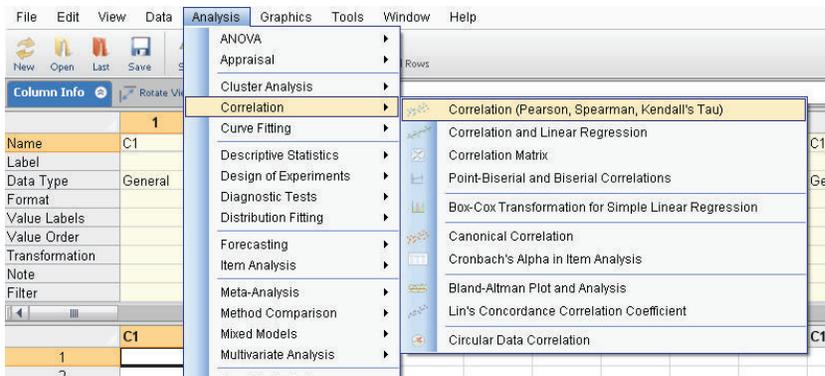
The status bar at the bottom indicates 'Row: 1 | Col: 1' and a zoom level of '100%'.

En la barra de herramientas al hacer clic en Analisis (Analysis) se despliega un menú donde se encuentran todas las herramientas disponibles.



## Correlación

Al seleccionar correlación según se muestra en la siguiente figura es posible luego introducir el par de variables para obtener la correlación seleccionando los ejes x, y según corresponda al caso.



Al hacer click en ejecutar (Run) se obtiene los resultados de la correlación similar al ejemplo mostrado.

## Correlation Report

Dataset: Untitled  
 Y Axis Variable: C1; X Axis Variable: C2

## Run Summary Section

Parameter	Value	Parameter	Value
Y Axis Variable	C1	Rows Processed	9
X Axis Variable	C2	Rows Used in Estimation	9
Frequency Variable	None	Rows with X Missing	0
Sum of Frequencies	9	Rows with Freq Missing	0

## Column Summary Section

Variable	Count	Mean	Standard Deviation	Minimum	Maximum
C1	9	5,22	3,07	1,00	10,00
C2	9	4,78	1,92	2,00	8,00

## Pearson Correlation Confidence Interval Section

Two-Sided Confidence Interval of  $\rho$

Pearson Correlation	Count	R Distribution 95% Confidence Limits		Normal Approximation 95% Confidence Limits	
		Lower	Upper	Lower	Upper
0,9194	9	0,6301	0,9792	0,6554	0,9832

## Pearson Correlation Test Section

H0:  $\rho = 0$

Alternative Hypothesis	Pearson Correlation	Count	df	T-Value	P-Value	Reject H0 at $\alpha = 0,05?$
$\rho \neq 0$	0,9194	9	7	6,1824	0,0005	Yes

